Jigsaw: Authoring Immersive Storytelling Experiences with Augmented Reality and Internet of Things

ANONYMOUS AUTHOR(S)*

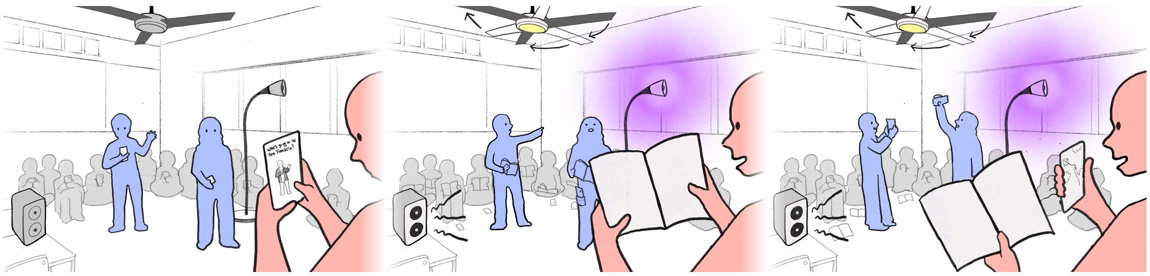


Fig. 1. An example story of Benjamin Franklin's Kite Experiment from Jigsaw that combines AR and IoT devices into one immersive experience: (a) The participants can assign themselves a character by waving. (b) As the narrator reads the story, trigger words from the narration enact changes in the physical environment (shown with a smart light, smart fan, and smart speaker). (c) Virtual effects or objects are shown in the AR view, such as a kite, clouds, and sparks.

Augmented Reality (AR) provides new opportunities for telling stories in more immersive ways. Yet, the utilization of AR for immersive storytelling encounters two primary challenges. First, the immersivenesss in AR is typically limited to visual augmentations, e.g., pixels displayed on a screen. Second, the process of creating immersive stories remains inaccessible to non-experts, owing to the requisite proficiency in advanced skills. We introduce Jigsaw, a system that enables novices to both consume and create immersive stories that harness virtual and physical augmentations. Jigsaw achieves this through the novel fusion of mobile AR with off-the-shelf Internet-of-things (IoT) devices. We evaluated the consumption and creation of immersive stories through a qualitative user study with 20 participants, and found that end-users were able to create immersive stories and felt highly engaged in the playback of three immersive stories. Nonetheless, sensory overload was one of the most notable challenges across all experiences. We discuss design trade-offs and considerations for future endeavors in immersive storytelling involving AR and IoT.

CCS Concepts: • **Human-centered computing → Interactive systems and tools**.

Additional Key Words and Phrases: storytelling, augmented reality, internet-of-things, authoring tool

ACM Reference Format:

Anonymous Author(s). 2023. Jigsaw: Authoring Immersive Storytelling Experiences with Augmented Reality and Internet of Things. 1, 1 (September 2023), 16 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

With the development of computational devices, the media for storytelling evolved from analog forms, such as printed books and verbal narration, to digital forms, such as images, videos, and animations. One important aspect of such an evolution is that storytellers often aim to offer more immersive and engaging modalities for their audiences.As an emerging technology, augmented Reality (AR) possesses the merits of better immersion by displaying 3D content spatially and engagement by enabling richer interactions such as direct manipulation of 3D content. AR thus opens up vast opportunities for richer storytelling experiences at the consumer level. For example, storytelling apps such as Wonderscope [45] can tell stories in AR. However, the immersion of these experiences is limited to the field of view of the computational device (e.g., pixels on the screen).

Much effort has been made to make storytelling experiences more immersive and engaging in theatres and performing arts. Adding multiple senses to the storytelling experiences has been explored to achieve better immersion. For instance, "4D" movies in immersive theaters and amusement parks can make storytelling experiences more immersive by adding not only visual and auditory senses but also haptic and smelling senses, etc. On the other hand, more interactivity and participation have been included in storytelling experiences in order to achieve better engagement. Immersive plays such as Sleep No More have been designed for the audience to better engage with the story by allowing them to move around in the physical environment and interact with the actors.

Yet, these immersive plays have a high entry barrier in both the creation and the consumption of stories, as they require a specialized setup and can only be consumed at specific times and locations. To make consumer-level AR storytelling more immersive and engaging, we incorporate typical household Internet of Things (IoT) devices, such as smart lights and smart fans, into AR storytelling experiences. While AR overlays 3D virtual content to the physical world, IoT devices complement this by acting as physical overlays and adding additional senses to the world. Research democratizing AR+IoT storytelling has enabled users to make custom-make IoT devices and event-based stories, which still poses a high barrier to novices as it requires knowledge of building hardware and block-based programming [20].

In this paper, we aim to explore storytelling experiences that are composed of AR and off-the-shelf IoT devices and to lower the barrier of creating these storytelling experiences for end-users. We designed and evaluated Jigsaw, an AR application on mobile phones that enable people with little to no technical background to easily author and experience AR+IoT stories. To lower the entry barrier for end-users, we draw inspiration from slide-based presentation tools, such as Powerpoint: our authoring mode allows users to create multiple scenes similar to multiple slides. Users can define AR content within each scene, such as 3D models and animations, and IoT behaviors, such as dimming the smart light or making the smart fan blow harder. To enhance the audience's engagement, users can also define triggers such as a spoken keyword or a touch interaction that helps transit from the previous scene to the next. The whole story can then be played for multiple users by going through the scenes sequentially.

This paper describes the design and implementation of Jigsaw, including the authoring mode and three AR+IoT storytelling experiences. We present a design space of AR+IoT stories and demonstrate the expressiveness of the authoring mode by creating three stories focused on different aspects of the design space. We also conducted a qualitative user study of Jigsaw with 20 participants (10 groups), in which they found these stories to be immersive, engaging, and memorable to experiences, and the authoring mode was easy to use and capable of creating a wide variety of stories. The study also revealed trade-offs of several design aspects in this new genre of storytelling experiences. The contributions of this paper, therefore, include 1) three immersive stories that explore the design space of storytelling experiences composed of AR and off-the-shelf IoT devices; 2) an authoring tool that lowers the floor of creating AR+IoT stories and is generalizable for creating various stories; 3) a user study that reveals the benefits and challenges of both the creation and consumption of AR+IoT stories.

2 BACKGROUND AND RELATED WORK

2.1 Immersive Storytelling

Immersion is an important aspect in any storytelling to engage the audience and to convey the emotion of the performance. Therefore, many fields of performing arts have adopted novel forms of performance to enhance the immersion. Immersive theater is a prominent example "which use installations and expansive environments, which have mobile audiences, and which invite audience participation" [44]. It can involve great levels of intimacy [21–23] or put more distinction between the audience and the actors [37, 38]. The experience also involves multiple senses such as the sense of space, touch, and smell [44]. Sleep No More, an immersive theater performance based on Shakespeare's Macbeth, enables the audience to be embedded within the performance, inviting them to walk around the stage and touch the props [1]. While there are some examples of technology use in this type of performance, e.g., headphones and speakers in Rotozaza's Autoteatro [24], it is rather limited, with many available technologies like projection mapping, AR, and robotics rarely being used if ever. We envision that technology has a great potential to enhance immersion and give a sense of magical events occurring in the world around us, as seen in literary genres like magical realism [11, 34].

Outside of theater, we see more frequent use of technology in storytelling. For instance, the Weather Channel frequently uses AR in their weather forecast programs [19]. Additionally, theme parks, such as Disneyland, use projection mapping, robots, and mechanically-moving rides to foster a high sense of immersion in the storytelling of their rides [6, 7]. Projection mapping, in particular, is a popular method for AR to convert a room or a building into a magical space filled with eye-catching visualizations (e.g., [3, 41, 43]). However, these storytelling experiences require a team of expert designers and engineers to configure, along with an expensive set of equipment. Therefore, they are difficult to access for most people and can only be experienced in special venues. In contrast, our work aims to democratize immersive storytelling to everyday users.

The emergence of AR-enabled mobile devices created a platform to tackle this issue. Wonderscope [45] is a prime example that uses mobile AR to recreate immersive theaters for children in their own rooms. Using this, children can be immersed in the magical storytelling experience, presented as ready-made narratives. SceneAR achieves a similar immersive storytelling experience, but adds the ability to remix the stories [13]. There are also examples of AR use for educational purposes, such as teaching chemistry [16]. While these systems provides a starting point and an inspiration for our work, they are largely limited to augmenting the world with pixels on a screen. This fundamentally limits the level of immersion compared to previous examples of immersive theater and higher-fidelity experiences that give a sense of magic throughout the entire surroundings. Rather, we want to incorporate both virtual and physical augmentations ("pixels and atoms") to foster a more complete immersion. Therefore, we examine the current landscape of ubiquitous computing and the possibility of using smart IoT devices as a latent infrastructure of immersive storytelling.

2.2 Ubiquitous Computing and Physical Augmentation

With the growing popularity of smart homes and IoT devices, ubiquitous computing is becoming a part of everyone's day-to-day life. Part of this effort aims to bring virtual augmentation to the world around us. For instance, RoomAlive [27] and Room2Room [36] achieve roomscale projection mapping with commercially-available devices such as consumer-grade projectors and Microsoft Kinect. Moreover, Olwal and Dementyev [35] showcase an ambient display hidden under common surfaces that can be seen in household furnitures. As for use cases that focus on storytelling, Healey et al. [25] demonstrates that mixed-reality systems can enable remote storytelling and that it helps promote engagement among children.

However, as discussed above, we wish to go beyond the pixel-based augmentation, and leverage devices that physically augments the environment. Examples of this include smart lights, smart fans, speakers, thermostats, vacuum robots, and wearable devices like a smart watch or smart glasses. Additionally, we want to seamlessly integrate the virtual and physical augmentations to create a holistic narrative. There are previous works that leverage conversational agents on smart speakers [14, 33], robots [26, 29], and tangible interfaces [42] to foster effective storytelling with physical augmentation, but they largely focus on one or a handful of IoT device types, and often do not incorporate virtual augmentation. Lin et al. [33] showcase a more comprehensive, low-cost system that recreates 4D theater experience using home appliances like the ones mentioned before. However, the system nonetheless do not incorporate immersive virtual augmentations like mobile AR or AR glasses, and the system's authoring environment is not suited for novice users, which fail to meet our goal of enabling anyone to create immersive experiences at home. StoryMakAR [20] enables children to create an immersive storytelling experience using robots and mobile AR, which aims to achieve similar goals as our work, but still do not incorporate other physical devices like fans, lights, thermostats, etc. to create a truly immersive experience that transforms the entire surroundings. To envision how to empower anyone to author this type of experience, we examine existing tools with similar goals.

2.3 Authoring Tools

Much of the efforts of the HCI community focuses on how to make the creation of digital contents easy, which helped transform what was considered "expert-only" task of programming to something that anyone can do. Scratch [40] is a prime examples that uses block-based programming to empower children to author their own programs and games. Since then, many commercial applications have sprung up (e.g., graph-based programming in Lens Studio [4], Blender shader graph [5], Unreal Blueprints [2]) with similar goals in mind. This effort to achieve easy authoring of contents also naturally came to focus on AR applications as they emerged. Lee et al. [31] shows one of the first examples of this, where they enable immersive authoring os AR applications using tangible UI. FlowMatic [47] is a more recent work that uses virtual reality (VR) and graph-based programming for in-situ authoring of interactive scenes in immersive VR environment. SceneAR [13] specifically targets viewing and remixing of immersive storytelling experiences, and uses mobile AR for immersive authoring of these contents. While these works inspires the design of our authoring tool, they are limited to authoring virtual augmentations, whereas we aim to encompass physical augmentations as well.

In contrast, there are authoring tools that target physical IoT devices. Ivy [18] enables users to visually program smart devices and configure relational events between them in an immersive environment. Although we take inspiration from their spatial programming UI, our system avoids such potentially complicated process in creating physical augmentations from IoT devices. Other works demonstrate the use of conversational agents [28], and proxemic and gestural interactions [46] for authoring an IoT environment. All of these works, however, limit their target augmentation to only the physical. StoryMakAR [20] lets users author immersive stories that involve both physical devices and virtual objects, which is much closer to what we aim to achieve. However, as discussed previously, the system does not handle a variety of smart devices to provide a full immersive experience. It also requires a custom built robotics toolkit and does not leverage the existing IoT infrastructure that could unlock easily-accessible multi-sensory immersion.

Jigsaw allows anyone to create and experience immersive storytelling that uses both virtual and physical augmentation by leveraging commonly available IoT infrastructure and AR-enabled devices.

Table 1. Design considerations for each immersive story.

Story	IoT Augmentation	User Role	Interaction
Benjamin Franklin	Background	Narrators and actors	narration, touching, moving in the space
The Wind and The Sun	Foreground	Audience	tapping on the phone
Goodnight Moon	Foreground	Narrators and audience	narration

3 JIGSAW SYSTEM

We designed Jigsaw to enable novices to both consume and create immersive stories using a mix of mobile AR and off-the-shelf IoT devices. Our research process was divided into two steps: exploring *what* immersive stories are, and investigating *how* to create these immersive stories. We executed our research in this order because we believe it is necessary to understand what the possible resulting experiences are before building an authoring system for creating these experiences.

For the first step, we used two motivating questions to guide our design of immersive storytelling experiences: "how to seamlessly integrate both virtual and physical augmentation to these immersive stories?" and "what are the levels of participation for users in interacting with each other and with the stories?" We identified five design considerations for immersive stories that involve AR and IoT devices. Based on these design considerations, we then conducted a serious of brainstorming sessions to discuss potential stories via sketches and low-fidelity prototypes. Lastly, we prototyped three representative stories (i.e., Benjamin Franklin, The Wind and the Sun, and Goodnight Moon) that cover different aspects of the design considerations. The design considerations for each story are specified in Table 1. During the prototyping process, the research team met regularly to discuss design iterations and ran pilot tests frequently with end-users outside of our team using the prototypes.

For the second step, we designed and built an authoring system for creating these immersive stories. We have three design goals in mind: *(i)* to lower the barrier of entry for novices to create these stories, *(ii)* to make the system generalizable for creating various stories including the three representative stories, and *(iii)* to make the system consistent with existing interaction modalities in mobile AR applications. We drew inspiration from presentation tools such as Google Slides and enabled novices to create a linear storyline by creating a series of scenes. Within each scene, users can edit the triggers for the scene and the behaviors of AR content and IoT devices, inspiring by existing event-trigger models (e.g. as seen in [20]). Finally, users can play the story that they create in the same system.

We detail the design and implementation of the immersive stories and the authoring system in the following sections.

3.1 Immersive Stories

Inspired by the style of comic strips in storytelling and slide decks in presentation, we represent each immersive story using a sequential number of scenes. The story proceeds as the system transitions from the previous to the next scene. The transition of scenes can be triggered by either keyword detection (e.g., when the user says a keyword) or touch interaction (e.g., when the user's hand collides with an AR object in the physical world), similar to how the transition in Google slides is triggered by pressing a button on the mouse. Each scene contains states of AR content (e.g., appearing/disappearing of a 3D model) and IoT devices (e.g., turning on/off a smart device). We use three types of off-the-shelf IoT devices in these immersive stories: smart lights, smart fans, and smart speakers. In the below sections, we present the design considerations of each immersive story and how they are built using the *trigger-and-scene* framework.

3.1.1 Benjamin Franklin Kite Experiment. The Benjamin Franklin Kite Experiment is a story of Benjamin Franklin and his son taking a kite out during a storm to see if a key attached to the string would draw an electric charge. To play the Benjamin Franklin story, three users are required to be colocated with each user holding a phone. At the beginning of the experience, users coordinate according to the app and assign each user a different role: a narrator, an actor of Benjamin Franklin, and an actor of Benjamin Franklin Jr. After assigning the roles, users can see the two actors' bodies augmented with virtual costumes on them. The app then asks the narrator to narrate the story by reading a physical book out loud. As the narrator reads the story on the physical book, the app uses voice recognition to transcribe what the narrator says and automatically detect keywords in the transcription. A list of keywords that are selected based on the physical book and their corresponding scenes are predefined in the app. As one keyword is detected, the app will jump to its corresponding scene and wait for the next trigger. In this story, we defined a total of six pairs of triggers and scenes, with the last trigger being a touch event between users' hands and a virtual key in AR.

In this experience, we focused on using IoT devices as background augmentation, having human actors interacting with the story, and having a narrator reading from a physical book. The augmentation is considered as in the background since IoT devices add to the story's ambience, such as dimming the lights to simulate cloudy weather. The visualization of costumes and interaction of touching the key are designed to facilitate the sense of participation. Lastly, the design of reading a physical book is to increase the sense of physicality, where the the physical interactions of flipping pages and reading printed text can be meaningful in the experience.

3.1.2 The Wind and The Sun. The Wind and the Sun story is from the famous Aesop's Fables where the two characters quarrel about which of them is stronger. Any number of users can play the story with each user holding a phone. The story is narrated through the smart speaker. Users can tap on the screen to make the story proceed. In this story, the IoT devices act as the driving actors of the story with visual augmentations from the phone's AR, where the Wind is superimposed on the smart fan, and the Sun is on the smart light.

Compared to the previous story, this experience is similar to a 4D movie with less focus on interactivity. The users take a passive role, and the primary user who would have been the narrator before now only has control of moving to the subsequent scene.

3.1.3 Goodnight Moon. We recreated Goodnight Moon [12], an American children's story written by Margaret Wise Brown in 1947. The narrator greets objects around the room (e.g., "Goodnight, red balloon. Goodnight lamp. Goodnight fan"). The greeting causes an appropriate change in the greeted AR or IoT device (e.g., the balloon goes down; the smart lamp turns off; and the smart fan stops blowing wind). The story progresses with keyword detection as the narrator speaks, similar to the Benjamin Franklin story. However in this experience, the story is shown on the screen rather than on a physical book.

Because the narrator explicitly greets the IoT devices, the devices take on more of a foreground role. While the audience still has a passive role, the narrator is now an active, interactive participant in driving the story. The one-to-one mapping of each narration line with an effect on some object (AR or IoT) creates a simple but magical experience.

3.1.4 Implementation. We implemented the three stories using Lens Studio, which provides AR tracking, multi-user co-located experience, body pose detection and speech-to-text/text-to-speech capability. We connect and interact with the smart light using Kasa Smart's API [10]. We use an IR-controlled fan, along with Bond Bridge [8] for networked control of IR devices. Lastly, we use Amazon's Echo smart speaker [9] connected to the narrator's phone via Bluetooth.

3.2 Authoring System

Following the *trigger-and-scene* framework that was used to build different immersive stories, we introduce the authoring system of Jigsaw which allows novices with little to no technical background to create immersive stories with AR and IoT devices. We first present a system walkthrough of how to build the Goodnight Moon story. Then we introduce the our scene-based authoring approach and in-situ editing in AR. Lastly, we demonstrate the generalizability of our authoring system by replicating the Benjamin Franklin and the Wind and the Sun stories.

3.2.1 System Walkthrough: Building the Goodnight Moon Story. We present a system walkthrough involving a general end-user who has little to no technical background and is motivated to use Jigsaw to create the Goodnight Moon story that the user can play with their families and friends.

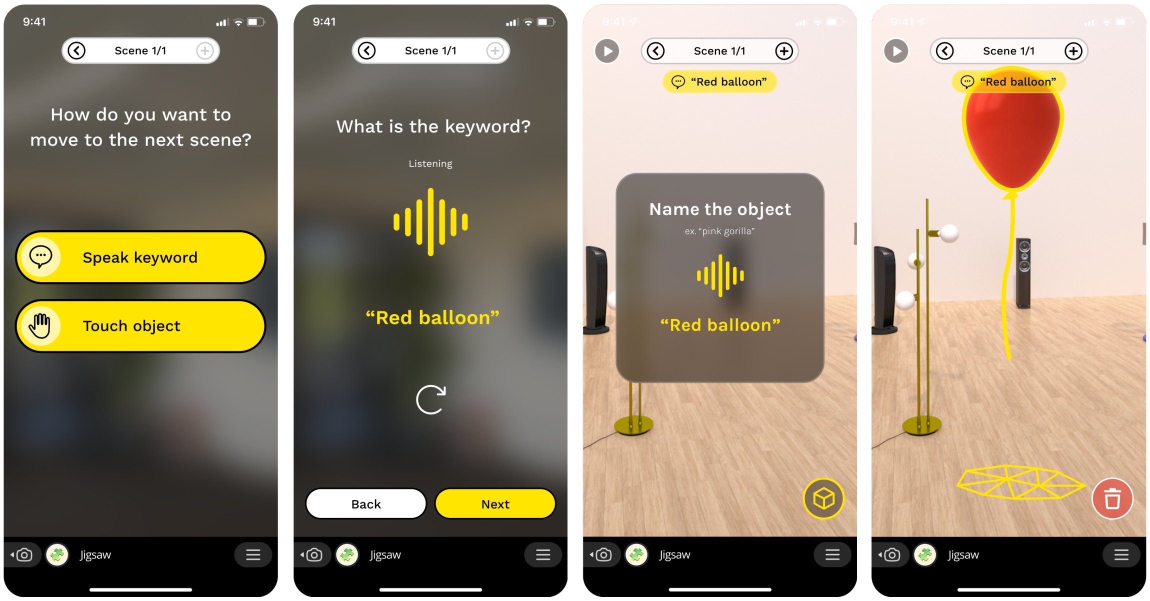
Opening Jigsaw's scene-based editor, the user first sees an empty scene with an initial state of IoT devices. The first scene that the user is going to build is dimming the smart lights when the user says "In a small, cozy room..." The user creates a new scene and defines the keyword "small" as the trigger by speaking the word directly to the system. After the new scene is created, the user can turn down the brightness of the smart lights by walking closer to one of the lights. An editing interface of the smart lights is shown in AR based on the user's proximity to the lights and allows the user to edit the status of the lights. In this way, the first pair of trigger (i.e., the keyword "small") and scene (i.e., changing the brightness of smart lights) is recorded in the system.

Next, the user creates more scenes in the editor till the end of the story. Within each scene, in addition to changing the status (e.g., on/off and intensity) of IoT devices, the user can also create, manipulate, and delete 3D models or animation. The user can create 3D content in AR by saying the model name (e.g., moon) directly to the system. They can place the content in the physical space by simple drag-and-drop interaction and delete it by dropping it to the trash icon on the screen. In this way, the user can author scenes such as a moon model appears and disappears. The system can automatically apply animation to the AR content of a scene. For example, when the user creates/deletes a moon model in a scene, the system can apply a fade in/out animation to the model when playing the scene. Similarly, the system automatically applies a translation animation to a 3D object by moving from the position of the previous scene, to the position of the next scene during playback.

When the user finishes authoring the story, they can "flip through" the scenes and make changes to previous scenes if needed. They can also play the experience immediately in the authoring environment.

3.2.2 Scene-based Authoring and In-situ Editing. Jigsaw enables novices to author immersive stories by creating a sequence of scenes in the editor, which drew inspiration from prior work [13] on scene-based authoring and presentation tools such as Google Slides, as seen in Fig. 2. The core of our authoring system is defining pairs of triggers and scenes that form the stories. We provide three types of triggers in the authoring system: keyword, touching, and tapping. The keyword trigger, as illustrated above, allows users to define custom keywords for entering each scene. This is incorporated since verbal narration has long been a key component of storytelling experiences. The touching trigger enables users to enter a new scene when their hands collide with an AR object. This is one of the basic interaction in AR applications in order to add to the interactivity of the stories. Lastly, tapping is the default trigger where the user enters a new scene by tapping on the screen.

In addition to different triggers, Jigsaw also enables novices to edit behaviors of AR content and IoT devices directly in AR, as seen in Fig. 3. This is aligned with the spirit of What-You-See-Is-What-You-Get that lots of authoring tools employ to lower the barrier for end-users to create immersive content [47]. Jigsaw currently supports in-situ editing of



(a) Scene-based authoring. (b) Keyword definition. (c) Model import. (d) Model manipulation.

Fig. 2. Interface of the authoring mode.

IoT devices including smart lights, smart fans, and smart speakers. A corresponding editing interface is popped up when the user's phone is close enough to the IoT device. For the smart light, users can edit its color, brightness, and special effect. The color is visualized in a semi-circular palette and the user can change its color by tilting the phone. Users can change the brightness using the same control, where the value of 0 will turn off the light. Lastly, users can add special effect of smart lights by saying the name (e.g. flickering) to the system. We envision delegating the low-level implementation of special effects to advanced developers and only expose the effect name to novice creators. Similarly, users can edit the on/off and intensity of the smart fan, and the on/off, volume, and special sound effects of the smart speaker.

3.2.3 Generalizability of the Authoring System. A key factor that makes our authoring system generalizable is that the the triggers and scenes that users create are relatively open-ended and not coupled with any specific story. For example, a user can define any keyword or touching with any AR object as the trigger and use the brightness of smart lights for any kinds of story. To demonstrate the generalizability of the authoring system, we use replicated examples [30] by replicating the other two immersive stories using the same system.

To replicate the Benjamin Franklin story, users can define pairs of keyword triggers and scenes like they do in creating the Goodnight Moon story. Two challenging parts of creating this story are allowing multiple behaviors in one scene and defining touch triggers. Different from the Goodnight Moon story, the Benjamin Franklin can have multiple behaviors triggered by one keyword in a scene. For example, when the user says “when the clouds first passed over”, the smart fan will turn on and the AR cloud model will appear. Jigsaw enables this by allowing users to define one state for multiple AR objects and IoT devices in a scene. The Benjamin Franklin story also incorporates touch triggers, where the user touches the virtual key model and triggers the flickering effect of smart lights. Jigsaw accomplishes this by

(a) Brightness Control of Smart Lights.

(b) Color Control of Smart Lights.

(c) Smart Speaker Control.

(d) Smart Fan Control.

Fig. 3. In-situ control interfaces for editing IoT devices.

allowing users to select trigger types when adding a new scene. After selecting touching as the trigger type, the user can then select the object for collision detection by touching it directly in AR.

The Wind and the Sun story is also different from the Goodnight Moon example since it requires the system instead of the user to narrate the story. Jigsaw also enables this by allowing users to select narrators at the beginning of the authoring experience. After selecting the system as the narrator, the user can add a narration to each scene by saying it directly to the system. The system can then transcribe the narration and play it through the speaker during the playback. When the system is assigned as the narrator, the user also doesn't need to define trigger when adding a new scene and the trigger type goes to default, i.e. tapping on the screen, as previously mentioned.

3.2.4 Implementation. The authoring system is implemented using Lens Studio for displaying and interacting with AR content. The connection with IoT devices is built using the same technique as the implementation of the immersive stories mentioned above. We used the built-in voice recognition and hand tracking component to enable users to define keyword and touch triggers. The playback of the final experience among multiple users is implemented using the connected lens component.

4 EVALUATION

We conducted an in-lab user evaluation with 20 participants (10 groups) with two goals: *(i) to investigate the benefits and challenges of immersive storytelling experiences with both virtual and physical augmentation, and (ii) to assess the usability and utility of the authoring system.* We recruited participants from a technology company in the United States through Slack posts and e-mails. Participants were randomly paired into 10 groups and volunteered for the one-hour study. Our study was approved by the privacy & legal team in our institution.

4.1 Study Procedure

Our study was divided into two sessions, each lasting 30 minutes. Prior to the study, participants first signed an agreement form that confirms their consent for our data collection and usage. They also completed a short entry survey about their demographic information and their prior experience of content creation. In the first session, we asked the participants to play the three immersive stories that we created in the order of Goodnight Moon, the Wind and the Sun, and Benjamin Franklin. Before playing each story, participants were debriefed about the story background and what their roles in the story. Since the Benjamin Franklin story requires three people to experience, one member from our study team acted as the Benjamin Franklin Jr. Our team member filled this audience role because participants already played the audience role in other stories such as the Wind and the Sun. After experiencing all the three stories, we then conducted a semi-structure interview with both participants together to ask about their experiences and the benefits and challenges in the experiences.

In the second session, we spent the other half an hour evaluating the authoring system with one randomly picked participant from the group. The participant was first given a tutorial introducing the functionalities of the authoring system. After ensuring the participant understood how the authoring system works, we then asked the participant to complete a task of recreating the Goodnight Moon story that they experienced in the first session. During the task, the participant was provided with the complete script on a piece of printed paper including keywords and the their corresponding behaviors. We provided this so that the participant could focus on *how* to create the story using the system instead of *what* to create. During the task, the study team was not allowed to provide any help unless the participant explicitly asked for help. Any question that the participant asked during the task was noted down. After completing the task, we conducted another semi-structure interview with this participant to ask about the usability and utility of the authoring system. Finally, we aggregated the entry survey data, transcribed and coded the recordings of the interviews.

5 RESULTS

Participants' experiences playing immersive stories during the study suggested that they felt that the storytelling experiences with both virtual and physical augmentation were immersive and memorable. In the evaluation of the authoring system, all participants were able to complete the recreation task and applauded the system for being easy to use and generalizable for creating a wide variety of stories. We center our findings around the two evaluation goals, *(i) benefits and challenges of immersive stories*, and *(ii) the usability and utility of the authoring system*

5.1 Benefits and Challenges of Immersive Stories

5.1.1 IoT devices made the experience immersive, engaging, and memorable. Our results suggested that IoT could make the storytelling experience more immersive by invoking more senses.

> "Because it's sensory, right? So it's not only are you seeing it on the screen again, it had panned out a little bit further and not only do you see it on the screen, but you're feeling it too, so it's 4D as opposed to just the AR experience of seeing it." -P11B

> "I just feel like if you just had AR simulated wind, there's something that's not as exciting about the fact of having real wind and having that effect come together. Whereas maybe for the lights for whatever reason, I don't know, because I guess you could just have lights in AR and have them change colours, but the wind being real because you can feel it, I guess."-P9B

Due to these senses, participants also found the stories more memorable with IoT devices involved.

> “I think it (without IoT) would’ve been less of an immersive experience and also probably little less memorable. So you remember, "Okay, at some point the moon went to sleep, the lights went out," and it’s a bit easier to remember visual cues than it is just reading it off. And then also staying with the story was a lot easier because things are changing around and keeping me engaged." -P2A

5.1.2 Users’ participation and interaction improves the sense of immersion and agency. We also found that by allowing users to act as different roles (e.g., narrators or characters in the stories) in the stories can enhance the sense of immersion. This is because when users act as a role in the stories and their interaction with the story elements triggers certain effects, it reinforce the impression that they are embodied as the character. When users act as a narrator, they will focus more on the storyline and feel immersed.

> “It makes you feel like you are in a theatre with the whole background noise, the light changes, and the fact that you can touch things virtually. And I mean, I had a physical book to read. I was very involved in everything going on, so I was definitely very immersed." -P5A

> “I think being the narrator helped me be immersed in the story more. I think maybe because I had to focus on the story more, if that makes sense. Since I’m narrating, I kind of have to think of what’s going to happen next or think what I have to say next. Whereas when I’m an audience member, I think it’s a little easier to get distracted by what’s going on and then lose what the computer’s saying." -P7A

We found that being a narrator also gives them more sense of agency since they can control the pace of how the story goes.

> “I definitely enjoyed being a narrator more because I felt more in control of the actions that I was committing throughout the story. So let’s say I was saying the story and then something came on the AR screen, I was able to look around and say, okay, I’m done looking at that. I’m ready for the next thing." -P7A

5.1.3 Sensory overload is the primary challenge in immersive stories. While Jigsaw’s immersive stories enabled a sense of immersion, agency, and engagement, we found that the design of immersive stories comes with the challenge of sensory overload. We found that participants’ attention was occupied by participation, mobile AR effects, and IoT effects, making them not being aware of all the changes.

> “I think if the experience doesn’t happen at the same time, but almost as if one glitch happened and then you see it in the world itself, that would make sense, because I feel like if you’re already dealing with something flashing in your phone and you’re going to kind of blink and miss what’s in your world, because you’re really focusing on what’s your phone there. Yeah, I think that might make more sense. Let’s say if my phone kind of flashed briefly, but then it made me panic, turn the flashing light and the light bulb so that way I’m not trying to focus on two things at the same time." -P11A

> “And then reading the book, honestly, I love reading books. I love reading to children. I didn’t need to have a phone. The reader in that case doesn’t need to have any devices. I don’t need to be a part of the immersive experience. I can just be the reader reading to the people in the room. So yeah, I didn’t really like having the phone sitting beside me because I’m kind of just like, well, now I’m tempted to pick up the phone, but I’m trying to read this story in a really nice way. So it just felt like there was this weird

> divide and I felt like I was being pulled in two different directions and not really being able to fully be in the moment. And then telling me, oh, if you pick up the phone, you can interact with the key too. Although it wasn't special to experience that, it didn't really make any sense. Why do I need to play with the key, I wasn't really a part of this story?" -P4B

5.1.4 Immersive stories are versatile for various population and scenarios. Lastly, we found that Jigsaw's immersive stories are versatile for both ambient and interactive scenarios, and for both kids and adults. For ambient scenarios, participants like to have IoT devices to set the mood of the environments.

> "So for me, I think one thing that I would love to see, so maybe for the adult story part, would be more horror type of stories. I could see a potential there because you have lights flickering, you have sort of sounds that come out of nowhere or wind that starts blowing. So I think there's a potential there to maybe have some horror based story like ghosts that you can only see with a phone and you can see with your eyes kind of thing." -P8A

> "So, I have a set bedtime and 30 minutes before that I'm supposed to do this end the day meditation, where I watch a sunset. And I just thought it reminded me of that, because it would be really cool to see that in AR and then maybe you power down your devices, because I'm supposed to be off my phone until bedtime. You power down your TV and electronics and screens and stuff, lower the lights maybe, that type of transition." -P10A

For interactive scenarios, participants like the interactions of Jigsaw's immersive stories and think they would be suitable for interactive experiences such as escape rooms where users can trigger certain effects via interacitons.

> "I can't tell you of how many escape rooms I know that do that kind of, you trigger one thing happens in the room, flickering lights like that, that would be perfect for it." -P11B

5.2 Usability and Utility of the Authoring System

5.2.1 The authoring system has a low barrier of entry. Participants commented that the authoring system is easy to learn because the in-situ interface that Jigsaw offers for users to control directly the IoT device in the space. They also think it requires little literacy in technology.

> "I think if I was a non-technical person, it was pretty easy to use because you just point at an object and then you can interact with it. So that mechanic, rather than me specifying things through a complicated user interface or a bunch of text, makes it a lot easier to just as a lay person to be, "Oh, I want the lights to turn on or to be red," and then you can just do it."-P8A

At the same time, participants think the authoring system requires less cost in terms of the devices to create immersive stories.

> "Especially at younger ages, you don't have that much production, you don't put too much production into it. So, just having, I would say, devices take care of a lot of that."-P2A

5.2.2 Open-ended triggers make the story generalizable and customizable. We found that the open-endedness of keyword triggers enables users to create a variety of stories.

> "Well, I guess the benefits are with the keyword part that's super open-ended versus just on a Google Slides, you know, have only preset options, and so you're somewhat limited to the options that they

> provide you with. But the keyword element being something where you could just really make it any word that you want to say, it makes that part fun for the interactions themselves"-P9A

5.2.3 Creating more complex stories is challenging using Jigsaw. We also found that the expressiveness of our authoring system is limited when participants wanted to author special animation and more complex behaviors of objects such as shaking animation and complex movements of objects.

> "Maybe if you wanted to make it feel like the ground was shaking or something like that, that would be hard."-P9A

> "[For the authoring tool] It'll have to be very basic type of AR models, not anything really moving or do a lot of movement, which kind of limits it to... Yeah, I don't think I'd be using the AR elements as much." -P11A

6 DISCUSSION

Overall, we found that Jigsaw enables immersive storytelling experience through the inclusion of IoT devices and participation. We found that the inclusion of IoT devices made the stories more immersive by adding more senses to the stories. The participation that Jigsaw enables in the stories also enhances users the sense of agency and immersion. At the same time, our findings suggest that immersive stories should be carefully designed as mobile AR, IoT effects, and participation can cause sensory overload. We also found that the immersive stories are versatile for various population and scenarios. We found participants were able to use our authoring tool to create immersive stories. Our authoring tool can also enable users to create diverse stories due to the open-endedness of the keyword triggers. However, our authoring tool has limited expressiveness for creating complex behaviors of objects or effects.

6.1 Design Implications and Future Directions

6.1.1 Adding immersion via physical augmentation and participation. We found that the augmentation of IoT devices and user participation enhanced the sense of immersion during storytelling experiences. This aligns with prior effort of bringing in physical augmentation and participation for immersive experiences such as 4-D theatres and participatory plays [6]. Our study indicates that enhancing readily available IoT devices, instead of using specialized, costly equipment, can effectively promote immersion. We encourage future researchers and practitioners exploring immersive stories integrate physical augmentation and participation to add to the immersion of stories. For example, future immersive stories might explore how more types of IoT devices such as rombots and projectors can be used to add to the immersion of stories. They can also explore more means of participation beside being narrators, actors, and audience and investigate how they can affect users' sense of immersion during storytelling experiences.

6.1.2 Creating immersive stories by editing open-ended attributes. Our findings demonstrate that Jigsaw's open-ended approach to keyword definition and behavior editing enables it to be widely adaptable for authoring diverse stories. Users can assign a single keyword (e.g. "balloon") or behavior (e.g. smart light brightness change) to serve different narrative functions across multiple stories. This flexibility empowers authors to get creative and develop inventive ways of linking keywords with behaviors to generate immersive experiences.

This underscores the importance for future research and design in immersive storytelling to embrace systems that do not enforce restrictive one-to-one mappings but rather enable many-to-many associations between narrative elements. Adopting adaptable frameworks like Jigsaw's could promote more creative authorial possibilities and accommodate wider storytelling applications. In addition, future researchers can consider ways for reusing or remixing [17] to make the authoring experience even more expressive for adapting to different storytelling contexts.

6.1.3 Supporting end-user creation of interactive behaviors. While being generalizable for creating various stories, one limitation we found about Jigsaw is its expressiveness for defining complex behaviors such as interactivity. Such complex behaviors typically requires coding to accomplish. We encourage future researchers to explore ways to raise the ceiling of expressiveness of the authoring tool for immersive stories while keeping the low barrier of entry for end-users. For example, in the context of editing animation, visual programming platforms such as Alice [15] and Scratch [39] could be incorporated into authoring tools like Jigsaw to enhance its expressiveness. In-situ programming approaches (e.g., [32, 48]) can also be considered as a way to enhance the while augmenting the existing intuitive in-situ controls of IoT devices in Jigsaw.

6.2 Limitations

Our results demonstrate the benefits and challenges of both the consumption and the creation of immersive stories. However, our work is not without limitations. First, Jigsaw currently incorporates a constrained set of modalities including mobile AR, smart lights, fans, and speakers. Further research should explore additional modalities like head-worn AR and diverse smart devices to fully uncover the possibilities of immersive storytelling. Second, our qualitative evaluation focused on adults; given children represent a key demographic for engaging with and creating stories, future studies could examine their perceptions and involvement. Follow-up work could also undertake more quantitative investigations into how immersion is impacted by blending virtual and physical augmentations. Finally, while designed for storytelling, Jigsaw's adaptability suggests applications in other domains like education, entertainment, and communication. Further research might examine how combining IoT and AR could play a role in these and other contexts. In summary, our work provides initial insights but warrants expanded investigations into users, technologies, and applications to elucidate the full potential of immersive storytelling.

7 CONCLUSION

In this work, we investigated immersive experiences created by integrating mobile AR with off-the-shelf IoT devices. We also introduced an authoring tool for crafting these immersive narratives. Our qualitative lab study revealed augmenting IoT devices and enabling user participation can enhance immersion in storytelling, though sensory overload poses a primary challenge. We also found our tool's open-ended approach to defining keywords and attributes enabled authoring diverse stories. These insights highlight opportunities and obstacles around AR+IoT storytelling. Our recommendations for future research and design include: furthering immersion via physical augmentation and participation; enabling immersive story creation by editing adaptable narrative components; and supporting end-user authoring of interactive behaviors.

REFERENCES

[1] 2011. Sleep No More. https://mckittrickhotel.com/sleep-no-more/
[2] 2022. Blueprints visual scripting. https://docs.unrealengine.com/5.0/en-US/blueprints-visual-scripting-in-unreal-engine/
[3] 2022. Digital Graffiti: Alys Beach – where immersive technology meets art and architecture. https://digitalgraffiti.com/
[4] 2022. Lens Studio. https://ar.snap.com/en-US/lens-studio
[5] 2022. Shader Editor. https://docs.blender.org/manual/en/latest/editors/shader_editor.html
[6] 2022. Space Mountain. https://disneyland.disney.go.com/attractions/disneyland/space-mountain/

[7] 2022. Star Tours: The adventures continue. https://www.disneylandparis.com/en-gb/attractions/disneyland-park/star-tours-the-aventures-continue/

[8] 2023. https://bondhome.io/

[9] 2023. https://www.amazon.com/Echo-4th-Gen/dp/B07XKF5RM3

[10] 2023. Kasa Smart. https://www.kasasmart.com/us/

[11] Isabel Allende. 1982. *La casa de los espíritus*. Plaza & Janés.

[12] Margaret Wise Brown. 1947. *Goodnight Moon* (50 ed.). Harper Collins, New York, NY.

[13] Mengyu Chen, Andrés Monroy-Hernández, and Misha Sra. 2021. SceneAR: Scene-based Micro Narratives for Sharing and Remixing in Augmented Reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 294–303. https://doi.org/10.1109/ISMAR52148.2021.00045

[14] Richard Cook. 2021. *Connecting the Echo Dots: An Exploratory Ethnographic Study of 'Alexa' in the Classroom*. Ph. D. Dissertation. https://doi.org/10.13140/RG.2.2.10861.72164

[15] Stephen Cooper, Wanda Dann, and Randy Pausch. 2000. Alice: a 3-D tool for introductory programming concepts. *Journal of computing sciences in colleges* 15, 5 (2000), 107–116.

[16] DAQRI. 2014. Elements 4D interactive blocks. https://www.kickstarter.com/projects/daqri/elements-4d-interactive-blocks

[17] Sayamindu Dasgupta, William Hale, Andrés Monroy-Hernández, and Benjamin Mako Hill. 2016. Remixing as a pathway to computational thinking. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 1438–1449.

[18] Barrett Ens, Fraser Anderson, Tovi Grossman, Michelle Annett, Pourang Irani, and George Fitzmaurice. 2017. Ivy: Exploring Spatially Situated Visual Programming for Authoring and Understanding Intelligent Environments. In *Proceedings of the 43rd Graphics Interface Conference* (Edmonton, Alberta, Canada) *(GI '17)*. Canadian Human-Computer Communications Society, Waterloo, CAN, 156–162.

[19] Brian Feldman. 2019. The best use of augmented reality right now is the Weather Channel's. https://nymag.com/intelligencer/2019/01/the-weather-channels-augmented-reality-segments.html

[20] Terrell Glenn, Ananya Ipsita, Caleb Carithers, Kylie Peppler, and Karthik Ramani. 2020. *StoryMakAR: Bringing Stories to Life With An Augmented Reality & Physical Prototyping Toolkit for Youth*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376790

[21] Ontroerend Goed. 2004. 'The smile off your face (2004)' - projects: Ontroerend goed. http://www.ontroerendgoed.be/en/projecten/the-smile-off-your-face/

[22] Ontroerend Goed. 2007. 'internal (2007)' - projects: Ontroerend goed. http://www.ontroerendgoed.be/en/projecten/intern/

[23] Ontroerend Goed. 2010. 'A game of you' - projects: Ontroerend goed. http://www.ontroerendgoed.be/en/projecten/a-game-of-you/

[24] Ant Hampton and Silvia Mercuriali. 2020. Autoteatro. https://rotozaza.co.uk/autoteatro.html

[25] Jennifer Healey, Duotun Wang, Curtis Wigington, Tong Sun, and Huaishu Peng. 2021. A Mixed-Reality System to Promote Child Engagement in Remote Intergenerational Storytelling. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 274–279. https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00063

[26] Layne Jackson Hubbard, Yifan Chen, Eliana Colunga, Pilyoung Kim, and Tom Yeh. 2021. Child-Robot Interaction to Integrate Reflective Storytelling Into Creative Play. In *Creativity and Cognition* (Virtual Event, Italy) *(C&C '21)*. Association for Computing Machinery, New York, NY, USA, Article 16, 8 pages. https://doi.org/10.1145/3450741.3465254

[27] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 637–644. https://doi.org/10.1145/2642918.2647383

[28] Sanghoon Kim and In-Young Ko. 2022. A Conversational Approach for Modifying Service Mashups in IoT Environments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 345, 16 pages. https://doi.org/10.1145/3491102.3517655

[29] Jacqueline Kory and Cynthia Breazeal. 2014. Storytelling with robots: Learning companions for preschool children's language development. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. 643–648. https://doi.org/10.1109/ROMAN.2014.6926325

[30] David Ledo, Steven Houben, Jo Vermeulen, Nicolai Marquardt, Lora Oehlberg, and Saul Greenberg. 2018. Evaluation strategies for HCI toolkit research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–17.

[31] G.A. Lee, C. Nelles, M. Billinghurst, and G.J. Kim. 2004. Immersive authoring of tangible augmented reality applications. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*. 172–181. https://doi.org/10.1109/ISMAR.2004.34

[32] Gun A Lee, Claudia Nelles, Mark Billinghurst, and Gerard Jounghyun Kim. 2004. Immersive authoring of tangible augmented reality applications. In *Third IEEE and ACM international symposium on mixed and augmented reality*. IEEE, 172–181.

[33] Yi-Bing Lin, Ming-Ta Yang, and Yun-Wei Lin. 2019. Low-Cost Four-Dimensional Experience Theater Using Home Appliances. *IEEE Transactions on Multimedia* 21, 5 (2019), 1161–1168. https://doi.org/10.1109/TMM.2018.2876043

[34] Gabriel García Márquez. 1981. *Chronicle of a Death Foretold*. La Oveja Negra.

[35] Alex Olwal and Artem Dementyev. 2022. Hidden Interfaces for Ambient Computing: Enabling Interaction in Everyday Materials through High-Brightness Visuals on Low-Cost Matrix Displays. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 504, 20 pages. https://doi.org/10.1145/3491102.3517674

[36] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andy Wilson. 2016. Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment. In *CSCW '16 Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 1716–1725. https://www.microsoft.com/en-us/research/publication/room2room-enabling-life-size-telepresence-in-a-projected-augmented-reality-environment/

[37] Joanna Resnick, Ruth Saxton, Matt Angove, and Alan Lane. 2010. Anthology. https://www.slunglow.org/anthology/

[38] Joanna Resnick, Ruth Saxton, Matt Angove, and Alan Lane. 2012. They only come at night: Resurrection. https://www.slunglow.org/come-night-resurrection/

[39] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, et al. 2009. Scratch: programming for all. *Commun. ACM* 52, 11 (2009), 60–67.

[40] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, and Yasmin Kafai. 2009. Scratch: Programming for All. *Commun. ACM* 52, 11 (nov 2009), 60–67. https://doi.org/10.1145/1592761.1592779

[41] Massimiliano Siccardi. 2022. Van Gogh exhibit Los Angeles: The immersive experience. https://www.vangoghla.com/

[42] Cristina Sylla. 2013. Designing a Tangible Interface for Collaborative Storytelling to Access 'embodiment' and Meaning Making. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, New York, USA) *(IDC '13)*. Association for Computing Machinery, New York, NY, USA, 651–654. https://doi.org/10.1145/2485760.2485881

[43] teamLab. 2021. Sketch Ocean. https://www.teamlab.art/e/asianart_sketchocean/

[44] GARETH WHITE. 2012. On Immersive Theatre. *Theatre Research International* 37, 3 (2012), 221–235. https://doi.org/10.1017/S0307883312000880

[45] Inc. Within Unlimited. 2018. https://wonderscope.com/

[46] Hui Ye and Hongbo Fu. 2022. ProGesAR: Mobile AR Prototyping for Proxemic and Gestural Interactions with Real-World IoT Enhanced Spaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 130, 14 pages. https://doi.org/10.1145/3491102.3517689

[47] Lei Zhang and Steve Oney. 2020. FlowMatic: An Immersive Authoring Tool for Creating Interactive Scenes in Virtual Reality. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 342–353. https://doi.org/10.1145/3379337.3415824

[48] Zhengzhe Zhu, Ziyi Liu, Youyou Zhang, Lijun Zhu, Joey Huang, Ana M Villanueva, Xun Qian, Kylie Peppler, and Karthik Ramani. 2023. LearnIoTVR: An End-to-End Virtual Reality Environment Providing Authentic Learning Experiences for Internet of Things. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.