"Let's Make A Story": Measuring MR Child Engagement

Duotun Wang
duotun@cs.umd.edu
University of Maryland
College Park, MD, USA

Jennifer Healey
jehealey@adobe.com
Adobe Research
San Jose, CA, USA

Jing Qian
jing_qian@brown.edu
Brown University
Providence, RI, USA

Curtis Wigington
wigingto@adobe.com
Adobe Research
San Jose, CA, USA

Tong Sun
tsun@adobe.com
Adobe Research
San Jose, CA, USA

Huaishu Peng
huaishu@cs.umd.edu
University of Maryland
College Park, MD, USA

Figure 1: "Let's Make A Story" system makes it possible for the grandparent and grandchild experience the story simultaneously but with different views and roles: a) and b) grandparents could control the story environment and see their grandchild's face and talk to them during MR story play; c) and d) grandchild could manipulate the can-controlled character and hear their grandparent's voice

ABSTRACT

We present the result of a pilot study measuring child engagement with the "Let's Make A Story" system, a novel mixed reality (MR) collaborative storytelling system designed for grandparents and grandchildren. We compare our MR experience against an equivalent paper story experience. The goal of our pilot was to test the system with actual child users and assess the goodness of using metrics of time, user generated story content and facial expression analysis as metrics of child engagement. We find that multiple confounding variables make these metrics problematic including attribution of engagement time, spontaneous non-story related conversation and having the child's full forward face continuously in view during the story. We present our platform and experiences and our finding that the strongest metric was user comments in the post-experiential interview.

CCS CONCEPTS

• **Human-centered Interface** → **Augmented Reality**; *Human Factors*; Network.

KEYWORDS

Reading, children, mixed reality, augmented reality, family communications

1 INTRODUCTION

In a world where people are often separated by distance, people crave meaningful experiences that can happen remotely. The "Let's Make A Story" system is designed to create an engaging collaborative experience for physically distant grandparents and grandchildren. We present our system and our insights from a set of three experiences with children ages 6,7 and 9. Our application builds on paradigm of storytelling, where an adult will read a story to a child and share a bonding experience. We address the challenge that it is difficult to engage children's attention for long periods of time over video calls [8] and leverage the fact that children today are often frequently engaged by phone video games. We designed a story sharing system that addresses both the needs of grandparents and grandchildren. For the grandparents, we created a desktop application that includes the words of the story and various controls for the story environment. The grandparents control the pace of the story and actions that are necessary for the story to advance. Grandchildren are given a phone based application that resembles a game and a manipulative, a can wrapped in an MR triggering image, to control the protagonist in the story. Grandparents see and control the game and also see their grandchild's face while children are game focused while hearing their grandparents Voice. The twin experiences are shown in Figure 1. While our initial plan was to measuring engagement by story time, amount of supplemental story generated by the user and facial expression analysis, we found that there are many practical considerations that need to be considered to make these viable metrics. We instead find that user response in the post-hoc interview is highly informative.

2 RELATED WORK

Video conferencing has been found to be more highly engaging than phone calls when it comes to connecting grandparents and

grandchilden, as children are often at a loss when it comes to sustaining conversation without expressing themselves physically [1]. Interacting with grandchildren was found to be the "the primary, if not sole, motivation for video" [1] for grandparents who would prefer the phone. Videochat allows kids to be more engaged[3] and allows them to assert their own participation more easily by putting something (an object, their body) in front of the camera rather than needing to find words to share[4].

Beyond videochat alone, systems such as Family Story Play system[7] and StoryVisit[8] allow videoconferencing to be augmented with remote storytelling. While StoryVisit uses a shared ebook, Family Story Play wuses a paper book with an augmented frame that plays videos of a popular television character at different parts of the story. Both these studies found that augmented storytelling increased child engagement using time in story as a metric.

The problem of remote grandchild-grandparent engagement still has not been solved [3–5]. Many of the reasons cited in prior work include the disparity in the levels of technical literacy and comfort between the generations[1, 3, 8]. We designed our mixed reality experience with these concerns in mind, providing an easy to use desktop interface for grandparents and a fun mobile phone based MR experience with tangibles for grandchildren.

3 SYSTEM OVERVIEW AND INTERACTION

Our system, inspired by prior work[6–8], consists of a desktop program for the grandparents (Figure 1a) and a mobile phone mixed reality application (Figure 1c). The grandparent reads the text of the story and controls story-related events (dropping apples, controlling endings). The child generates 3 MR scenes (Figure 2c) using a 3-page printed paper book (Figure 2a) and controls a manipulative character using a 12oz can wrapped in an MR-activating sheath (Figure 2b). The grandparent begins the application using the desktop app to invite the child to "join". The grandparent has reads the text of the story and controls the story-related events. The child uses the can to make the MR character interact with the virtual background displayed on the phone. Each page ends with MR activity initiated by the grandparent and completed by the child.

4 USER STUDY

We conducted a pilot study with three participants, two boys aged 6 and 9 and one girl aged 7. All participation was remote. An Android APK for the MR experience was customized for each participant's phone as camera resolution is an important factor. Each participant's parents were given instructions on how to assemble all paper materials which were printed by the participants. This pilot was focused on measuring child engagement, so for convenience the role of the grandparent was played by one of the experimenters. Each child participant read two stories, one paper based story as a benchmark and one story with the MR game. Both stories were

Figure 2: a) Printed book pages were used for generating c) interactive virtual background for the MR Story Play; b) Cans were used as MR Characters (little red and wolf). Book images were redesigned from Behance [1]. 3D Models were from Unity Asset Store [2].

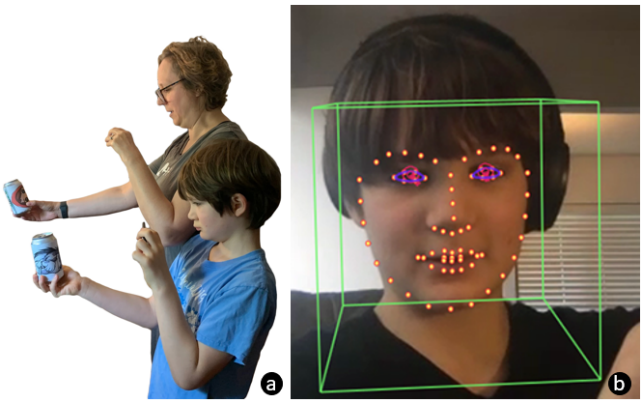

Figure 3: a) Children with short arm has difficulties positioning the phone camera in related to the can. b) Using OpenFace to measure engagement.

versions "Little Red Riding Hood," one version told from Little Red's perspective and the other from the Wolf's. The MR background, shown in Figure 2c was the same for both stories, but the can based characters, shown in Figure 2b and active elements were protagonist specific (e.g., apples vs. bunnies, different paths through the forests, different antagonist and action choices in final scene).

The grandparent leads the interaction and the child follows along. For the paper story interactions include answering questions about the story visuals and lifting paper flaps to find a hidden items. In the MR experience, activities include making the MR background appear and summon the character with the can. The child is free to play with the character while the grandparent reads the text. Each page ends with a MR activity like "catch the bunnies", "follow the path" or "choose an action." At the end of the third page of both stories the child is asked to create an ending. Children and their parents (if present) are briefly asked about their experiences at the end of the session.

[1]House Interiors: https://assetstore.unity.com/packages/3d/environments/fantasy/retro-dungeons-house-interiors-170705
Isometric Pack 3D: https://assetstore.unity.com/packages/3d/environments/fantasy/isome-tric-pack-3d-62262
[2]Little Red Book Cover: https://www.behance.net/gallery/9151145/Little-Red-Riding-Hood-(Book-Covers)
[3]Werewolf Cute Series: https://assetstore.unity.com/packages/3d/characters/creatures/werewolf-cute-series-177868

5 MEASUREMENT CHALLENGES

Our initial hypothesis was that, similar to prior studies, we use time as an engagement metric. We also wanted to explore measuring user generated content (the story ending) and facial expressions as potential engagement metrics. During the pilot study we encountered several challenges that caused us to reconsider how best to measure engagement, described below.

5.1 Time Metric

Interaction time is the most common metric of engagement. The amount of time an experience take, however, can reflect more than just the participant's interest. For example, we found that children often took a lot of time to get the MR can-controlled character to appear in the scene. Factors including lack of adequate lighting, children occluding part of the overlay with their fingers and children's shorter arms lengths illustrated in Figure 3a all seemed to contribute to the difficulty. In two of the three cases, simultaneous can and phone manipulation was too difficult and the strategy of placing the can on the table and moving the phone to move the character was adopted (shown in Figure 1). The paper stories also had technical difficulties where sometimes pieces were lost or not assembled properly. To alleviate these confounds, we plan to go through the videos by hand and eliminate all time periods related to technical difficulties. We also found having the children test making the character appear before starting the story uncovered possible issues in advance.

5.2 Participant Generated Story

Another metric we had considered was user generated conversation. Throughout the story the grandchild was given opportunities to interact with the grandparent. We found that in the programmed interaction none of the three children elaborated on the story beyond simple answers to the questions. In the final scene, where the children need to make up an ending children's responses varied, but almost all of them were confused at first and needed prompting. One child was reluctant to say anything in either the paper or the MR ending. Another child was proud that they knew the actual story of Little Red Riding Hood story form memory at the end of the paper story which was presented first. In the third case the child was more interested in giving commentary on the difference between the wolf characters in the two stories at the end of the paper story which in this case was presented second. To alleviate this we plan to sort coments related to the story versus non-related comments.

5.3 Facial Expression Analysis

Another promising engagement metric is facial expression analysis, using software such OpenFace [2] which can quantify gaze and facial action units used for emotion recognition (Figure 3b). Quantitative metrics such as the duration, frequency, or pattern of emotion involvement can provide more context about the overall experience. We noticed that children in MR conditions smiled for a longer period of time, but sometimes facilal occlusions confounded automated analysis. We plan to use automated analysis when possible and use human raters when faces are occluded to judge affect.

5.4 Post-Hoc Interviews

Despite out hope for objective metrics, we believe that the most explicit evidence for engagement came from the subjective post-hoc interviews. When asked about the experience, both the children and the adults co-located with the children expressed a high degree of enthusiasm for the MR experience. Comments such as "That was epic," "This is awesome" and a follow up text comment "It was SO FUN!" all lead us to believe that the experience was engaging. We plan to make more a more formal analysis of these comments, potentially using sentiment analysis on transcripts.

6 CONCLUSION

We believe that our system has the potential to create engaging cross-generational experiences that both grandparents and grandchildren will love. In our initial pilot we tested the hypothesis that we could measure engagement through story time, user generated story content and facial expression analytics. We found that confounds needed to be considered for each of metrics. We found that the post-hoc user interviews were far richer and more descriptive than we expected and we plan to rely more on these in the future. We had initially wanted to avoid a subjective assessment from children, but from our pilot we found that co-located parents also observed the child's interaction and were able to have an informed opinion on their child's engagement.

REFERENCES

[1] Morgan G. Ames, Janet Go, Joseph 'Jofish' Kaye, and Mirjana Spasojevic. 2010. Making Love in the Network Closet: The Benefits and Work of Family Videochat. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 145–154. https://doi.org/10.1145/1718918.1718946

[2] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, 59–66.

[3] Sean Follmer, Hayes Raffle, Janet Go, Rafael Ballagas, and Hiroshi Ishii. 2010. Video Play: Playful Interactions in Video Conferencing for Long-Distance Families with Young Children. In *Proceedings of the 9th International Conference on Interaction Design and Children* (Barcelona, Spain) *(IDC '10)*. Association for Computing Machinery, New York, NY, USA, 49–58. https://doi.org/10.1145/1810543.1810550

[4] Azadeh Forghani and Carman Neustaedter. 2014. The Routines and Needs of Grandparents and Parents for Grandparent-Grandchild Conversations over Distance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 4177–4186. https://doi.org/10.1145/2556288.2557255

[5] Janine Hacker, Jan vom Brocke, Joshua Handali, Markus Otto, and Johannes Schneider. 2020. Virtually in this together - how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. *European Journal of Information Systems* 29, 5 (2020), 563–584. https://doi.org/10.1080/0960085X.2020.1814680

[6] Anuruddha Hettiarachchi and Daniel Wigdor. 2016. Annexing reality: Enabling opportunistic use of everyday objects as tangible proxies in augmented reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1957–1967.

[7] Hayes Raffle, Rafael Ballagas, Glenda Revelle, Hiroshi Horii, Sean Follmer, Janet Go, Emily Reardon, Koichi Mori, Joseph Kaye, and Mirjana Spasojevic. 2010. *Family Story Play: Reading with Young Children (and Elmo) over a Distance.* Association for Computing Machinery, New York, NY, USA, 1583–1592. https://doi.org/10.1145/1753326.1753563

[8] Hayes Raffle, Glenda Revelle, Koichi Mori, Rafael Ballagas, Kyle Buza, Hiroshi Horii, Joseph Kaye, Kristin Cook, Natalie Freed, Janet Go, and Mirjana Spasojevic. 2011. Hello, is Grandma There? Let's Read! StoryVisit: Family Video Chat and Connected e-Books. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1195–1204. https://doi.org/10.1145/1978942.1979121