Engaging Children in Storytelling Through Tabletop Play: Exploring Construction of Story Ideas through Enactive Actions and Vocalizations

Ting Liu
tingliu@tamu.edu
Texas A&M Univeristy
USA

Francis Quek
quek@tamu.edu
Texas A&M Univeristy
USA

Gabriela Gomez
geg2289@tamu.edu
Texas A&M Univeristy
USA

ABSTRACT

This paper describes a Tabletop Imaginative Play as Enactive Storytelling (TIPES) system for children that captures their tabletop enactment and transforms them into animated cartoons in real time. In this paper, we compare children's storytelling support using TIPES (AR condition) and the raw tabletop-play recordings without using TIPES (TR condition). We conducted a study involving 18 pairs of 7-13 year-old children. Each pair engaged in creating two stories in the AR or TR conditions, respectively. We coded for *actions* and *vocalizations* and grouped these action-vocalization compounds into *enactment idea units*. We identified a statistically significant positive correlation between total actions and vocalizations of each story in AR but not in TR and a statistically significant negative correlation between *action complexity* (actions/idea unit) and *vocal richness* (vocalizations/idea unit) in TR but not in AR. It can be inferred that children manifest distinct engagement in constructing story enactment ideas by demonstrating a more balanced management of semantic loads between executing actions and uttering vocalizations in AR compared to TR. This implies varying levels of incorporating multimodality in expression across different conditions. Furthermore, the analysis of interview responses highlighted greater enjoyment in utilizing TIPES.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; **Interactive systems and tools**.

KEYWORDS

Storytelling, Tabletop play, Enactment, Physical actions, Vocalizations, Engagement, Enjoyment

ACM Reference Format:
Ting Liu, Francis Quek, and Gabriela Gomez. 2024. Engaging Children in Storytelling Through Tabletop Play: Exploring Construction of Story Ideas through Enactive Actions and Vocalizations. In *Interaction Design and Children (IDC '24), June 17–20, 2024, Delft, Netherlands.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3628516.3655751

IDC '24, June 17–20, 2024, Delft, Netherlands

ACM ISBN 979-8-4007-0442-0/24/06.
https://doi.org/10.1145/3628516.3655751

1 INTRODUCTION

Expressive communication requires a level of sophistication in language to articulate and convey ideas. Such communication requires ideas to express, the structuring of ideas for expression, and the necessary language skills. Children, especially elementary school students who are still in the process of attaining mastery of language often struggle with this combinatorial process [16]. However, when engaged in imaginative play or role-play (also referred to in literature as pretend play [4], or story enactment [9]), children reveal remarkable capability for story creation and expression. Children's enactive play is replete with imagination and stories. They form imaginative ideas and create their own narratives when giving voice to their toys and building story fragments in their play.

Our research investigates the effect of employing Digitally-Augmented Enactment (DAE) [37] on children's idea generation and formulation (structuring) in relation to their expressive abilities. Instead of having children go directly from generating ideas to structured verbalization, our approach leverages technology as a 'bridge' to decompose the storytelling activity into two stages: story creation and structuring through embodied enactment and the encoding of the story into linguistic expression. In this paper, we introduce the Tabletop Imaginative Play as Enactive Storytelling (TIPES, pronounced TȳPES) system that augments physical tabletop play with realtime digital animations. This animation is combined with children's voice-acting at the time of play during playback.

This paper compares children's storytelling support using TIPES (Animation Recording, AR, condition) and raw video recordings of tabletop play without TIPES (Tabletop Recording, TR, condition). The children were given story prompts for their tabletop story play. The children wrote their stories after either watching the TIPES-generated cartoons (AR) or the raw tabletop video recordings (TR). These recordings are multimodal, combining the physical enactment through the tabletop toys with the vocalizations (voice-acting, sound effects, narrations). By separately analyzing the physical actions and vocalizations in the story recordings, we acquire a deeper understanding of the children's enactment idea construction process.

2 BACKGROUND AND RELATED WORK

In this section, we present a brief review of the previous research on the effects of story enactment on children's cognitive development and digital storytelling approaches.

2.1 Storytelling and Enactment

Storytelling fulfills various functions, encompassing education, entertainment, and cultural preservation. According to Bruner [6], narratives are an essential means of organizing and interpreting experiences, empowering individuals to derive meaning from their surroundings. Utilizing narrative structures simplifies intricate concepts and ideas, fostering enhanced comprehension and engagement. Engaging in play encourages individuals to explore fictional worlds, fostering the exercise of imagination and creative thinking abilities [5, 17]. Individuals immersed in enactive story worlds during play envision various scenarios, construct mental images, and craft elaborate narratives and characters [3, 13, 17]. This imaginative engagement not only has the potential to enhance creative thinking but also to facilitate the development of critical thinking skills, adaptability, and emotional intelligence [2]. Additionally, our research aims to provide children with a collaborative storytelling experience by asking players to cooperate and negotiate during the play, leading to the cultivation of social skills [29].

Piaget's stages of children's cognitive development outline key milestones in phases such as sensorimotor, preoperational, concrete operational (COP), and formal operational (FOP) thinking [28]. While questions have been raised about the exact points of transition, that children generally move from a more concrete form of thinking (e.g., using physical manipulatives for arithmetics) to more abstract forms of thinking around the 7 - 12 years age range (COP to FOP) is undisputed. We situate our research within this general stage of transition.

2.2 Digital Storytelling

Digital augmentation in storytelling integrates digital technologies to enhance traditional narrative experiences. Jager et al. [14] conducted a systematic review of digital storytelling in research, involving 25 articles that reported its benefits on professional development, education, and community building. Various researchers have significantly contributed to the development of digital storytelling applications. Helmes et al. [19] and Alofs et al. [1] presented storytelling applications allowing children to interact with 2D scenery and story characters on a multi-touch tabletop screen. For 3D experiences, Song et al. [32] provided a mixed reality storytelling environment based on Kinect and HoloLens. Other works relied on tracking devices like RFID [21] and body motion tracking equipment [8, 38] to support children in telling imaginative stories in a mixed reality environment. Kim et al. [22] generated an immersive and interactive augmented reality running on a miniature based on real-time camera tracking. Juan et al. [20], utilizing physical markers for camera tracking, incorporated audio of an invented narrative into storytelling. Park and Moon [27] presented an AR-based interaction system utilizing ARToolKit markers attached to tangible objects. Our approach involves developing a storytelling system offering a 3D experience, employing real-time camera tracking of physical fiducial markers, and implementing a simple yet effective tabletop setup.

3 TIPES SYSTEM

The TIPES system (Figure 1(a)) enables the coherent linkage of physical figurines with virtual objects and provides real-time 3D graphical support displayed on a screen. Children generate story ideas and structure as they voice-act and manipulate their tangible figurines on the tabletop space (Figure 1(b)). The system captures physical figurine movements on the tabletop (Figure 2(a)) through a top-down web camera. The figurines are tracked using ArUco-based marker tracking [26] to produce realtime animated story scenes (Figure 2(b)) through a Unity-based [36] animation subsystem. Further details on system design and development can be found in a preceding paper [23].

4 THE STUDY

We conducted a user study with a total of 36 participants, 16 boys and 20 girls, ranging from 7 to 13. Each study trial lasts approximately 1 to 1.5 hours. The following section details the selection and recruitment of the participants, how the study was designed, and a brief overview of various data collected.

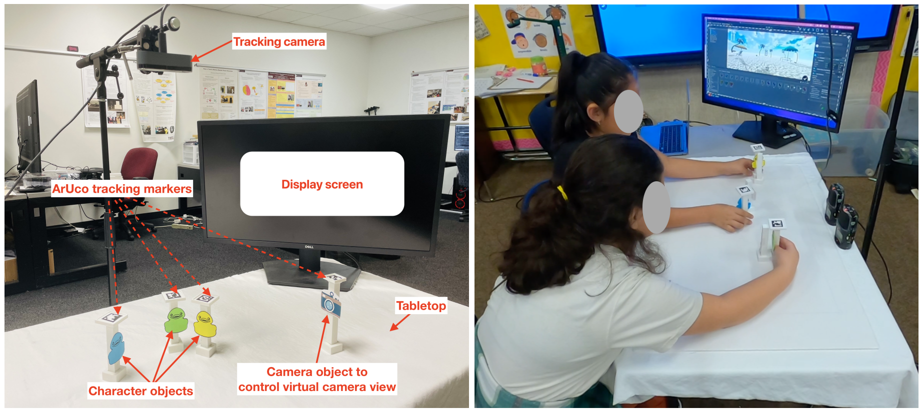
4.1 Selection and Recruitment of Participants

The participants of this study were selected off the following criteria: between the ages of 7 to 14, have a reasonable understanding of the English language. We recruited participants for the study with the help of a local elementary school teacher with deep ties to a local Hispanic community. This teacher informed people in his community of the study and allowed the research team to talk with prospective parents of the participants. The participants had varying understandings of the English language and additionally Spanish. An assent form and a consent form were presented, respectively, to the child and his or her legal guardian. The forms were both available in English and Spanish, as some of the parents had a limited understanding of English. Both the participants and the legal guardians were made aware that the data collected is confidential to the research team, and face blurring would be applied for any potential publications. The participants were thanked for their time, and they received a small gift of their choosing at the end of the study.

4.2 Study Design

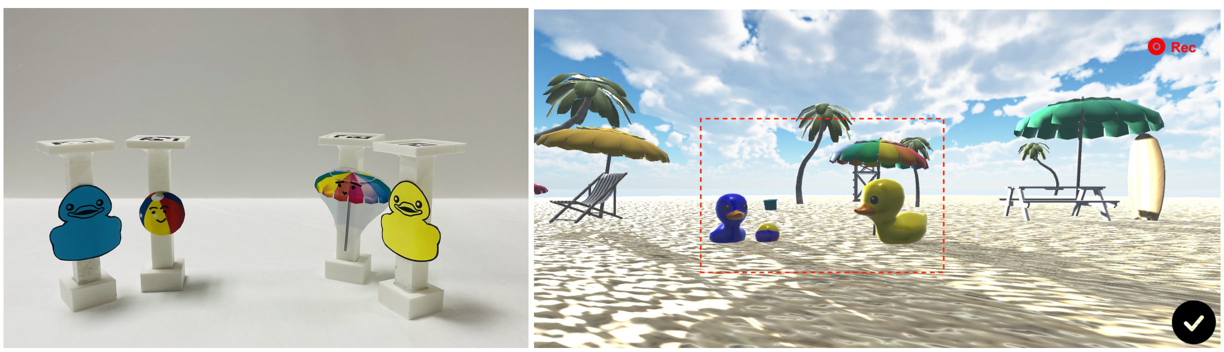
The children were paired up by similar age and same gender. Each pair was asked to create two stories in a counter-balanced order of AR and TR conditions (see Figure 1(b)). To prevent children from using identical characters and thus generating similar stories, we provided different sets of figurines for the AR and TR conditions. Each set comprised six characters. While three figurines (a yellow, blue, and green duck) were common to both conditions, the remaining ones differed. The AR set included a beach ball, an umbrella, and a palm tree, whereas the TR set had a sofa, a couch, and a lamp. These varied figurines were used to inspire different stories.

The participants were first asked to fill out a pre-survey including the Mini Big Five [31] and some pre-questions to gauge how they felt about storytelling before being presented with the AR and TR conditions. As a point of note, we reformulated the questions in the Mini Big Five questionnaire to ensure they're comprehensible for young children since they were originally designed for adults. The children were then given a brief explanation of the setup and encouraged to start story planning with a prompt sentence given to them by one of the researchers: "Imagine that you are at a friend's



(a) TIPES system setup (b) Children's tabletop playing (TR)

Figure 1: Collaborative Story Enactment approach



(a) Figurines' layout on the tabletop (b) Corresponding virtual scene in the screen (default camera view)

Figure 2: Transformation from a Tabletop scene to a Virtual scene

house/on the beach and you find out all the furniture/beach environment objects can move and talk." Each pair was then given 8 minutes for TR and 10 minutes for AR to plan a two-minute story with their counterpart. The extra two minutes for AR were provided to allow children to get familiar with the TIPES system. The children then stated to a researcher that they were ready to record.

After the children completed their story enactment, the recorded enactment was played back to them, and they were tasked with writing a story based on it. In addition, to mitigate challenges arising from spelling and motor difficulties in writing [12], we provide the option of dictation, enabling children to verbally express their thoughts of what to write about the story, which we then assist in typing. Children have the flexibility to review and make edits to their dictated story at any point. After watching the recording, a researcher prompted the child to start with the dictation process by asking, "What do you want to write? Tell me the words and sentences, and I'll type them down for you." Occasionally, the child chose to write it themselves. A post-survey followed; it was used to assess their enjoyment and opinions on the story enactment and writings for different conditions. The story-creation process was repeated for the opposite condition. Finally, a post-interview was conducted to measure the child's overall experience. The whole study process was recorded using a Go-Pro camera, which was used for observing the children's physical actions in later data analysis.

4.3 Data Collected

The data collected from the various stages of the study are displayed in Table 1. Each pair of children planned and enacted their story together but wrote it out separately. Of the 72 written stories in both conditions, 64 were dictated and 8 were handwritten. And each child in the pair was surveyed and interviewed separately. During the pre-survey and post-interview stages, we didn't distinguish the conditions and repeated the processes.

5 ANALYSIS

To comprehend the impact of our approach on children's engagement in both story enactment and retelling stages, we conducted three analyses utilizing different sets of data: enactment analysis focused on story enactment recordings; writing analysis on written story texts; and enjoyment analysis derived from questionnaires

Table 1: Different types of Data collected from the Study

	Participants	Pre-surveys	Post-surveys	Post-interviews	Story enactment recordings	Writings
AR	18 pairs	36	36	36	18	36
TR	18 pairs		36		18	36

and interview responses. Enactment analysis aims to elucidate the process of children's idea construction during play, while writing analysis assesses children's capabilities in retelling stories.

5.1 Enactment Analysis

We base our cross-modal analysis of the enactments on the theory of motion descriptions by Leonard Talmy [34]. According to Talmy, motion descriptions contain multiple aspects (e.g., manner, ground, direction, subject, object) [33–35]. McNeill and Quek showed these motion descriptions are represented multimodally through visual expression (e.g., gesture and enactment) and speech [24, 25, 30]. We analyze how the visual performance in the story enactments relates to the vocalizations as an indicator of the integrated story imagination. We further analyze how these enactment-vocalization compounds combine into sequences, providing insights into the cohesion of the constructed story scenes.

We obtained the raw video recordings of children's story enactment in both AR and TR conditions and transcribed them. Among all 18 pairs of children, we obtained 15 pairs of valid data. The raw story enactment recordings of two pairs using our system (AR) were incomplete due to technical recording problems. Additionally, in one pair, one child participated in story planning and rehearsals with her partner but was too shy to complete the formal recording. And a researcher intervened to assist the other child with the recording. Therefore, these three pairs of video recordings were excluded from the enactment analysis due to practical issues. In total, we analyzed approximately 102 minutes of video recordings, as quite a few groups produced stories exceeding the required duration of 2 minutes. The enactment analysis results from both partners within the same pair were combined since they collaborated to create the same story, thus preventing overlap.

5.1.1 Analysis methods. We applied three approaches derived from prior research to extract vignettes (sequential combinations of multimodal compounds) [7, 9], descriptors (adverbial/adjectival characterizations in the verbal accounts that are not contained in the physical enactments) [38], and story digests (decomposition of stories into individual idea units) [9, 38]. We adapted and integrated these in coding story enactment recordings for physical *actions*, *vocalizations*, and *enactment idea units*. An *action* is a minimal meaningful enactive action. Some examples are jumping, flying, hitting, jiggling a character figurine while it's walking, changing a character's orientation to talk to other characters (indicating "Who" [18]), moving to a specific location [33], and so on. *Vocalizations* are represented mainly by vocal descriptors. A vocal descriptor is an identified descriptive word or phrase in the video transcription, including adjectives, adverbs or phrases together with prepositions that function as adverbs indicating "When, Where, Why and How" [18], descriptive verbs like "sleep, run, faint, etc.", descriptive nouns like "love" in "a love story," and sound effects. An *enactment idea unit* is a single complete story idea that combines a sequence of actions and vocalizations.

Open coding was conducted by two coders on a randomly selected 20% of the total data. A final agreement level of approximately 90% was reached for both coding physical actions and idea units. Prior to analysis, both coders reached an agreement on the coding protocol for vocalizations. Then the data was divided into two halves and assigned to the respective coders.

Based on the segmented enactive idea units, we analyzed the degree of sophistication of these units in terms of the complexity of the physical actions and the richness of the accompanying vocalizations. Accordingly, we defined these two measurements: **Action complexity**, operationalized as the number of observed physical *actions* in the *enactment idea units*, and **Vocal richness**, the occurrence of identified *vocalizations (vocal descriptors)* in the *enactment idea units*. We further analyzed the relationship between complex physical actions and rich vocalizations per idea unit in the AR and TR conditions.

5.1.2 Statistical analysis results. The overall descriptive statistics showed little difference in the average numbers of action complexity (mean=1.1733, sd=.4850 in AR, mean=1.0025, sd=.4431 in TR) and vocal richness (mean=1.2636, sd=.3351 in AR, mean=1.1907, sd=.5284 in TR) across the AR and TR conditions. Moreover, ANOVAs [15] were performed to explore variations in actions, vocalizations, enactment idea units, action complexity, and vocal richness between the AR and TR conditions. No significant differences were observed in any of these aspects across the two conditions. We also compared different datasets based on age and gender but found no discernible differences between genders or age groups. However, one notable observation emerged regarding children aged 10 and older: their motivation to play with the figurines appeared to be relatively lower compared to younger children, particularly in the TR condition.

We conducted a Pearson correlation [10] to investigate the relationship between the total numbers of Actions and Vocalizations for each story and the Action complexity (actions per idea unit) and Vocal richness (vocal descriptors per idea unit). The results are displayed in Table 2. Notably, a statistically significant positive relationship between the total numbers of Actions and Vocal descriptors was found only in the AR condition with a higher correlation coefficient (r=+0.689, p=0.004) compared to the TR condition (r=+0.008, p=0.997) (see Table 2(left)). It's not hard to understand intuitively that as time progresses, more actions and vocalizations tend to occur, leading to an overall positive correlation between them. However, the result reveals that this positive correlation is only evident in the AR condition, not TR. This suggests that even under the influence of the time factor in the TR condition, the total numbers of actions and vocalizations remain significantly divergent.

Table 2: Two sets of Pearson Correlation results (**. Correlation is significant at the 0.01 level)

	Total Actions with Vocal descriptors		Action complexity with Vocal richness	
Overall	$r = +0.528^{**}$	$p = 0.003$	$r = -0.543^{**}$	$p = 0.002$
AR	$r = +0.689^{**}$	$p = 0.004$	$r = -0.364$	$p = 0.182$
TR	$r = +0.008$	$p = 0.997$	$r = -0.750^{**}$	$p = 0.001$

Furthermore, an overall (without separating AR and TR conditions) statistically significant negative relationship between action complexity and vocal richness was found with a large r value [11] of -0.543 (see Table 2(right)). This result implied a complementary relationship between the semantic loads of performing actions and speaking vocalizations in constructing story enactment idea units. On top of that, a statistically significant negative relationship between action complexity and vocal richness was only found in the TR condition (r=-0.750, p=0.001). These correlations further demonstrate that while constructing a story idea in TR, the number of physical actions decreases as the number of vocalizations increases, and vice versa. To elaborate, in the TR condition, the participants tended to express their ideas separately across modes (i.e., some idea units are expressed more richly through actions with fewer vocalizations, and others may be expressed more richly through vocalizations with fewer actions). On the contrary, the significant negative correlation doesn't exist in the AR condition, indicating that actions and vocalizations are used more together to construct ideas. Since the average numbers of complex idea units were similar across both TR and AR conditions, it can be inferred that the expression from the different modes was distributed in a more balanced way in the AR condition, which suggests a higher level of multimodality.

5.2 Writing Analysis

Similar coding analyses were performed in children's written story texts obtained from dictations or their hand-writings. For text analysis, we also coded for written idea units and descriptors. We divided the numbers of written idea units and descriptors by the total word count of the story text and defined two normalized measurements, *Written story complexity* and *Written story richness*, in assessing children's story writing performance. We also conducted a series of ANOVAs on story complexity and richness. No significant differences in any of the aspects between the AR and TR conditions were found in the written story texts. Some possible explanations for this finding are discussed in the limitations subsection of Section 6.

5.3 Enjoyment Analysis

We collected qualitative data, including pre- and post-questionnaires and post-interview question responses from all 18 pairs of children, 36 participants in total. No statistically significant differences were found in the questionnaires. For interviews, we focused on the answers to these questions: "Which story you wrote do you prefer, and why?" "Which approach (with or without using our system, AR or TR) do you prefer, and why?" and "What is your favorite part of today's experience?" We classified all their responses for reasons why they prefer one writing or one approach over another and their favorite experience during the whole participation into different categories, respectively.

We observed a higher preference for story writings in AR (20 out of 36) than in TR (15 out of 36) (see Table 3). The predominant reason categories are: Technology attraction (16/36), Better content (8/36), and Unspecified (7/36). In addition, there was a much higher preference for AR approach (32 out of 36) than TR approach (3 out of 36). And the three most commonly mentioned reason categories are: seeing the animation/like a movie (18/36), more imaginative information (11/36), and Unspecified (3/36) (see Figure3(b)). For favorite experience, the top three experiences that children liked the most are: Using TIPES system (19/36), Making stories (9/36), and Working with friends (4/36) (see Figure4).

6 DISCUSSION

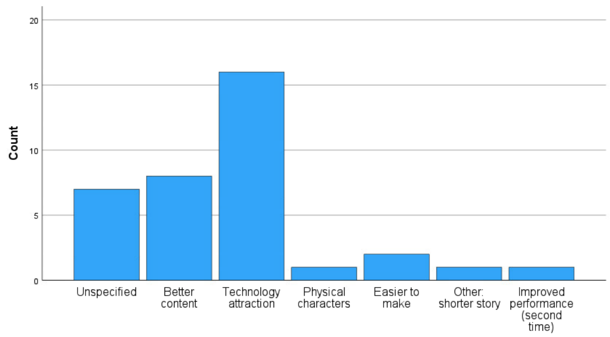
Interrelationship between actions and vocalizations. Although the writing outcomes are comparable across different conditions, we identified a distinction in the process of how children enacted their stories and built their ideas across the AR and TR conditions. Our findings suggest greater unimodal expression in either action or vocal modes in TR, and more reinforced co-expressions occurred in the AR condition. It is likely that the presence of animation in AR might enhance multimodality in enactment, which could imply a greater embodied imagination when utilizing TIPES.

Presently, we understand that our TIPES system engages children differently in story enactment, but the reasons behind it remain unknown. Hence, a more fine-grained analysis of the nuanced meanings embedded in actions and vocalizations is needed. So far, our analysis involves instances of actions and vocalizations. However, for future investigation, we could introduce subcategories for vocalizations, such as scene-related narratives, action-related descriptors, and dialogues. This analysis could yield two potential outcomes: first, a scenario with more actions coupled with more action-related vocal descriptors in one condition compared to the other, signifying a stronger activity enactment; second, a scenario with more verbal scene-related elements than physical actions in one condition compared to the other, indicating a better affordance for narration.

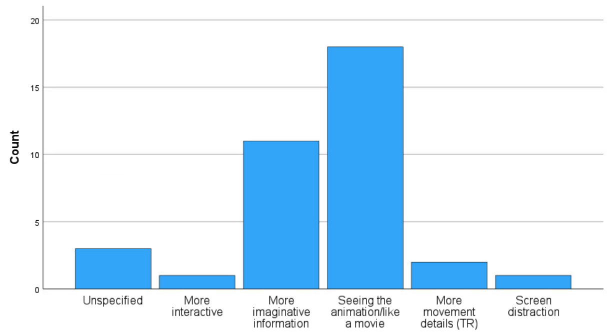
Greater enjoyment. Out of 36 children surveyed, 32 expressed a preference for the AR approach (1 child mentioned both, and 3 preferred TR). Three predominant reasons are: seeing things moving on the screen like a movie was interesting; being able to see the screen helped with how to set up scenes and move characters; and the real-time animated video provided more information that could help them imagine the story. As a point of note, among the 20 children who favored their story writings in AR (see Table 3), 14 pointed to technology attraction as the primary reason (see Figure 3(a)). Sample responses included "I like it because it's so cool, I

Table 3: Children's preference on story creation

Condition	Preferred written story	Preferred approach
AR	20 (55.6%)	32 (88.9%)
TR	15 (41.7%)	3 (8.3%)
Both	1 (2.8%)	1 (2.8%)



(a) Story writings



(b) Approach

Figure 3: Reasons for children's preferences

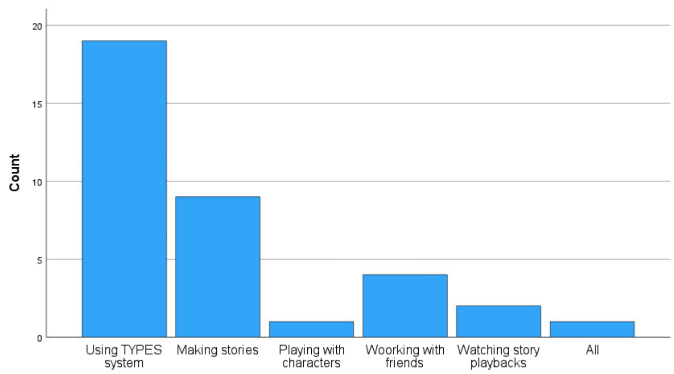


Figure 4: Children's favorite experience

can play with it the whole day" and "It was so fun that I could see things moving on the screen," indicating a fondness for our TIPES system. Their top-rated experience was being able to control the characters and see the animation on the screen using the TIPES system, providing them with a sense of novelty and enablement to make movies. Common responses also included enjoying the act of creating stories and collaborating with friends, showcasing their appreciation for creative activity and social interaction.

Limitations. Opinions from the three children who preferred the TR approach indicated that the AR approach might impose additional mental effort on children. There were concerns about accidentally blocking markers on top of figurines while playing and meticulous manipulation of the figurines to achieve desired virtual character movements. This could explain their preference for the story writings in TR because "it was easier to make" (see Figure 3(a)). Another child preferred watching the TR story recording because it accurately reflected their actions during story creation. From these responses, it can be inferred that the AR approach presented tracking restrictions (e.g., loss of tracking when markers are blocked or moved too quickly) and couldn't fully replicate all their manipulations in the animation.

Additionally, our study is not powerful enough to detect results in terms of story quality outcomes, partially due to the limited sample size. More importantly, we were working with a local community where our participants came from diverse backgrounds, encompassing too many variations in age, grade level, language preference, writing ability, pre-existing familiarity between the child pair, and so forth. For future studies, a potential strategy could be to recruit all participants from a single class to reduce variability.

7 CONCLUSION

This paper presented an approach by utilizing the TIPES system to support children's tabletop enactment in storytelling. We conducted a study with 18 pairs of children and analyzed their story enactment recordings, story writings, and interview responses to explore engagement differences in AR and TR conditions (i.e., with and without using TIPES). Statistically significant results indicate that our approach (AR) distinctively engages children in story enactment, altering the dynamics of their semantic loads compared to traditional tabletop play (TR). In their expression of an enactment idea, AR displays a relatively more equitable distribution between using multiple modes of physical actions and vocalizations, whereas TR exhibits a greater attendancy towards unimodality. We aim for this discovery to open up new directions to better understand children's multimodal enactment behaviors. Additionally, both qualitative and quantitative findings from interviews suggest that our approach contributes to greater enjoyment, which not only enriches children's immediate experience but also plays a crucial role in their overall development.

8 FUTURE WORK

In future endeavors, we will further explore the interrelationship of the actions and vocalizations as discussed in Section 6. Furthermore, we will investigate the impact of our TIPES system on children's embodied interaction while they are playing with the figurines from various dimensions. This investigation will include understanding how children immerse themselves in story-playing from different narrative perspectives (e.g., first-person or third-person). Other possible directions involve analyzing children's eye gazes while playing, improving the categorization of their diverse actions, and examining their conversations related to the system.

ACKNOWLEDGMENTS

This research has been partially supported by the National Science Foundation, Cyberlearning-1736225, DRL-1934173, and Grant-2008800.

REFERENCES

[1] Thijs Alofs, Mariët Theune, and Ivo Swartjes. 2012. A tabletop board game interface for multi-user interaction with a storytelling system. In *Intelligent Technologies for Interactive Entertainment: 4th International ICST Conference, INTETAIN 2011, Genova, Italy, May 25-27, 2011, Revised Selected Papers 4*. Springer, 123–128.
[2] Teresa M Amabile. 2018. *Creativity in context: Update to the social psychology of creativity*. Routledge.
[3] Craig A Anderson and Karen E Dill. 2000. Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of personality and social psychology* 78, 4 (2000), 772.
[4] Doris Bergen. 2002. The role of pretend play in children's cognitive development. *Early Childhood Research & Practice* 4, 1 (2002), n1.
[5] Tom Boellstorff. 2020. The ability of place: Digital topographies of the virtual human on Ethnographia Island. *Current Anthropology* 61, S21 (2020), S109–S122.
[6] Jerome S Bruner. 2009. *Actual minds, possible worlds*. Harvard university press.
[7] S Chu and Francis Quek. 2013. Maia: A methodology for assessing imagination in action. In *CHI 2013 Workshop on Evaluation Methods for Creativity Support Environments*.
[8] Sharon Lynn Chu, Francis Quek, and Kumar Sridharamurthy. 2015. Augmenting Children's Creative Self-Efficacy and Performance through Enactment-Based Animated Storytelling. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. 209–216.
[9] Sharon Lynn Chu, Francis Quek, and Theresa Jean Tanenbaum. 2013. Performative authoring: nurturing storytelling in children through imaginative enactment. In *Interactive Storytelling: 6th International Conference, ICIDS 2013, Istanbul, Turkey, November 6-9, 2013, Proceedings 6*. Springer, 144–155.
[10] Israel Cohen, Yiteng Huang, Jingdong Chen, Jacob Benesty, Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. *Noise reduction in speech processing* (2009), 1–4.
[11] Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences*. Academic press.
[12] Vincent Connelly, Julie E Dockrell, and Anna L Barnett. 2012. Children challenged by writing due to language and motor difficulties. *Past, present, and future contributions of cognitive writing research to cognitive psychology* (2012), 217–245.
[13] Jennifer Grouling Cover. 2014. *The creation of narrative in tabletop role-playing games*. McFarland.
[14] Adèle De Jager, Andrea Fogarty, Anna Tewson, Caroline Lenette, and Katherine M Boydell. 2017. Digital storytelling in research: A systematic review. *The Qualitative Report* 22, 10 (2017), 2548–2582.
[15] Ellen R Girden. 1992. *ANOVA: Repeated measures*. Number 84. sage.
[16] Steve Graham and Karen Harris. 2005. *Writing better: Effective strategies for teaching students with learning difficulties*. ERIC.
[17] I Granic, A Lobel, and RCME Engels. 2013. The Benefits of Playing Video Games. American Psychological Association, 69 (1), 66–78.
[18] VS Gupta. 2003. Elements of news story-5Ws and 1H. *Handbook of Reporting and Communication Skills* 25 (2003).
[19] John Helmes, Xiang Cao, Siân E Lindley, and Abigail Sellen. 2009. Developing the story: Designing an interactive storytelling application. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. 49–52.
[20] Carmen Juan, Raffaela Canu, and Miguel Giménez. 2008. Augmented reality interactive storytelling systems using tangible cubes for edutainment. In *2008 Eighth IEEE International Conference on Advanced Learning Technologies*. IEEE, 233–235.
[21] Nuri Kara, Cansu Cigdem Aydin, and Kursat Cagiltay. 2013. Investigating the activities of children toward a smart storytelling toy. *Journal of Educational Technology & Society* 16, 1 (2013), 28–43.
[22] Kiyoung Kim, Youngmin Park, and Woontack Woo. 2010. Digilog miniature: real-time, immersive, and interactive AR on miniatures. In *Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry*. 161–168.
[23] Ting Liu, Larry Powell, and Francis Quek. 2023. Catching Imagination: Enabling Children to Capture Imaginative Tabletop Play to Support Storytelling and Writing. In *2023 IEEE Frontiers in Education Conference (FIE)*. IEEE, 01–08.
[24] David McNeill. 2019. *Gesture and thought*. University of Chicago press.
[25] David McNeill, Susan Duncan, Amy Franklin, James Goss, Irene Kimbara, Fey Parrill, Haleema Welji, Lei Chen, Mary Harper, Francis Quek, et al. 2009. Mind merging. *Expressing oneself/expressing one's self: Communication, language, cognition, and identity: essays in honor of Robert Krauss (2009)* (2009), 143–164.
[26] OpenCV. [n. d.]. *Open Source Computer Vision Library: Aruco marker detection*. https://docs.opencv.org/4.x/d9/d6d/tutorial_table_of_content_aruco.html
[27] Hyung-Jun Park and Hee-Cheol Moon. 2011. AR-based tangible interaction using a finger fixture for digital handheld products. *Korean Journal of Computational Design and Engineering* 16, 1 (2011), 1–10.
[28] Jean Piaget. 2000. Piaget's theory of cognitive development. *Childhood cognitive development: The essential readings* 2 (2000), 33–47.
[29] Anne Marie Piper, Eileen O'Brien, Meredith Ringel Morris, and Terry Winograd. 2006. SIDES: a cooperative tabletop computer game for social skills development. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. 1–10.
[30] Francis Quek, David McNeill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E McCullough, and Rashid Ansari. 2002. Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction (TOCHI)* 9, 3 (2002), 171–193.
[31] Beatrice Rammstedt and Oliver P John. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality* 41, 1 (2007), 203–212.
[32] Yingjie Song, Nianmei Zhou, Qianhui Sun, Wei Gai, Juan Liu, Yulong Bian, Shijun Liu, Lizhen Cui, and Chenglei Yang. 2019. Mixed reality storytelling environments based on tangible user interface: Take origami as an example. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1167–1168.
[33] Leonard Talmy. 1975. Semantics and syntax of motion. In *Syntax and Semantics volume 4*. Brill, 181–238.
[34] Leonard Talmy. 2000. *Toward a Cognitive Semantics. Volume II, Typology and Process in Concept Structuring (Language, Speech, and Communication)*. MIT Press.
[35] Leonard Talmy. 2018. Fictive motion in language and 'ception'. In *Ten Lectures on Cognitive Semantics*. Brill, 125–162.
[36] Unity. [n. d.]. *Unity3D: Real-time Development Platform*. https://unity.com/
[37] Sharon Lynn Chu Yew Yee. 2015. *Performative authoring: Nurturing children's creativity and creative self-efficacy through digitally-augmented enactment-based storytelling*. Texas A&M University.

[38] Niloofar Zarei, Sharon Lynn Chu, Francis Quek, Nanjie'Jimmy' Rao, and Sarah Anne Brown. 2020. Investigating the Effects of Self-Avatars and Story-Relevant Avatars on Children's Creative Storytelling. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–11.