Interacting With Stories

Ashley Robinson, Chao Peng, Francis Quek, and Yong Cao
Center for Human-Computer Interaction, Virginia Tech
2202 Kraft Drive, Blacksburg, Va, USA
{arrobin, chaopeng, quek, yongcao}@cs.vt.edu

ABSTRACT
In today's media-saturated world, students are consuming media both actively and passively. To facilitate active interaction with media, we address a specific kind of audio-visual media interaction in which we call a hyper-drama. We address hyper-drama interaction preferences across two age groups: grades one to five and grades 6 to 9. These hyper-drama interactions include a token on a horizontal display versus mouse on a desktop display for story navigation, desktop display versus tablet display for scene viewing, and virtual buttons versus speech for character interaction and decision making within the hyper-drama. We conducted a within-subjects pilot study to evaluate these interaction techniques.

Categories and Subject Descriptors
K.3 [Computers and Education]: General; H.5.2 [User Interfaces]: Input devices and strategies

General Terms
Human Factors, Experimentation

Keywords
Storytelling, Hyper-drama, Interaction Techniques

1. INTRODUCTION AND CONTEXT
That new generations of children are growing up in ever increasing media saturation is more than an empty cliche [26, 28, 27, 33, 24, 23, 30]. A recent survey, conducted by the Kaiser Family Foundation [23], indicates that the children ages 8-18 are spending an average of 6.5 hours per day on various types of media, such as TV, movies, the internet, music, and video games. With multitasking (watching TV while chatting over the internet), children are exposed to the equivalent of 8.5 hours of media content daily. “Across the seven days of the week, that amount is the equivalent of a full-time job.” Of this, 4 hours and 26 minutes per day are spent on activities in which the audio-visual media content is passively consumed, including watching TV, DVDs and movies in cinema. In terms of exposure to the interactive media activities, such as Internet surfing, online chatting and playing video games, the duration is 1 hour and 51 minutes. Compared against a similar survey performed 5 years earlier [24], passive media maintained its domination of children’s daily life, while interactive media exposure increases dramatically (e.g. Computer usage increases from 27 minutes to 1 hour and 2 minutes). With the rapid progress of online services this ‘passive-active ratio’ may change with media consumption on the internet such as YouTube and online movies.

Our core research explores end-user authoring of a form of audio-visual media, in which we call hyper-drama, and its ability to facilitate creative imagination. A hyper-drama is a kind of hyper-narrative where each node of the narrative is presented in the form of an animated dramatic scene. This follows on the use of dramatic performance in education. For example, drama has been used to structure powerful learning contexts [9]. One might think that media exposure may have a negative effect on imagination, arguing that audio-visual presentation removes from the consumer the need to visualize mentally. Gerrig and Prentice [11] showed, however, that while children viewing audio-visual media may not have to imagine how the specific scenes or characters appear, they engage in imagination about the characters and the story as a whole. In fact, it may be argued that the whole point of sexual content in advertising is to engage product association with sexual fantasy [29]. Other studies along the same theme show mixed results concerning claims of imagination reduction [31, 32]. Jenkins [15] shows, furthermore, that new interactive, social, and communicative media that engage children as participants rather than passive consumers can have positive benefits in learning and literacy.

The engagement of children as active participants in media creation as opposed to passive consumers may, in fact, be the solution to the question of how one might support grounded imagination development in children. Two basic tenets of constructivist thinking about learning are that the learner has to be an active constructor of knowledge (as opposed to a passive receiver), and that this construction must draw its raw material from the prior experience of the learner [10]. In this, both Piaget (to whom the origin of constructivism is often credited) and Vygotsky would agree. An important component of creative imagination involvement is self-actualization and the self-expression that ac-
companies it [5]. Hence to support creative authoring, a means to present the resulting dramatic creations to the authors’ peers is important. We research this self-expression aspect by investigating age-related preferences in interaction techniques for hyper-dramas. Our subjects were children between grades 1 to 9. Our hyper-drama model supports multiple audience members, each of whom may follow a different character through a story graph.

In this paper, we investigate means of interacting with these hyper-dramas and how these interaction methods affect the process of creative imagination. The techniques that we are investigating are token on a horizontal display versus mouse on a desktop display for story navigation, desktop display versus tablet display for scene viewing, and virtual buttons versus speech for character interaction and decision making. We conducted a within-subjects pilot study comparing eight different combinations of techniques for interacting with a hyper-dramatic story. Our goals in this study are:

- To provide empirical guidelines for interacting with applications that contain multi-threaded story-lines, such as games
- To provide usage pros and cons for each interaction technique
- To address the question of novelty versus familiar interfaces and interactive convenience

STAGES, a hyper-drama storytelling system, is the testbed on which we are using to evaluate these various interaction techniques. While our study employs STAGES as a testbed, our results are extensible across a broader set of applications for active media interaction and storytelling. This includes applications such as online games, where there tend to be story-lines, plots, and other players with which the user interacts.

2. STORY INTERACTION

There are many ways in which people can interact with non-linear stories. These interactions include keyboard/mouse, gesture, and tangible interaction. Facade is a storytelling system that presents its story as a one-act play. Players interact with Artificially Intelligent agents with keyboard and mouse. The keyboard is used to talk to the agents and the mouse is used to interact with objects [17]. Kidsroom is a gesture-based storytelling system where children are physically able to move freely in the story space. Children navigate the story by making different gestures in different places of a play room, following the instructions of a virtual character [2]. Tangible Viewpoints is a tangible story interaction system. In Tangible Viewpoints, children use physical pawns to interact with the story. The pawns enable the children to experience the story from different character viewpoints [18]. Tangible interaction was included in the STAGES system because children can relate to physical object manipulation at an early stage in life [6].

3. HYPER-DRAMA APPROACH

3.1 Drama

There is long history for drama being performed in theater in the form of stage play and narrative dialogue. Dramatic performance and authoring has been used widely in education [4], and applied in such disparate areas as language learning [16] and ethnic studies [25]. Storytelling in dramatic form has been described as a technique for teaching [12, 13]. Robbins [22], for example, makes a compelling case of drama as a teaching tool in the language arts, summarizing various effective approaches for teaching and learning, and Edmundson [9] shows how drama can be used to structure powerful learning contexts. Bolton [3] further contrasts the vivid experience of drama as a concrete portrayal of abstract social and ethical concepts against the ‘accrual of facts’ typical of much of school instruction.

By employing drama as our medium of storytelling, we hope to ground the creative activity of our students within a larger cultural context. In our system, the students produce dramatic scripts that are performed by synthetic animated actors. The ‘stage set’ takes the form of a graphical backdrop. The scripted dialogue is enunciated by the actors using standard text-to-speech technology. This allows the students to enact their dramatic scripts immediately and as many times as they wish. The goal is to encourage further engagement by the student authors in the dramatic creations. “Is the resulting scene as they imagined?” “Would their characters actually say those words?” Our hypothesis is that if the students begin to explore these broader images beyond the specific dialogue they create, they will have engaged more comprehensively in the world they are constructing.

3.2 Hyper-Narrative

The hyper-narrative is the natural extension of hypertext and media to narrative [8, 19, 20, 21]. As a departure from the age-old linear story-line, hyper-narratives entertain multiple threaded stories in which the ‘reader’ participates in story-line selection and their decisions determine the outcome of the story. These stories are often presented solely as written text or with text and pictures. They differ from traditional linear stories in that hyper-narratives are written with multiple possible story arcs or paths through a story tree or graph. Hyper-narratives allow the ‘reader’ to experience these different narrative arcs as they navigate through the graph or tree. This enables information to be presented non-linearly, and gives the user control of the information they see and the order in which they see it. Hence, hyper-narratives allow the author to create virtual worlds full of characters, places, and events that interact in time and space as they would in the real world. This allows the reader to explore this world, choosing their own paths to create their own personal experiences [1]. Hyper-narratives also provide implicit structure because it flows in a logical order. Hence, they are able to communicate cause and effect by enabling the reader to make decisions that affect the outcome of the story.

3.3 Hyper-Dramas

In STAGES, the hyper-narrative is further extended to hyper-drama where each node of the story tree is occupied by a story fragment that may occur in one or more places. Each place is presented in the form of a dramatic animation. This differs from the static pictures and/or text found in hyper-narratives. A hyper-drama is a multi-threaded story presented as dramatic performances (full of characters, places, and events), where the story participant would choose their
4.1 Hardware

4. SYSTEM DESIGN RATIONAL

4.1 Hardware

4.1.1 Horizontal Display

The drama gameboard containing the story map is displayed on an interactive horizontal display table. It acts as a shared space between multiple users or ‘readers’ of the STAGES System. A significant weakness of hyper-text systems is the lack of readily perceptible structure. Without the aid of perceived structure, readers who are unfamiliar with the concepts of the text may become overwhelmed and find it difficult to maneuver through the links. This brings forth confusion about where they are in the network [7]. The concomitant concern with hyper-narratives is that they tend to result in disembodied presentation with respect to place because context is easily lost. The story map is used to ground the hyper-drama. Unlike hyper-text on the web, which can result in a disembodied presentation with result to place, the audience will always know their location within the context of the story. It also alerts the audience of where they should place their game token by surrounding the location with a red square. The horizontal display contains a 30 inch Apple monitor embedded in a wooden coffee table. This monitor was chosen because of its robust aluminum frame, which does not require extra support when placed inside of a table.

4.1.2 Tablet PC

A Tablet PC is used to display the dramatic scenes that correspond to hyper-drama nodes. It is the private domain for scene viewing. The portability and hand-held nature of the tablet PCs enhance the perception that the player is intimately involved with the goings on in the scenes, and that the story characters are interacting with them individually. It also permits the player to move around the game-board and view it without disengaging with the temporal events presented in the scenes.
4.1.3 Desktop Display

The desktop display along with the associated mouse and keyboard are most familiar to users. It is for this reason that we provide a means to interact with hyper-dramas using this technology. Sole desktop interaction would involve two monitors, where one would be used for contextual navigation, while the other is used for scene viewing. However, one monitor could also be used in combination with the other interaction devices.

4.2 Software

4.2.1 Speech Recognition

We used the Microsoft Speech SDK for speech recognition. In order for the system to distinguish between commands and general conversation, users were required to say the keyword, ‘select’ before saying the command. The purpose was to stabilize the system by reducing false-positives.

4.2.2 Character Animation

Our system is implemented with XNA in CSharp. XNA provides more flexibilities for integrating other interaction features. By using Windows .NET framework and XNA SDK, we integrate the game engine with text-to-speech, speech recognition and socket networking. The story is simulated with 2D graphical interaction and is separated by several fragments, each of which contains several scenes. Character actions are included in each scene.

By using XNA SDK, the 2D pictures are also easy to be imported into the story simulation.

4.2.3 Tracking

Vision-based technology, ARToolkit [14] (Augmented Reality Toolkit), is used to track the tokens on the game table. ARToolkit uses a camera to track unique markers, which are referred to as tokens. We chose to use ARToolkit because of its’ portability and low cost.

4.3 Interaction Design

The audience can interact with the hyper-drama using a game table, tangible interaction, a tablet PC, and/or desktop displays, which are displayed in Figure 1. We compare each component against the standard desktop with mouse with which students are typically familiar. In addition, we investigate the use of speech for making story choices. When interacting with the hyper-drama, the audience member first places a tracked game-piece/mouse on the first location on the story map (located on a horizontal display). This location is marked with a red square. The scene associated with that place appears on the tablet PC (in their hands) and they are able to see the characters interact with each other as well as with the audience. At some points in the story, the audience will be able to provide input (via speech or clicking a button). This input can determine the next scene for the story. Once a decision is made, and outcomes occur, a red square will mark the spot of the next scene. This cycle continues until the end of the story. Since this story is author-centric, this story can have as many paths as the author creates.

5. USER STUDY

The pilot study involved fifteen participants, grade levels 1-9. Eight participants were in grades 1-5 and seven participants were in grades 6-9. Participants were asked to view and interact with an age-appropriate, pre-written story, using the STAGES System. We have two conditions for dramatic scene interaction (desktop vs. tablet), two conditions for contextual map interaction (tabletop with token vs. desktop with mouse), and four conditions for story path selection (speech with tablet, speech with desktop, button and pen on tablet, and button and mouse on desktop), giving a total of eight conditions. We tested all eight possible combinations. The order in which students used various conditions was randomized. At the conclusion of each condition, students were given a questionnaire to complete.

To evaluate the data, we split the students into two groups by grade level and compared the results. Students in grades 1-5 are referred to as Group 1 and students in grades 6-9 are referred to as Group 2. Each condition was evaluated based on four major criteria: story enjoyment, ease of use, naturalness of use, and fun. We used a five-point Likert scale for each criteria. We also asked participants to discuss what they like most and least about interacting with the story. Since some younger students did not understand the concept of the word ‘natural’, it was explained as their ability to use the system without prior instruction (e.g. to figure it out on their own). At the conclusion of the study, we asked participants which interaction techniques they liked the most in terms of story navigation, scene viewing, and decision making. The following discusses our results.

**Story Enjoyment:** The average story rating was 4.57/5 points for Group 1 and 3.89/5 points for Group 2.

**Ease of Use:** Group 1 indicated the mouse, tablet, and speech were easier to use, while Group 2 preferred mouse, tablet, and button. Group 1 preferred to use the mouse over the token by 0.040 points (4.61 vs 4.57), while Group 2 preferred the same by 0.179 points (4.21 vs 4.04). Group 1 preferred to use the tablet over the desktop by 0.353 points (4.76 vs 4.41), while Group 2 preferred the same by 0.179 points (4.21 vs 4.04). Group 1 preferred the speech over the button by 0.201 points (4.69 vs 4.49), while Group 2 preferred button over speech by 0.107 points (4.18 vs 4.07).

**Naturalness of Use:** Group 1 indicated that mouse, desktop, and speech were more natural to use, while Group 2 preferred mouse, tablet, and button. Group 1 preferred to use the mouse over the token by 0.281 points (4.41 vs 4.13), while Group 2 preferred the same by 0.214 points (4.07 vs 3.86). Group 1 preferred to use the desktop over the tablet by 0.094 points (4.31 vs 4.22), while Group 2 preferred to use tablet over desktop by 0.286 points (4.11 vs 3.82). Group 1 preferred to use speech over button by 0.406 points (4.47 vs 4.06), while Group 2 preferred to use button over speech by 0.143 points (4.04 vs 3.89).

**Fun:** Group 1 indicated the token, tablet, and speech were more fun to use, while Group 2 preferred mouse and tablet with no preference towards button or speech. Group 1 preferred to use the token over the mouse by 0.085 points (4.87 vs 4.78), while Group 2 preferred to use the mouse over token by 0.286 points (4.21 vs 3.93). Group 1 preferred to use the tablet over desktop by 0.103 points (4.88 vs 4.77), while Group 2 preferred the same by 0.143 points (4.14 vs 4.00). Group 1 preferred to use speech over button by 0.228 points (4.94 vs 4.71), while Group 2 did not have a preference towards button or speech (4.07).

Table 1 shows a set of correlations to determine, for example, if the subjects evaluation of fun was related to the
degree to which they liked the story.

Table 1: Correlations – Group1/Group2

<table>
<thead>
<tr>
<th>Mouse use for story navigation on game-board</th>
<th>Ease of Use</th>
<th>Naturealns</th>
<th>Fun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Like</td>
<td>0.223/0.159</td>
<td>0.085/0.326</td>
<td>0.549/0.534</td>
</tr>
<tr>
<td>Ease of Use</td>
<td>1</td>
<td>0.419/0.305</td>
<td>0.365/0.344</td>
</tr>
<tr>
<td>Naturalness</td>
<td></td>
<td>1</td>
<td>0.678/0.545</td>
</tr>
<tr>
<td>Speech for story navigation on game-board</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Like</td>
<td>-0.003/0.256</td>
<td>0.262/0.356</td>
<td>-0.089/0.597</td>
</tr>
<tr>
<td>Ease of Use</td>
<td></td>
<td>0.419/0.305</td>
<td>0.665/0.413</td>
</tr>
</tbody>
</table>

6. DISCUSSION

Our results show a difference in preference of interaction techniques between children of different age groups. Group 1 preferred to use mouse (62.5%), desktop (62.5%), and speech (75%). Some children reported that having to reach across the horizontal table made token selection harder. The majority preferred the desktop for drama viewing although the tablet was marginally more fun. This was because the tablet was heavy for the young children, and most just placed it on the table – making it a desktop with a smaller screen. Some children said they preferred the desktop because the screen was bigger. Speech was preferred for decision making and was indicated to be more fun, easier, and more natural to use than the button. Students like the idea of talking to the computer and not having to physically click a button.

Group 2 preferred to use token (57.14%), tablet (57.14%), and button (57.14%). They preferred token interaction although they indicated that the mouse was marginally more fun, easier to use, and more natural to use. A primary factor for choosing the token was that it was cool and reminded them of an electronic board game. This indicates that Group 2 enjoyed moving their token but did not like it with the combination of other interaction techniques. The tablet was chosen by the majority for drama viewing, which was consistent with that fact that they indicated that the tablet was fun, easy, and more natural to use than the desktop. The primary reason for choosing the tablet was the fact that it is portable. Button use was preferred for decision making (buttons were judged marginally easier to use and more natural). The button was chosen primarily because it was more stable than the speech. Sometimes, speech would not respond correctly to a student’s voice, of which Group 2 was more sensitive.

As for the correlation analysis, we note that Group 1 had a large variance for ‘liking the story’ because it contained an argument/conflict that some did not like. However, they judged ease-of-use and fun highly when using speech, hence there was no correlation between liking the story with either measure. Likewise they thought speech was uniformly fun even with the variance in ‘naturalness’.

7. CONCLUSION/FUTURE WORK

This study shows that there is a difference in the preferred story interactions techniques amongst two distinct age groups, grades 1 to 5 and grades 6 to 9. The results of grades 1 to 5 indicate that they have fun using new and unfamiliar technology but prefer not to adapt to change. They are used to using a desktop computer and mouse and most of them indicated that they talk all the time. They seem to be so fascinated by familiarity that the glitches in the speech recognition system did not have much of an affect on this population. Since grades 6 to 9 are older, they have had more technology exposure and are somewhat used to adapting to different interaction techniques. They have the ability to look past the ‘cool’ factor of the technology and evaluate the total experience. In the future, we will increase the significance of the token. For example, the token will become a mechanism for making decisions rather than solely an outcome of the decision itself. We will also like to conduct a full study with an increase population of students.

8. REFERENCES


