Electric Agents: Combining Collaborative Mobile Augmented Reality and Web-Based Video to Reinvent Interactive Television

GLENDA REVELLE, University of Arkansas, Human Development and Family Sciences
EMILY REARDON and KRISTIN COOK, Sesame Workshop
LORI TAKEUCHI, Sesame Workshop, Joan Ganz Cooney Center
RAFAEL BALLAGAS, KOICHI MORI, HIROSHI HORII, HAYES RAFFLE,
MARIA SANDBERG, and MIRJANA SPASOJEVIC, Nokia Research Center, North America Lab

Electric Agents is a transmedia game that presents new ways for children to actively engage with television content. A typical-looking educational television show transforms into an interactive game in which children collaborate through a mobile augmented reality experience to find and collect vocabulary words that are missing from the show. The players return the words to the show by throwing them back into the television using their mobile devices. This blend of a linear video narrative and an interactive game strives to make educational television content more engaging and participatory while fostering collaborative play with vocabulary words. We describe the technical implementation to support this collaborative mobile augmented reality experience and report findings from a pilot user study. Results demonstrate patterns of collaborative activity, scaffolded learning, and parasocial relationships that have been linked in previous literature to educational benefits.

Categories and Subject Descriptors: H5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical user interfaces*

General Terms: Design, Human Factors

Additional Key Words and Phrases: Interactive television, mobile phones, gesture input, mobile augmented reality

ACM Reference Format:
Glenda Revelle, Emily Reardon, Kristin Cook, Lori Takeuchi, Rafael Ballagas, Koichi Mori, Hiroshi Horii, Hayes Raffle, Maria Sandberg, and Mirjana Spasojevic. 2014. Electric agents: Combining collaborative mobile augmented reality and Web-based video to reinvent interactive television. ACM Comput. Entertain. 12, 3, Article 1 (December 2014), 21 pages.
DOI: http://dx.doi.org/10.1145/2702109.2633413

1. INTRODUCTION

Studies show that the average American (age 15 and over) spends over half their leisure time watching television (2.8 hours of TV per day on average), much more than they spend socializing with friends [U.S. Department of Labor 2009]. Television dominates the media lives of younger Americans even more, totaling about 4.5 hours per day [Rideout et al. 2010]. In fact, total media consumption (including TV, music, games, etc.) of 8- to 18-year-old Americans is even more striking:

> On a typical day, 8- to 18-year-olds in this country spend more than 7.5 hours (7:38) using media—almost the equivalent of a full workday, except that they are using media seven days a week instead of five.

Author's email: Glenda Revelle: revelleg@gmail.com.

DOI: http://dx.doi.org/10.1145/2702109.2633413

Fig. 1. (Bottom left) Children use their communicators to capture words through a mobile augmented reality experience. (Center) Then they can throw the captured words back into the show.

> Moreover, since young people spend so much of that time using two or more media concurrently, they are actually exposed to more than 10.5 hours (10:45) of media content during that period. And this does not include time spent using the computer for schoolwork, or time spent texting or talking on a cell phone [Rideout et al. 2010].

Similarly high levels of media consumption have been reported around the globe [Kimura 2007].

Technology has the potential to change the way we consume television/video [Barkhuus and Brown 2009]. The Internet allows people to choose what they watch, as well as when they watch it. Based on data collected in 2009, the Kaiser Family Foundation reported that a total of 20 percent of young people's media consumption occurs on mobile devices, and that another 11 percent is "old" media (such as music or TV) consumed via new pathways (e.g., through a browser) on a computer [Rideout et al. 2010]. The same study reports a 29 percent multitasking proportion among 8- to 18-year-olds, which is the proportion of media time spent using more than one medium concurrently. Thus, it is likely that many of today's youth are already using mobile devices while watching television. More recent data indicate that as many as 75 percent of 2- to 9-year-olds now have access to a smartphone [PlayScience 2011], so the percentage accessing media over mobile devices is probably higher now. Both the Web and mobile platforms are highly interactive, but there are few examples of systems that make use of that interactivity with video content beyond simple program selection. How can we leverage interactivity to change the experience of watching television from a passive experience to an active, engaging interactive experience?

In our research, we've designed a new interactive, transmedia television experience to take advantage of the interactive opportunities inherent in new video-delivery platforms. Electric Agents is a prototype system that combines mobile augmented reality, embodied interaction and television content (see Figure 1). In Electric Agents, mobile augmented reality is used to facilitate social interaction and collaboration among players and to engage players in meaningful interaction with educational television content.

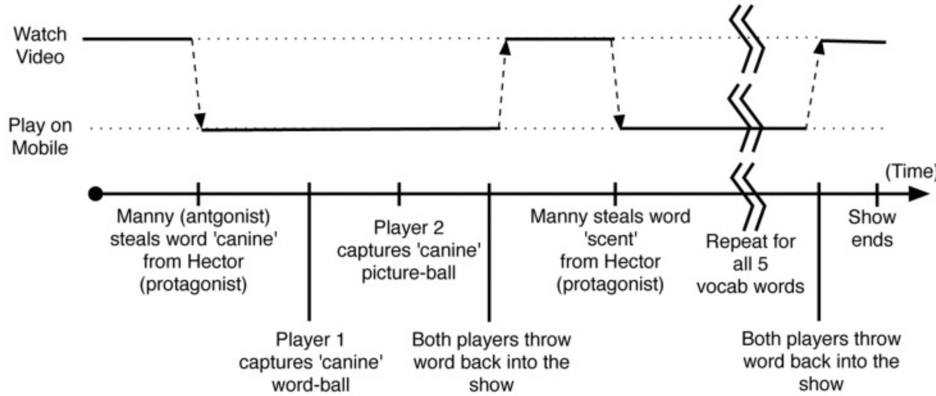

Fig. 2. The experience fluidly transitions between watching the video narrative and playing the mobile augmented reality experience. The video narrative is interrupted when Manny, the antagonist, steals words from Hector, the protagonist. The show resumes after players successfully retrieve a word from an augmented reality view and return the word to the show using throwing gestures. This flow repeats for all vocabulary words (5 in total) in the curriculum for each episode.

2. ELECTRIC AGENTS

The Electric Agents experience is designed around the educational television series "The Electric Company," a recent winner of five Emmy awards including Outstanding Children's Series and New Approaches – Daytime Children's for "The Electric Company" website.[1] "The Electric Company" (TEC) uses the powerful medium of television narrative to encourage the development of literacy skills among children ages 6 to 10 years old. The show is aimed primarily at populations of children who are not succeeding in the classroom. TEC has a curriculum that includes phonics, vocabulary, and comprehension of connected text. In this research, we build on the TEC narrative and curriculum by adding a game-like element in which the child viewers become Electric Agents, or agents of "The Electric Company" team. As Electric Agents, children engage in tasks to help TEC characters in the narrative, creating the illusion that the viewers are part of the show. Electric Agents consists of a multi-player game layered on top of an existing TEC television episode. The game experience is broken up into alternating phases of watching the show narrative and actively playing the game. Figure 2 diagrams the experience flow during different phases of gameplay.

During the watch phase, the experience is very similar to traditional video viewing, except the video is streamed from the Web. The handheld devices are inactive during this phase, displaying "The Electric Company" logo (see Figure 3). At various points during the show Manny, an antagonist and member of the Pranksters, interrupts viewing using a special gadget to zap words out of the mouth of Hector, a protagonist and member of "The Electric Company" (see Figure 4). This marks the transition from the passive watching phase to the active playing phase using the mobile devices in the Electric Agents experience. The show cannot continue until the players return the missing words back to Hector.

Just after the word is stolen, an onscreen member of "The Electric Company" appeals directly to the viewers to help Hector get his words back using their communicators. The mobile devices (referred to as communicators in the game narrative) activate, showing an augmented reality window. The mobile devices serve as a window to a shared virtual space revealing word-balls (similar to the word-balls featured prominently on "The

[1]http://pbskids.org/electriccompany/.

Fig. 3. A view of the controller in the ‘watch’ mode. The controller is inactive and displays a passive logo for the show.

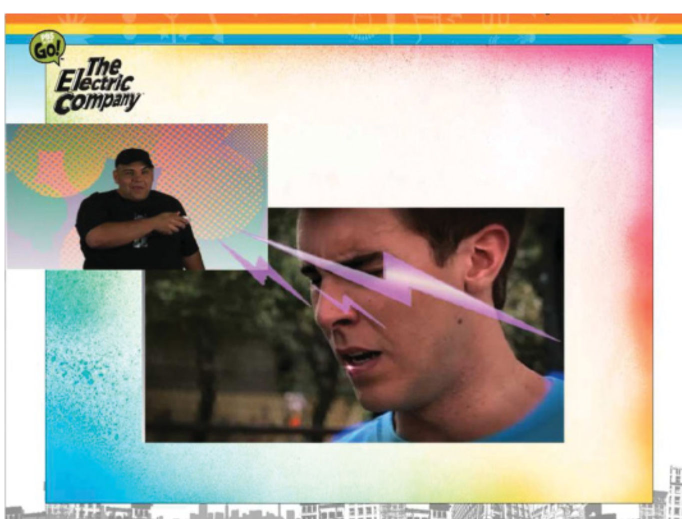

Fig. 4. Manny interrupts the show to steal words out of Hector’s mouth, marking the transition from the passive watching phase to the active playing phase of the experience.

Electric Company” show) and picture-balls floating through the air (see Figure 5). There is a word and picture pair to represent each vocabulary word from the show curriculum. Although both players can see all of the floating balls, they are each assigned a specific role: Agent Red (holding the red device) can only capture red word-balls, and Agent Blue (holding the blue device) can only capture blue picture-balls. Players have to work together to capture matching words and pictures using their communicator devices. To capture a word- or picture-ball, players must align the ball within the rectangular target region on their mobile display (see Figure 6). When the target object is properly

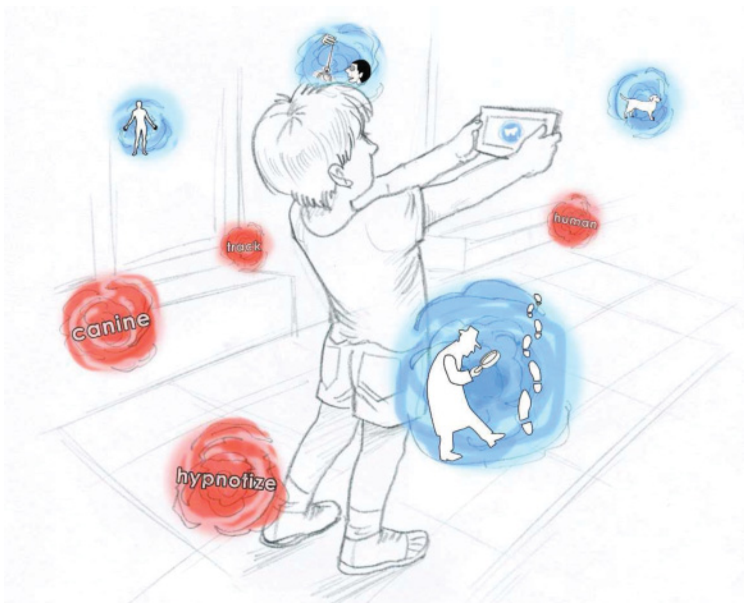

Fig. 5. Artistic rendering depicting the virtual word-balls and picture-balls in the virtual space shared by both players.

Fig. 6. An image of the actual view of the scene through the communicator of Agent Blue during gameplay. In this view the blue picture-ball for *canine* and a partial view of the red word ball for *human* and *scent* are visible.

aligned, the target region blinks and the communicator beeps. Once this alignment is achieved, the player presses a trigger mechanism (physically located on the back of the communicator housing) to capture the floating ball. The ball then gets “sucked into” the communicator, taking over the player’s screen (see Figure 7).

Fig. 7. After the matching word and picture are captured, a textual reminder to 'Throw the ball now!' is displayed on screen below the ball.

Once the balls are captured, players return the words back to the show by throwing the word- and picture-balls towards the screen (see Figure 8). Conjuring and throwing word-balls is a special power of "The Electric Company" characters in the show. In Electric Agents, players are afforded the same special powers through their "communicators".

If a player throws an incorrect word, the game offers progressive hints for finding the correct word. Initially, a text hint is provided showing the context of the word within a fill-in-the-blank clue. For example, when Manny steals the word 'canine' from Hector the text of his spoken sentence appears on the screen: "------ *like... like a dog?*" (see Figure 9). After an incorrect word is thrown to the screen, Jessica (a member of the Electric Company team) comes full-screen on the television to give players a hint *"I don't think that was the right word, try another one."* If another incorrect word is thrown, Jessica returns to give an age-appropriate definition for the word: *"Hector needs a word that means an animal in the dog family, like a pet poodle, a wolf or a coyote."* These progressive hints scaffold players to return the correct missing word so that the show can continue.

If the players correctly return the missing word and its associated picture, the episode advances after a short video payoff (see Figure 9). The payoff is an animation that is playful and used commonly as a transition between show elements in the television program. During this transition, the communicator goes back to a dormant state showing only a static logo for the Electric Company. This pattern repeats for all five vocabulary words in the episode curriculum.

Note that in the Electric Agents design no points or badges are awarded for correct answers. The reward for correctly returning words is the players' success in helping the TEC team and advancing to the next piece of the episode narrative.

The main contribution of Electric Agents is to demonstrate the first example of a transmedia system that combines mobile augmented reality with television episodes. Interaction through mobile devices allows the traditional educational television content

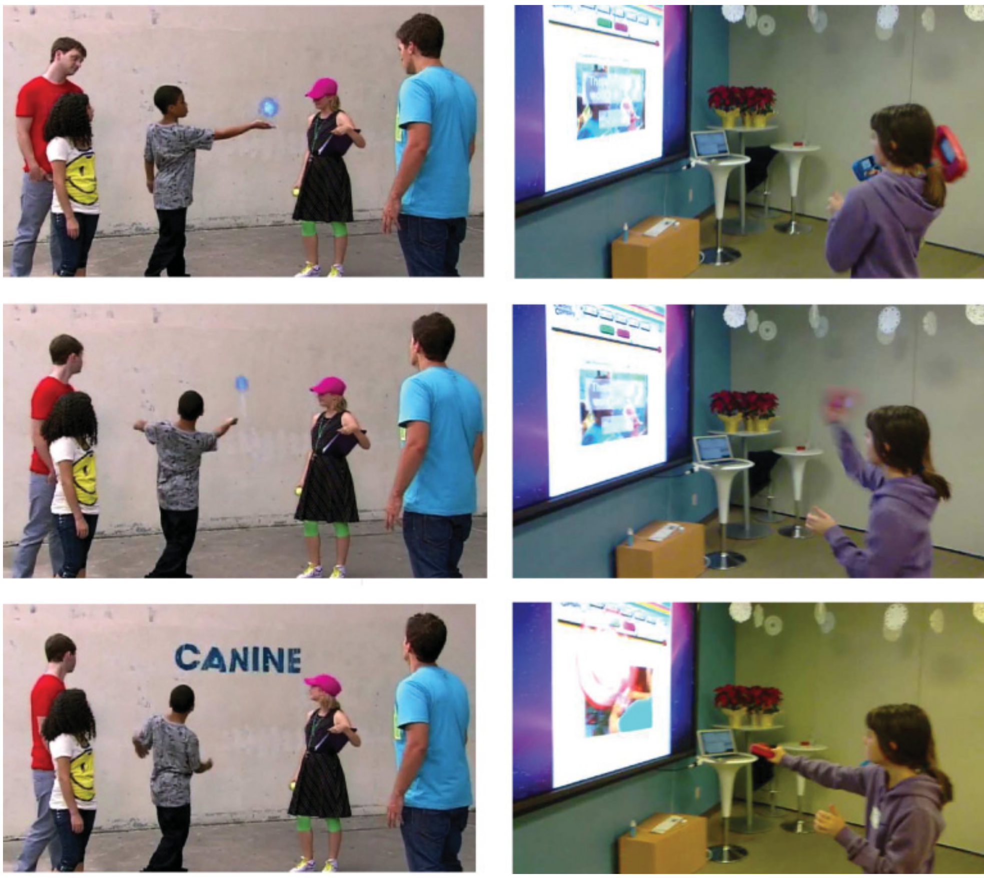

Fig. 8. Player activities in the game mimic abilities of the characters in the show. (Left) One of the actors in the Electric Company TV Show conjures a word-ball and throws it against the wall. (Right) During playtesting, one of the game players throws a captured word-ball back into the show. (Bottom-right) The thrown word-ball splats onto the screen.

to be transformed into an active and collaborative experience. A pilot user study of the system with children demonstrates patterns of collaborative activity, scaffolded learning, and parasocial relationships that have been linked in previous literature to educational benefits.

3. RELATED WORK

3.1. Related Mobile Augmented Reality Games

Studies of social interaction in multiplayer handheld gaming experiences have found that players tend to sit together, but focus on their own devices [Szentgyorgyi et al. 2008]. However, research on handheld mobile augmented reality has demonstrated potential to support co-located social play. AR Tennis [Henrysson et al. 2005] allows players to use their devices as both viewports and paddles to engage in a virtual tennis game. Markers on both handheld devices and the tabletop create shared perspective for the collaboration. Cows vs. Aliens [Mulloni et al. 2008] explores tradeoffs between mobility and social involvement. Bragfish [Xu et al. 2008] also uses a shared physical board augmented with visual markers to create a shared playing space. Art of Defense [Huynh et al. 2009] combines tangible manipulation of game board pieces

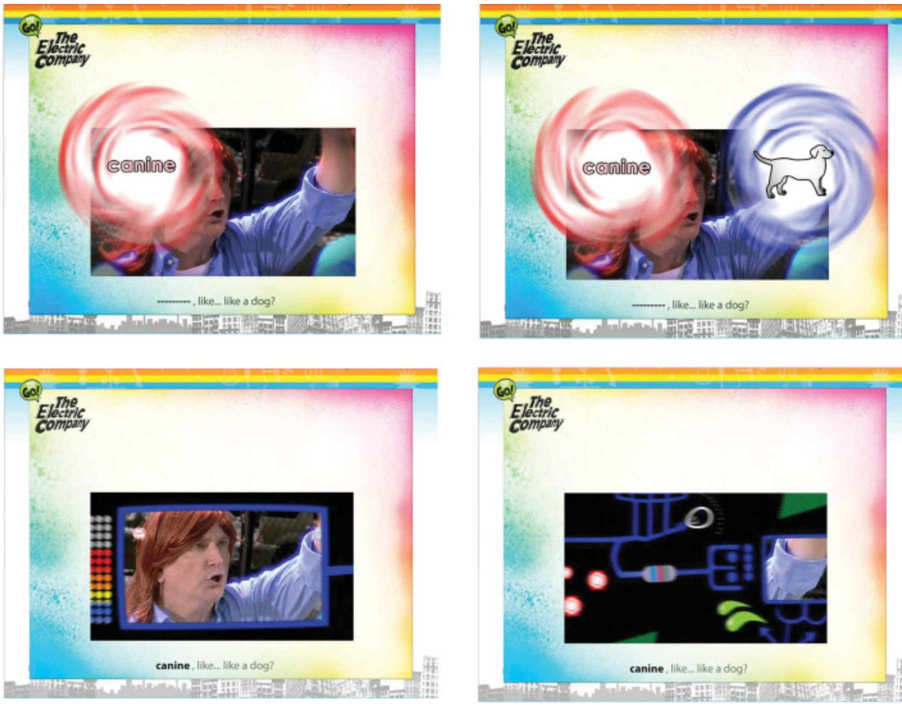

Fig. 9. (Top left) Agent Red throws the word *canine* up to the screen resulting in a splat of the wordball over the video. Notice a text hint below the video ("- - - - - -, like. . . like a dog?"). (Top right) Agent Blue throws the matching picture-ball resulting in a second splat. (Bottom left, right) If the word correctly matches the scene, a payoff in the form of an animated transition to the next scene is displayed.

with handheld augmented reality. Several benefits of handheld augmented reality games have been identified including increased social awareness and increased social interaction during gameplay [Huynh et al. 2009; Xu et al. 2008]. Similar to these systems, Electric Agents uses handheld augmented reality to promote co-located social play.

The Electric Agents augmented reality system differs from some previous systems in a number of ways. Unlike "augmented reality educational gaming" systems designed for classroom settings (such as Environmental Detectives [Klopfer and Squire 2007]), Electric Agents is specifically designed to be used in the home so that gameplay is interwoven into the everyday existing rituals [Bell et al. 2006] of media-entrenched youth. In contrast to previous mixed reality experiences that engage players in learning about historical sites through location-based narratives (e.g., Archeoguide [Vlahakis et al. 2002], Voices of Oakland [Dow et al. 2005], REXplorer [Ballagas et al. 2008]), the Electric Agents mobile augmented reality system is specifically designed to be location agnostic and function in any living area.

3.2. Related Interactive Television Experiences

Several historical television shows have reached beyond the "fourth wall," explicitly engaging viewers to participate in scaffolded interactive activities during viewing. Examples of this include the children's shows "Winky Dink and You" and "Picture Pages," in which viewers are invited to engage in drawing and learning activities throughout the show.

“Winky Dink and You” is widely considered to be the first interactive television show [Gawlinski 2003]. It originally aired on CBS from 1953 to 1957. Children were equipped with special Winky Dink Kits at home. The kits retailed for 50 cents and included a Magic Window (a vinyl sheet that was attached to the television screen using static electricity), Magic Crayons, and an erasing cloth. In every episode, Winky would arrive at a scene that contained a connect-the-dots puzzle. He would then prompt the children at home to help him complete the picture by drawing something that would help him continue the story (e.g., a cage to trap a dangerous lion or a bridge to cross a river). Another use of the Magic Window was to decode secret messages—an image would be displayed, showing only the vertical lines of the letters of the message, which viewers at home were asked to trace onto their Magic Window, and then a second image would display the horizontal lines to complete the text.

“Picture Pages”[2] was an educational segment that debuted nationally on Captain Kangaroo and aired on CBS from 1978 to 1984. Captain Kangaroo would use a magic drawing board (and later, from 1980 to 1984, Bill Cosby would use a musical marker named Mortimer Ichibod) to teach children basic arithmetic, geometry, and drawing through a series of structured activities. During the broadcast, children could follow along with the lesson using a workbook that was given away at local supermarkets. Electric Agents extends this history of using television to scaffold interactive learning activities and further enables technologically interactive experiences.

Children’s media are still leading the way in bridging the gap between traditional television and interactive media. Grimes [2008] examined the expansion of three children’s television programs into the world of massively multiplayer online games (MMOGs): (Nickelodeon’s “Nicktropolis,” Cartoon Network’s “Big Fat Awesome House Party,” and Corus Entertainment’s “GalaXseeds”). In addition, PBS has migrated almost all of its children’s programming from television to include online and mobile platforms [PBSKids Raising Readers 2011].

3.3. Related Educational Literature

Children’s media producers have been using television to educate and entertain children for decades, starting with programs like “Sesame Street” and “Mister Rogers’ Neighborhood” in the 1960s in the U.S., the effects of which have since been reproduced around the world. It is well established that children learn from watching well-designed educational television programs [Calvert and Wilson 2011; Comstock and Scharrer 2007]. Moreover, educational research is beginning to document that learning gains are greater when multiple media platforms are used than when any one medium is used alone (e.g., video narrative combined with interactive Web game [Fisch et al. 2011]). Electric Agents uses proven educational television content combined with mobile interactive vocabulary activities, so there is reason to suspect that this system may have similar educational benefits.

Children’s programs like “Sesame Street” and “Dora the Explorer” foster interactivity by leveraging parasocial relationships (mock social interaction with onscreen characters) to encourage children to treat characters on the show as social partners [Reeves and Nass 1996; Richter et al. 2011]. In a recent study of “Dora the Explorer,” preschool children who interacted more with Dora while Dora was onscreen were more likely to understand the content of the story [Calvert et al. 2007]. Moreover, children were more likely to learn a problem-solving strategy when they perceived themselves as more similar to Dora. Older children by the age of 7 are even more likely to form parasocial relationships with characters on screen [Hoffner 1996]. Parasocial relationships are

[2]http://www.tvguide.com/tvshows/cosbys-picture-pages/200073.

characterized by the ways in which children respond to onscreen figures as if they were actually in a social relationship (e.g., feelings of companionship, empathy). These parasocial relationships and feelings of similarity are linked to attention and learning gains [Bandura 1989; Fisch 2004]. These findings considered together suggest that children are most likely to learn from onscreen characters when they identify with them or treat them as social partners [Richter et al. 2011]. In fact, a recent study comparing young children's learning when a socially meaningful character, rather than an unfamiliar character, demonstrated the task found that even very young children (age 2 and under) can learn cognitive skills from television, but only when the onscreen character is socially meaningful to them [Lauricella et al. 2011].

Electric Agents encourages feelings of similarity and parasocial relationships with the protagonists on the show by inviting viewers to become part of the Electric Company, and giving them the same powers (i.e., the ability to capture and throw word-balls) as the characters from the show. In addition, Manny and Jessica speak directly to the camera, as if in direct social interaction with the child, and Jessica refers to elements in the child's environment (the communicators, the floating word and picture balls), further supporting the conceit of an actual social interaction.

Embodied cognition is the notion that the nature of human thinking and learning is determined in large part by the physical form of human bodies. Embodied cognition theory (e.g., [Wilson 2002]) posits that cognition is mediated by the body because cognition is situated in a real-world environment, cognition is time-pressured by interaction with the environment, and that the flow of information between cognitive systems and the physical environment is always ongoing. Another assumption of this theory is that the human mind has evolved to learn in interaction with the environment, as seen in infants' and toddlers' coordination of sensory inputs with motor responses they were born with as a central part of cognitive development [Piaget and Inhelder 1969].

Thus, embodied interaction (using natural physical body movements as a means of interacting with technology) has interesting potential to support learning and cognitive growth. Lakoff and Núñez [2000] argue that quantity is spatial extension, and Howison et al. [2011] have built on this idea by creating an embodied-interaction system for helping young students learn about proportional equivalence. Similarly, Cress et al. [2010] supported young children's learning of relative numerical magnitudes using whole body gestures on a number line projected onto the floor, and found that children's learning in this condition was significantly greater than for children who used a tablet PC and stylus for the same task. In another domain, Antle et al. [2009] created a system that enables children to learn musical concepts by moving their bodies. In addition to supporting learning, full-body physical movement has also been linked in the literature to increased engagement [Reeves and Nass 1996; Richter et al. 2011] and social interaction [Bianchi-Berthouze et al. 2007; Isbister and Dimauro 2011]. The Electric Agents game design leverages large-scale motor movement to engage children in social play and learning about vocabulary words.

Finally, children learn from each other when engaged in cooperative or collaborative activities [Rogoff 1990; Slavin 1996]. Recognizing the importance of this kind of social learning, researchers have contributed numerous analyses of cooperative interactions around games designed for learning using computers, handhelds, and other devices [McConnell 2000; Roschelle et al. 2010]. In addition, research suggests that well-designed games can make remarkably good learning environments because they embody many important learning principles [Gee 2003; Prensky 2005], and there is some initial evidence to suggest that children are able to transfer knowledge from games to the real world [Stevens et al. 2007]. Electric Agents is a two-player game that is specifically designed to support collaborative play and cooperative learning.

Fig. 10. Form factor of the communicator. (Left) Each communicator glows to represent the role of the player, either Agent Red or Agent Blue. (Right) A clear version of the housing makes the Ariane sensor box visible to the right of the N900.

4. TECHNICAL ARCHITECTURE AND IMPLEMENTATION

4.1. The Communicators

Although the basis of the communicator is a commercially available Nokia N900 smartphone, the housing is distinctive and transforms the look and feel of the device (see Figure 10). The communicators were designed to support suspension of disbelief—we wanted players to believe that these devices were more than "just a smartphone," but rather, special communicators with properties related to the game narrative.

Conceptually, the housing serves a similar purpose as plastic accessories for the Nintendo Wii controller (e.g., the attachment transforming the Wii controller to look like a tennis racquet). However, our housing is not passive, and contains several active elements. First, the controllers use LEDs to create a striking glow that helps players remember their roles as Agent Red or Agent Blue. Additionally, the controllers contain an Ariane sensor box, which is a wireless sensor module featuring many sensors. Notably for this project, it contains a 3-axis magnetometer, compensating for the N900's lack of an onboard compass orientation sensor. We also use buttons on the Ariane sensor for the trigger mechanism of the controller, allowing users to comfortably grasp the housing with the trigger located on the back of the device. The Ariane sensor module communicates orientation and button information to the Nokia N900 mobile phone over Bluetooth. Since many mobile phones today have compass orientation sensors, the game could be played without the supplemental housing as long as one is willing to sacrifice the benefits of the distinctive form factor.

AssaultCube—a screenshot of the game is shown in Figure 3a—is an open source, first person shooter game based on the Cube engine and game. It debuted in 2004 with the name ActionCube, a member of the Cube community was behind its launch. The official release date was in November 2006, and in May 2007 the name was changed for AssaultCube to avoid ambiguity with the name of another game, Action Quake. The game is a more realistic version of the original Cube game, which keeps the simplicity and velocity of the original one.

4.2. Mobile Augmented Reality

Our mobile augmented reality solution is different from previous handheld augmented reality games in that it relies solely on compass orientation and does not require visual markers or geographic location to create a shared play space. Each ball is situated at a fixed compass orientation as depicted as a 2-D plane in Figure 11. This 2-D plane is navigated using horizontal and vertical orientation of the device. Tilting the device up and down corresponds to translation on the vertical orientation axis. As the player spins around, the horizontal orientation of the device changes, corresponding to translation

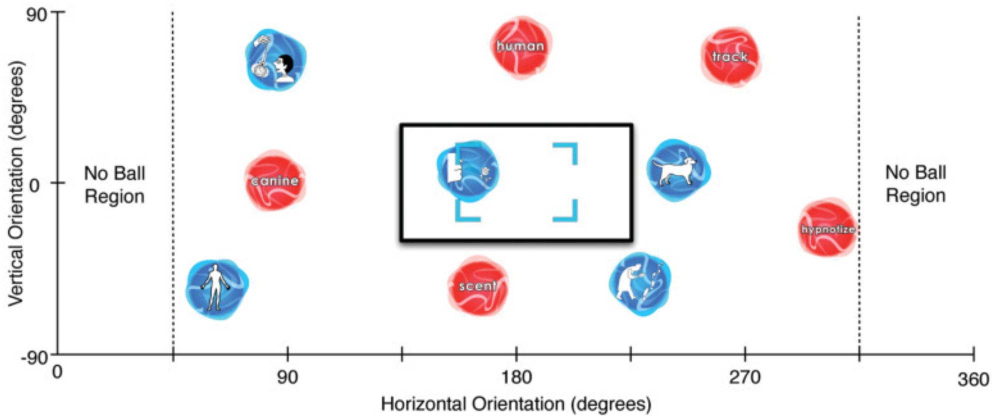

Fig. 11. Orientation map for the word-ball and picture-ball targets is consistent for both Agent Red and Agent Blue. Each target is positioned at a fixed location. The targets are spaced such that there is almost always at least one word-ball on the screen, but no way to include multiple word-balls in the target region. The No Ball Region is exactly the width of the device and allows the orientation map to wrap around without visual disruption as the user moves the orientation of the device across the 360-degree boundary.

of the view along the horizontal axis of the orientation map. The location of the targets in the orientation map is consistent for both players, allowing them to help each other find appropriate targets.

The target map was built using Qt Quick, a declarative cross-platform UI toolchain that makes it easy to specify interfaces with fluid animations. We take advantage of the platform animation support in two ways: first to fluidly animate translation of the orientation map between disjoint orientation sensor readings, and second, to animate each word-ball such that the background ball graphics spin behind the word or picture to approximate the way word-balls spin in the show. We are able to overlay the Qt Quick canvas on top of the live camera video stream at high frame rates (>20 fps on N900 hardware).

4.3. High-Level Architecture Considerations

Our prototype architecture is based on the assumption that television will be delivered through the Internet. The Web browser is becoming ubiquitous on devices connected to the television, especially gaming consoles. In addition, some high-end television or set-top-boxes already have Web browsing capabilities. Electric Agents was designed to leverage the fact that Web technology is more commonly available on televisions, and is built to run in a Web browser.

In order to use mobile phones to interact with the video content across devices in the way we envisioned, real-time communication is required. Instead of peer-to-peer connections, our architecture routes communication through the Internet (see Figure 12), simplifying connection-pairing processes. The messages are routed through the server based on user login from the phone and the website. This login information can be stored in browser cookies to eliminate the need to log in every play session. An additional benefit with the server-client approach is that the server can absorb potential constraints and incompatibilities in the messaging between the Web browser and smartphones.

4.4. Implementation Details

The interactive video episode in Electric Agents consists of the sequentially played main video content from a TEC episode, some pop-up narrative video content with

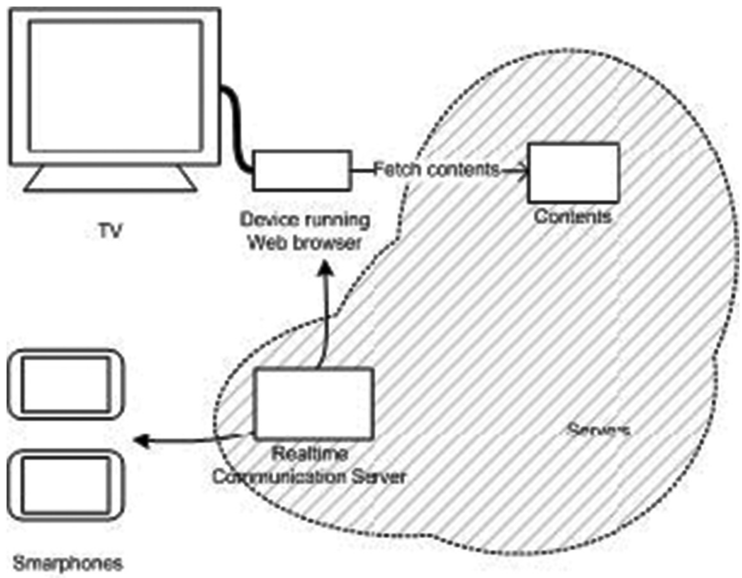

Fig. 12. High-level architecture of our Web-based system implementation.

transition animations, and some visual effects. We chose to use Flash to create the interactive video episode content over HTML5 because it is a powerful tool when creating interactive content involving animated graphics and video with a good tool chain, API stability, good integration of videos, and browser compatibility. HTML5 and related standard technologies are certainly evolving as an alternative for building interactive, rich media content and could be suitable to building this kind of interactive video system in the future. But the technology is still young, and we felt that Flash was better suited for the requirements of this project at the present.

The primary way to realize real-time communication with Flash is using the Flash Media Server. In order to create the real-time communication between the smartphones and the interactive video episode in Flash, we wanted the application on each of the smartphones to make a connection with the Flash Media Server. However, Flash Media Server's security policy prevented this. We overcame this limitation by providing a proxy server that bridges communication from/to the application running on the smartphones.

4.5. Gesture Recognition

In Electric Agents, the only gesture used is throwing the captured ball back to the screen. We first tried a very complex motion recognition library that can precisely distinguish motion using dynamic time warping and template matching (similar to [Akl and Valaee 2010]). However, in our early tests with children, we found that each child had different throwing motions, and, therefore, template matching was not well suited. Instead, we applied a very simple heuristic using a low-pass filter and threshold values for movement (see Algorithm 1).

Basically, if $|z_{avg}|$ is low (below a constant threshold value *threshOrientation*), the gesture started in an orientation with the mobile screen roughly parallel to the television screen (perpendicular to the floor). Then if the instantaneous z accelerometer reading is much higher than the running average (more than constant threshold value *threshThrow*), we assume the user is performing a throw gesture. We were able to use

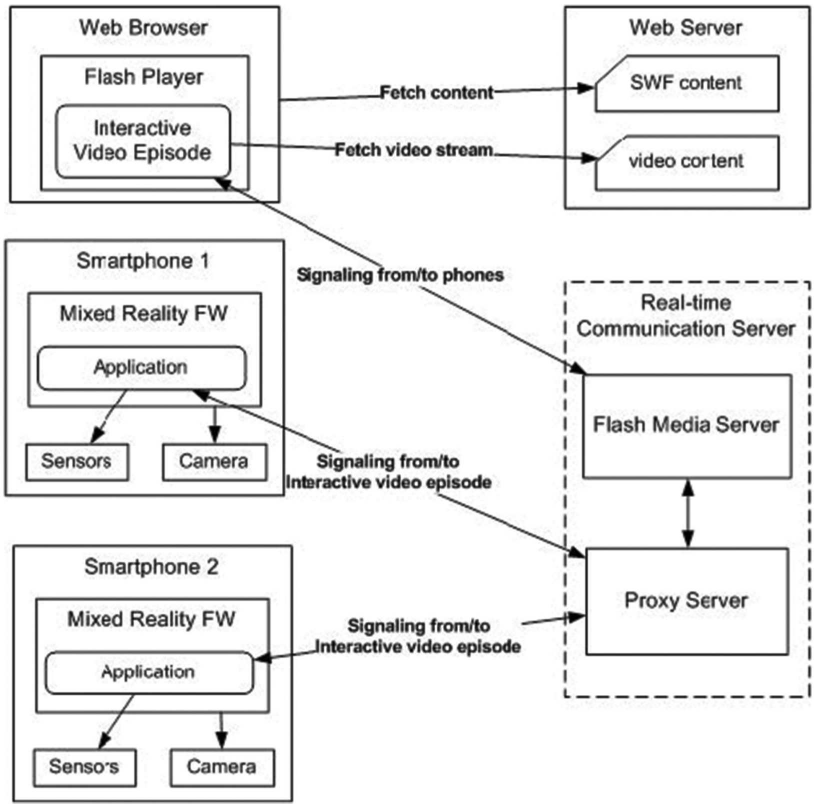

Fig. 13. Detailed architecture of our system implementation.

ALGORITHM 1: Simple Throw Gesture Recognition Algorithm

Given accelerometer readings *x,y,z* where *z* is the dimension coming out of the screen
z_t is the accelerometer reading at sample *t*.
$z_{avg} = \text{average}(z_t, z_{t-1}, z_{t-2}, z_{t-3}, z_{t-4})$
if $|z_{avg}| < threshOrientation$ **and** $z_t > (threshThrow + |z_{avg}|)$ **then**
isThrow = *true*

---

such a simple algorithm because in our application, the gesture recognition algorithm was activated only under special circumstances (after players captured a ball). Also our application required only a single gesture, so there is no possibility of confusing multiple gestures. This simple approach worked robustly against a variety of throwing motions and resulted in an 88 percent (N = 93) gesture recognition rate in our user trials.

The ball-throwing interaction is where network latency potentially becomes an issue, because it involves a message passing over the real-time communication channel in between the user's motion and the visual feedback on the television screen. In an informal study, we also tested on the regular high-speed Internet (DSL), where the average latency for networking from the real-time communication server from the smartphone was about 150 msec. Using video recordings to calculate the time from the end of the throwing motion to the appearance of the word-ball on the television, we found an average latency of around 240 msec. This perceived latency is likely lower than the actual communication latency because our algorithm recognizes the gesture earlier in the throwing motion. Also in our design, the players expect some latency due to the analogy of physical propagation of the ball from the phone to the screen.

5. PILOT USER EVALUATION

In order to validate our design decisions, we conducted a pilot evaluation of the user experience. We took a prototype of our system to a Boys' & Girls' club in the East Harlem neighborhood of New York City. We recruited 34 children (17 pairs of players), ranging from 6 to 10 years old, in order to match the target audience of "The Electric Company" show. We grouped all children in same-age, same-gender pairs. Videos were recorded of the play sessions and subsequent interviews. We provide here a descriptive analysis of emergent patterns of interaction.

5.1. Collaboration

In our trial, we observed players collaborating by helping each other find word-balls, discussing which word or picture to look for, and encouraging each other.

[Pair 8, girls 8 years old]
Hector (on screen): *It's the [Manny zap] of Annie Scrambler.*
Jessica (on screen): *Throw Hector a word-ball now!*
Agent Red: *OK, what do we have to do now?*
Agent Blue: *OK, I think scent.*
[Both players looking around the room using mobile augmented reality.]
Agent Blue: *Scent, scent... scent*
Agent Red: *Scent*
Agent Blue: *I saw it here somewhere... Here's scent!*
[Agent Blue throws scent picture-ball to the screen]
Agent Red: *Man... I keep on finding human.*
Agent Blue: *I, oh, just try to find scent. [...] I found scent, it's right next to the dog* [referring to the dog picture-ball]*... and hypnotize*
Agent Red: *Wait, wait, wait...*
Agent Blue: *You found it?*
Agent Red: *Wait.... I found it.*
[Agent Red throws scent word-ball to the screen.]

In this interaction, it was clear that Agent Blue was engaging her spatial memory to find the picture-ball she had seen while searching for a different word earlier in the game. Agent Blue was able to help Agent Red find the matching word-ball by describing what was next to the target word-ball. These interactions are enabled by our implementation that places the target balls in fixed compass directions. The arrangement of balls for both players match, and both players can see all red word-balls and blue picture-balls even though they can capture only balls matching their device color.

[Team 13, girls age 9]
Jessica (on the screen): *Throw Hector a word-ball now!*
Agent Red: [Reading the hint on the screen "What are we here to learn? instinct"] *We are here... we are... what are we here to learn?*
[...]
[Agent Blue throws scent picture-ball.]
Jessica (on the screen): *I don't think that was the right word. Try another one!*
Agent Red: *What?*
[Both players look for a new word.]
Agent Blue: *Oooh, I think I found it!*
Agent Red: *This is so cool, man!*
Agent Blue: [Throws track picture-ball]

Fig. 14. Artist's rendering depicting observed collaboration in play session. Here players are helping each other find appropriate words and pictures.

Jessica (on the screen): *Hector needs a word that means a person like you and me.*
Agent Red: *Oh I found it, I found it! Come!* [speaking to the word-ball that is moving away from her]
Agent Blue: *Ugh, it leaves.*
Agent Red: [Captures and throws the word-ball for human] *I got it!* [Starts dancing for a few seconds then continues searching for picture-balls]
Agent Red: *It's right there, it's right there, it's right here...it's right here, right here—come! It's here!*
[Agent Blue turns toward where Agent Red is pointing from across the room.]
Agent Red: *You don't see it right here?*
Agent Blue: *Um, I see hypnotize...*
Agent Red: *Come!*
[Gestures for Agent Blue to come over]
[Agent Blue walks over to Agent Red and points in the same direction]
Agent Red: *You see it right here? Look!*
Agent Blue: *Oooh, I see it!*
[Agent Blue captures picture-ball for human and throws it at the display.]

This rich collaboration sequence has some interesting elements that deserve to be dissected. First of all, this sequence demonstrates how the progressive hints helped scaffold the players to find the correct matching words. Agent Red reads the textual hint out loud, and later Jessica's child-friendly definition helps Agent Red find the correct word-ball.

Next, Agent Red proceeds to help Agent Blue find the correct picture-ball. Agent Red finds the correct picture-ball and points Agent Blue toward it. There is a slight breakdown at this stage because the mental model of the players seems to be that if they both point to the same area of the room, they'll see the same thing (see Figure 14). However, in this case the same area of the room meant different compass directions. The players were able to relatively quickly adapt by getting close enough together that they had a shared perspective that also happened to point in the same direction (see Figure 14). This might indicate that, if the technology can support it, a higher fidelity shared augmented reality experience that can model objects located in 3-D space (as opposed to compass orientation) might help collaboration. However, the players (especially older ones) were relatively adept at changing collaboration strategies once their initial interpretation proved faulty.

For the next word, the same players took a different collaboration approach:

Jessica (from the show): *Hector needs a word that means the smell of something.*
[Agent Red throws the word "scent" on the screen, waits to see if Manny zaps it.]
Agent Red: *Yay! Okay, now eh...* [Sits down briefly]
Agent Red: *Look for uh...* the nose. [Stands up and walks over to Agent Blue]
[Agent Blue finds the scent picture-ball and throws.]

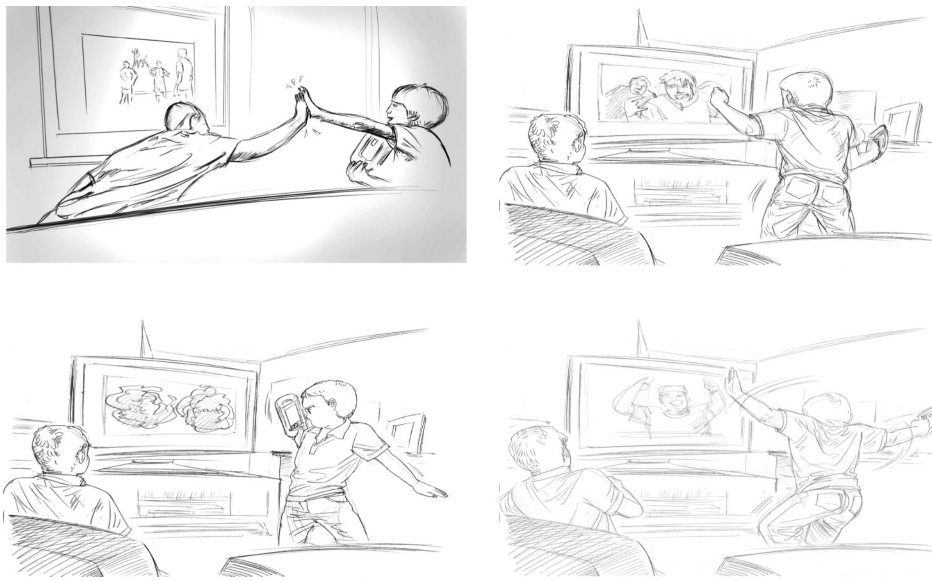

Fig. 15. Artistic renderings (to preserve child anonymity) from video of example celebrations: (Top Left) Team 7 congratulate each other with a high-five gesture after successfully advancing the episode. (Other pictures) Agent Blue from team 5 starts dancing to the closing music after the final word-ball.

5.2. Sense of Accomplishment

We observed several examples of players expressing a sense of accomplishment right after finding and throwing the correct word-ball, or when getting the word- and picture-balls to match up. Common celebratory outbursts included fist pumping, high-fives with the other player, and dancing (see Figure 15), often in conjunction with excited verbal exclamations. We interpret these expressions of accomplishment as a sign that the game was appropriately challenging for the players and that helping the Electric Company team and advancing the episode felt like an exciting success.

[Team 13, girls age 9]
[Agent Red throws a word-ball that matches the picture-ball previously thrown by Agent Blue. The episode advances.]
Agent Red: [Dancing] *Yay! I got it!*

5.3. Age Appropriateness of Task

Of the 17 pairs in our pilot, we had three pairs of children where both were 6 years old. All three of these pairs had tremendous difficulty with the task of finding and matching word- and picture-balls. They seemed to employ a strategy of randomly throwing word- or picture-balls. One pair tried a total of 15 throws between them on a single word challenge before researchers had to intervene to help them get through the study. This initial evidence indicates that the structure of the word searching task and/or the interface may not be appropriate for the youngest players in this age group. Interestingly, even the youngest children in this group demonstrated tremendous enthusiasm (e.g., jumping up and down after each throw) despite the repeatedly getting their word-balls rejected by Manny. This indicates that we may need easier tasks, or more aggressive scaffolding for players of this age in order to help them be successful. For the rest of the age group (ages 7 to 10), the tasks and content seemed appropriate, and progressive hints were sufficient for players to complete the tasks eventually.

5.4. Active, But Not Necessarily Physically Active

In the trial, we provided chairs for the players to sit down during the passive viewing parts of the experience, lasting two to three minutes each. However, we expected them to stand up and move around during the word-ball capture stages of the experience. Some of the players resisted moving from those chairs, contorting their bodies in awkward positions to reach word-balls that might be behind them. Despite their resistance to standing, these players were still clearly engaged in the task, eagerly chasing augmented reality balls and visibly celebrating after successfully throwing the correct word.

On the other hand, we saw many examples of players jumping around with enthusiasm and dancing in celebration. This physical activity was not at the level where it could be considered exercise but was encouraging as a sign of physical engagement.

5.5. Suspension of Disbelief

Feedback from players indicated that the combination of augmented reality with the design of the physical controllers was relatively successful in helping them buy into the experience. Some players even thought the word-balls were real.

[Team 11, boys 9 years old after playing Electric Agents]
Agent Blue: *For real, there are words around?*
Researcher: *Yes, they're all around you. They're everywhere.*
Agent Blue: *Like words like in real life?* [Looks around where the word-balls were in the experience]
Researcher: *Yeah! You just need to pull them into your head*.
Agent Blue: *Cool*!

5.6. Parasocial Relationships with the Characters

We saw some evidence of children identifying with and relating to the onscreen characters through the experience.

[Team 13, girls age 9]
Jessica (in the show): *Some instincts are human* [A member of the Electric Company in the show conjures a word-ball for human and throws it against the wall]
Agent Blue: *Oh, we're doing that!*
Jessica (in the show): *Some instincts are canine* [Conjures a word-ball for canine and throws it against the wall]
[Agent Red mimics the throwing action of the character in the show as she throws the word-ball].

Players who feel that they are part of the Electric Company are more likely to experience educational benefits of parasocial relationships found in previous literature.

5.7. Coping with Lag

There is some lag in updating the word-ball layer over the camera view as the orientation changes. Logically, the word-balls are placed at fixed orientations. However, in the experience the lag and noise in the sensor actually supported to the sensation that word-balls are floating through the air. These factors made word-balls a little more difficult to catch, but this challenge seemed to be relished by the children. If the lag or sensor noise were reduced, we would want to compensate by adding some randomness to the ball position to help maintain some difficulty in capturing the words.

6. DISCUSSION

The Electric Agents game exhibits some of the same design patterns identified earlier in our discussion of related interactive television projects. The way the show pauses to allow children to collect matching word- and picture-balls is very similar to the interaction patterns of "Winky Dink and You" and "Picture Pages." However, the use of interactive technology enables the inclusion of a tight feedback loop, which supports richer and more challenging engagement with educational content. In earlier low-tech television systems, the show would simply continue after a certain period of time even with no input from viewers, sometimes with nonsensical results (e.g., encoded messages from “Winky Dink and You” not being solved but narrative proceeds as if they were). Interactive feedback allows for adaptive scaffolding and contingent feedback, which includes progressive hints, tracking of progress and answer verification. Although not implemented in our prototype, our architecture also supports adapting content to the skill level of the user based on user performance.

Electric Agents provides a solution to balance viewer agency and production overhead through a layered meta-narrative. In our prototype, the meta-narrative is the story of Manny stealing words from Hector with his special gadget. Additional footage was also required to teach players how to use their devices, as well as to scaffold the game play (e.g., progressive hints). Creating the game story on top of the television show is relatively low budget, and many of the elements of the meta-narrative layer can be reused to create game versions of other episodes of the show. Initially we had concerns that the meta-narrative approach might be too abstract for the children, leading to confusion. However, there was no indication in our trial that children had trouble understanding or interpreting the layered story.

One unique aspect of the Electric Agents project is the use of broadcast-quality children’s educational media content in a research prototype. High-quality content is key to creating compelling television and interactive experiences. Typically, the biggest issue to overcome in a project of this type is covering the production costs associated with creating high-quality media. In the case of Electric Agents, the amount of custom content required was minimal due to our layered narrative design. Most of the video content had already been produced for “The Electric Company” television episodes. The remaining costs, including studio recording and post-production, were covered as a relatively small part of the prototype production budget.

One important direction for future research is addressing the educational effectiveness of the Electric Agents system, which was beyond the scope of the pilot user study reported here. In order to adequately assess educational benefits, participants would need to engage with more than one episode of Electric Agents, over a much longer period of time than one session. The pilot study reported here demonstrates patterns of behavior that have been linked to educational benefits, such as collaborative activity, scaffolded learning, and parasocial relationships.

Many readers may recognize similarities of the gesture interaction combining phones and televisions to today’s commercial game systems like Nintendo Wii and Microsoft Kinect. The key difference in Electric Agents is that the phone has much richer input and output capabilities (e.g., screen, speakers, cameras, and haptic feedback) than these systems as well as significant processing power. This allows the experience to migrate away from the television and into the real world (even beyond our living rooms). In our game, the phones use the display, camera, speakers, and rich sensing capabilities to create the illusion that word-balls were floating around the children in the real world.

7. CONCLUSION

Television consumption is migrating to the Web, and people are increasingly using their mobile phones concurrently while watching television. However, most current video applications are not taking advantage of the potential for interactivity on these mobile and Web platforms. Our transmedia system architecture demonstrates how mobile phones can communicate in real-time with Web-based video. This technology is used to explore a novel, educational, blended interactive television and mobile experience that pushes the boundaries of current conceptions in this design space, offering an innovative model for using mobile augmented reality in support of collaboration and learning. We believe that this is just the beginning of a transformation that will help promote active media participation and viewer engagement.

REFERENCES

Akl, A. and Valaee, S. Accelerometer-based gesture recognition via dynamic-time warping, affinity propagation, & compressive sensing. *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, (2010), 2270–2273.

Antle, A. N., Corness, G., and Droumeva, M. What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments. *Interacting with Computers 21*, 1–2 (2009), 66–75.

Ballagas, R., Kuntze, A., and Walz, S. P. Gaming Tourism: Lessons from Evaluating REXplorer, a Pervasive Game for Tourists. *Pervasive '08: Proceedings of the 6th International Conference on Pervasive Computing*, Springer (2008), 244–261.

Bandura, A. Social cognitive theory. In R. Vasta, ed., *Annals of Child Development*. JAI, 1989, 1–60.

Barkhuus, L. and Brown, B. Unpacking the television: User practices around a changing technology. *ACM Transactions on Computer-Human Interaction (TOCHI) 16*, 3 (2009), 1–22.

Bell, M., Chalmers, M., Barkhuus, L., et al. Interweaving mobile games with everyday life. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, ACM Press (2006), 417.

Bianchi-Berthouze, N., Kim, W. W., and Patel, D. Does Body Movement Engage You More in Digital Game Play? and Why? In A. C. R. Paiva, R. Prada and R. W. Picard, eds., *Affective Computing and Intelligent Interaction*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, 102–113.

Calvert, S. L. and Wilson, B. J. *The handbook of children, media, and development*. Wiley-Blackwell, 2011.

Calvert, S. L., Strong, B. L., Jacobs, E. L., and Conger, E. E. Interaction and participation for young Hispanic and Caucasian girls' and boys' learning of media content. *Media Psychology 9*, 2 (2007), 431–445.

Comstock, G. and Scharrer, E. *Media and the American child*. Academic Press, Burlington, MA, 2007.

Cress, U., Fischer, U., Moeller, K., Sauter, C., and Nuerk, H. C. The use of a digital dance mat for training kindergarten children in a magnitude comparison task. *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1*, (2010), 105–112.

Dow, S., Lee, J., Oezbek, C., Maclntyre, B., Bolter, J. D., and Gandy, M. Exploring spatial narratives and mixed reality experiences in Oakland Cemetery. *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology - ACE '05*, ACM Press (2005), 51–60.

Fisch, S. M., Lesh, R., Motoki, B., Crespo, S., and Melfi, V. Cross-platform learning: Children's learning from multiple media. *Proceedings of the 10th International Conference on Interaction Design and Children - IDC '11*, ACM Press (2011), 46–51.

Fisch, S. M. Characteristics of effective materials for informal education: A cross-media comparison of television, magazines, and interactive media. *The design of instruction and evaluation: Affordances of using media and technology*, (2004), 3–18.

Gawlinski, M. *Interactive television production*. Focal Press, 2003.

Gee, J. P. *What Computer games have to teach us about learning and literacy*. Palgrave Macmillan, New York, 2003.

Grimes, S. M. Saturday morning cartoons go MMOG. *Media International Australia 126*, (2008), 120–131.

Henrysson, A., Billinghurst, M., and Ollila, M. Face to face collaborative AR on mobile phones. *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, IEEE (2005), 80–89.

Hoffner, C. Children's wishful identification and parasocial interaction with favorite television characters. *Journal of Broadcasting & Electronic Media 40*, 3(1996), 389–402.

Howison, M., Trninic, D., Reinholz, D., and Abrahamson, D. The Mathematical Imagery Trainer: From Embodied Interaction to Conceptual Learning. *To appear in CHI 2011*, (2011).

Huynh, D.-N. T., Raveendran, K., Xu, Y., Spreen, K., and MacIntyre, B. Art of defense: A Collaborative Handheld Augmented Reality Board Game. *Proceedings of the 2009 ACM SIGGRAPH Symposium on Video Games - Sandbox '09*, ACM Press (2009), 135.

Isbister, K. and Dimauro, C. Waggling the Form Baton: An Analysis of Body- Movement-Based Design Patterns in Nintendo Wii Games, Toward Innovation of New Possibilities for Social and Emotional Experience. *Games and Culture*, (2011), 4–6.

Kimura, H., ed. *Basic Research on Academic Performance - International Survey of Six Cities Preliminary Report.* Kenichi Arai, Benesse Educational Research and Development Center, Benesse Corporation, 2007.

Klopfer, E. and Squire, K. Environmental Detectives—the development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development 56*, 2 (2007), 203–228.

Lakoff, G. and Núñez, R. *Where mathematics comes from: How the embodied mind brings mathematics into being*. Basic books, 2000.

Lauricella, A. R., Gola, A. A. H., and Calvert, S. L. Toddlers' Learning From Socially Meaningful Video Characters. *Media Psychology 14*, 2 (2011), 216–232.

McConnell, D. *Implementing computer supported cooperative learning*. Routledge, 2000.

Mulloni, A., Wagner, D., and Schmalstieg, D. Mobility and social interaction as core gameplay elements in multi-player augmented reality. *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts - DIMEA '08*, ACM Press (2008), 472.

PBSKids Raising Readers. "A Story of Success." http://pbskids.org/read/files/raising_readers_a_story_of_success.pdf. Accessed on Oct. 14, 2011.

Piaget, J. and Inhelder, B. *The psychology of the child*. Basic Books, 1969.

PlayScience. Mobile Playgrounds: Kids, Families and Mobile Play. *A PlayScience Lab Report 2*, 1 (2011), 1–4.

Prensky, M. Computer games and learning: Digital game-based learning. *Handbook of computer game studies*, (2005), 97–122.

Reeves, B. and Nass, C. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press New York, NY, USA, 1996.

Richter, R. A., Robb, M. B., and Smith, E. I. Media as Social Partners: The Social Nature of Young Children's Learning from Social Media. *Child Development* 82, 1 (2011), 82–95.

Rideout, V., Roberts, D. F., and Foehr, U. G. *Generation M2: Media in the Lives of 8-18 Year-olds*. Kaiser Family Foundation, 2010.

Rogoff, B. *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press New York, 1990.

Roschelle, J., Rafanan, K., Estrella, G., Nussbaum, M., and Claro, S. From handheld collaborative tool to effective classroom module: Embedding CSCL in a broader design framework. *Computers & Education 55*, 3 (2010), 1018–1026.

Slavin, R. E. Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology 21*, 1 (1996), 43–69.

Stevens, R., Satwicz, T., and McCarthy, L. In-game, in-room, in-world: Reconnecting video game play to the rest of kids lives. In K. Salen, ed., *The Ecology of Games: Connecting Youth, Games, and Learning. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning.* The MIT Press, 2007, 41–66.

Szentgyorgyi, C., Terry, M., and Lank, E. Renegade gaming: Practices Surrounding Social Use of the Nintendo DS Handheld Gaming System. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, ACM Press (2008), 1463.

U.S. Department of Labor. *American time use survey.* Bureau of Labor Statistics, 2009.

Vlahakis, V., Ioannidis, M., Karigiannis, J., et al. Archeoguide: an augmented reality guide for archaeological sites. *IEEE Computer Graphics and Applications 22*, 5 (2002), 52–60.

Wilson, M. Six views of embodied cognition. *Psychonomic Bulletin & Review 9*, 4 (2002), 625–636.

Xu, Y., Bolter, J., MacIntyre, B., et al. BragFish: Exploring Physical and Social Interaction in Co-located Handheld Augmented Reality Games. *Proceedings of the 2008 International Conference in Advances on Computer Entertainment Technology - ACE '08*, ACM Press (2008), 276.