Mixed Reality Virtual Pets to Reduce Childhood Obesity

7 authors, including:

Kyle John Johnsen
University of Georgia
81 PUBLICATIONS **907** CITATIONS

Sun Joo-Grace Ahn
University of Georgia
36 PUBLICATIONS **441** CITATIONS

James N. Moore
University of Georgia
243 PUBLICATIONS **5,480** CITATIONS

Aryabrata Basu
Emory University
17 PUBLICATIONS **96** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project Design, development, and testing of virtual patient stimulation View project

Project Visualizing Ancient Sanctuary of Samothrace View project

Mixed Reality Virtual Pets to Reduce Childhood Obesity

Kyle Johnsen, Sun Joo Ahn, James Moore, Scott Brown, Thomas P. Robertson, Amanda Marable, Aryabrata Basu

Fig. 1. A child interacts using speech and gestures with a mixed reality virtual pet kiosk. A wearable USB activity monitor is currently plugged into the kiosk that identifies the user and allows real world physical activity to serve as input into the simulation.

Abstract—Novel approaches are needed to reduce the high rates of childhood obesity in the developed world. While multifactorial in cause, a major factor is an increasingly sedentary lifestyle of children. Our research shows that a mixed reality system that is of interest to children can be a powerful motivator of healthy activity. We designed and constructed a mixed reality system that allowed children to exercise, play with, and train a virtual pet using their own physical activity as input. The health, happiness, and intelligence of each virtual pet grew as its associated child owner exercised more, reached goals, and interacted with their pet. We report results of a research study involving 61 children from a local summer camp that shows a large increase in recorded and observed activity, alongside observational evidence that the virtual pet was responsible for that change. These results, and the ease at which the system integrated into the camp environment, demonstrate the practical potential to impact the exercise behaviors of children with mixed reality.

Index Terms—Virtual reality, user studies, field studies, gestural input

---◆---

1 INTRODUCTION

Childhood obesity has reached epidemic levels in many regions around the world. For example, in the United States, 18% of children have been classified as obese [1]. This trend is foreboding, given that obesity in childhood is significantly correlated to obesity

• *Kyle Johnsen is with the University of Georgia. E-mail: kjohnsen@uga.edu.*
• *Sun Joo Ahn is with the University of Georgia. E-mail: sjahn@uga.edu.*
• *James Moore is with the University of Georgia. E-mail: jmoore@uga.edu.*
• *Scott Brown is with the University of Georgia. E-mail: sbrown01@uga.edu.*
• *Thomas P. Robertson is with the University of Georgia.*
  E-mail: tomrob@uga.edu.
• *Amanda Marable is with the University of Georgia.*
  E-mail: mmarable@uga.edu.
• *Aryabrata Basu is with the University of Georgia. E-mail:*
  basuarya@uga.edu

in adulthood [2] and that obesity is a leading risk factor for cardiovascular disease, diabetes, sleep and psychological disorders, and liver disease [3, 4]. Furthermore, the financial impact of obesity can be startling, as annual obesity-related healthcare costs in the United States alone have been estimated to be $147 billion dollars (10% of total healthcare spending) [5]. As a result, stopping the childhood obesity epidemic is a goal of extreme importance.

At the heart of the problem are increasingly sedentary lifestyles and poor dietary choices, which over time have led to an increase in childhood obesity. Along with the rise in obesity, children have been increasingly exposed to and utilizing interactive media. Though scientific evidence is incomplete as to the correlation between media use and childhood obesity, the majority of the literature and the commonly held belief suggest that if children would substitute exercise for low-activity interaction with media, they would be more likely to achieve a healthy weight [6].

Children must be motivated to make this behavior change, and the existing interactive media systems are a potential motivational mechanism. Our conjecture in this work was that engaging interactive media that included intense, real world, physical exercise as a vital mechanism to earn rewards could allow for such a

Table 1. Theoretical Framework Mapping

Dimension	Theoretical Element	Virtual Pet Feature	Expected Outcome
Individual	Self-efficacy (Confidence that one can exercise regularly)	The child sets his or her own exercise goal each day and can check progress.	By setting and meeting his or her own goals, each child will be able to experience success. This mastery experience, or the experience of success, is the greatest driver of self-efficacy. Heightened self-efficacy will motivate children to set incrementally more challenging goals for themselves.
Social	Vicarious experiences (Learning through observation, learning through competition with peers)	The child's physical exercise in the real world directly affects the appearance (e.g., looking fit or overweight) and the abilities (e.g., performing tricks) of their virtual pet.	Children learn that regular exercise is important by observing the health consequences of their virtual pets. Even if they are unable to meet their exercise goals, they will be able to see their peers and the performance of their peers' virtual pets to observe and learn from their successes/failures.
Environmental	Incentives (Rewards for meeting goals, consequences for failing to meet goals)	Children will earn tricks for reaching goals. Lower goals will unlock fewer, simpler tricks (e.g., sit, beg), whereas higher goals will unlock more challenging tricks (e.g., spin around, moonwalk).	The incentive to work hard to unlock the next level of difficult tricks for the virtual pet. Consequences will be evident when his/her virtual pet is only able to do simple tricks whereas another child's healthier virtual pet performs more sophisticated tricks.

substitution to take place. This alone is not a novel concept; research and commercial examples exist (See Section 2).

The innovation in our work was in the use of a tightly coupled mixed reality system designed to promote increased activity and in incorporating the interactive experience into an authentic, active space. Additionally, it focused more on intrinsic rewards for exercise (e.g. increased self-efficacy, personal satisfaction, positive feelings towards exercise), rather than extrinsic rewards (e.g. points, prizes).

In this paper, we describe the application as aligned with our theoretical framework (Social Cognitive Theory [7]) as well as the detailed mixed reality system design. We additionally report the encouraging results of an initial user study, which involved young children enrolled at a local camp using the application over the course of a week.

2 RELATED WORK

Our application and system build upon the strategies, technology, and research in serious video games for health. Serious video games for health are designed specifically to help increase attributes such as fitness or improve diet, and have received much attention in the research literature [8]. We subdivide these into two categories direct and indirect.

2.1 Direct Games

One variety is *direct*, where healthy activity within the game or healthy choices result in better game outcomes. Nintendo's Wii fitness games built for the Wii Balance Board fall into the direct category. The user's goal in these using these games is typically to exercise, and the purpose of the game makes exercise more enjoyable through a variety of classic gaming mechanisms such as points, levels, leader boards, story, etc [9]. In addition, many games that have been repurposed or rebranded as fitness games, because of their high degree of physical activity, such as Konami's Dance Dance Revolution series [10, 11]. These types of games can be identified by a lack of fitness related feedback or similar supports within the game, e.g. calories burned.

2.2 Indirect Games

The other variety is *indirect*, where real-world activity that is otherwise unrelated to the game (e.g. walking, consuming less calories) is measured and used as game input. These games are less common, but are increasing in both research and commercialization. For example, the game Fish N' Steps used pedometers worn by users grouped as teams to determine the weight and activity of virtual fish avatars, with each fish representing the activity of a user [12]. This approach successfully increased the user's level of physical activity in the real world, although only for a short period of time. A related approach, using mobile technology, allowed children to feed virtual pets using pictures taken of their own food which were graded for nutritional quality by the research team [13]. Results showed that players of the game ate a healthy breakfast 52% of the time, significantly higher than of the 20% rate of healthy breakfast consumption in the control group.

Commercial variations of this concept now exist, where users pay for both an activity monitor (usually a 3-axis accelerometer that logs activity and transmits that activity to the internet with a computer interface) and for prizes that can be earned by meeting activity goals (often gift cards or similar high-value goods purchased by a parent as extra incentive). HopeLab's Zamzee, features an online portal (www.zamzee.com) that tracks fitness goals and provides a large variety of free rewards for meeting goals in addition to paid rewards. Clinical trials conducted by HopeLab claim an average of 59% increase in physical activity for children who had access to the portal relative to children who only had the Zamzee accelerometer [14]. A more recent commercial system, GeoPalz's iBitz platform (www.geopalz.com) additionally links user activity to the health of an interactive virtual pet. Users exercise and then interact with a virtual pet (partially through an active interface) on a Bluetooth connected mobile device. As of this writing, no research results have been published pertaining to the iBitz platform.

3 APPLICATION

Our game used similar intent and component technologies as much of related work discussed, but combined these elements into a unique and cohesive mixed reality game experience.

3.1 Game Design

The focal point of the game is the virtual pet. Virtual pets are common in popular children's games domain, where children play games with the virtual pet and often nurture it through a button or touch interface, although more advanced interfaces feature gesture and speech interaction, such as the game Kinectimals (Microsoft Xbox 360) or EyePet (Sony Playstation 3). Using interaction with a pet to treat and prevent obesity has support in the health research literature. The prevalence of pet obesity tends to be positively correlated with owner obesity [15]. Moreover, pets can be used to promote additional exercise in direct clinical interventions, as with the Paired Pets program [16]. However, there exists little evidence to suggest that actual pet ownership reduces childhood obesity

directly through additional exercise [17]. Rather, the benefits are more related to companionship, which may be used to promote more positive behaviors.

The game was designed around the concept of a virtual animal clinic that had obese pets that needed exercise. A player would be "helping out" by exercising a pet for the clinic, and in return would be able to name and customize their pet, play with their pet, and help the pet learn new tricks by reaching exercise goals. As more activity was accrued, the pet would visibly lose weight, would become more

Fig. 2. The clip on activity monitors used in the study plugged into the USB port of the Kiosk. The USB interface on the activity monitor is retractable.

energetic, and would be happier. Pets would travel with the player "inside" an activity monitor that measured their exercise. At any time, the player could go to a designated kiosk to check their progress and interact with their pet in a virtual play area, where verbal and gestural commands could make their pet perform tricks. Additional tricks could be learned through a process of setting and reaching goals. The activity goal needed to be reached before new tricks could be learned, but the pet could be played with at any time after checking progress. The target players for the game were children between the ages of 8 and 11, corresponding to grade levels 4 through 6 in the United States.

Support exists in the virtual reality literature to suggest that players may not want their virtual pet to be obese, and would exercise more to prevent this event. Fox and Bailenson's research on virtual self-modelling showed that users would exercise more to keep a personalized avatar thin, but would not for a model of another [18]. We consider a pet avatar to be between these two extremes. Expanding on this concept, the virtual pet game was designed based on the framework of Bandura's Social Cognitive Theory to promote exercising behavior in children. This theory incorporates elements of individual cognition, social interactions, and influence from the external environment to create a comprehensive approach to changing behaviors. See Table 1 for a mapping between features of the virtual pet experience, its theoretical foundation within the Social Cognitive Theory, and the anticipated effects on the child owner of the virtual pet.

3.2   System Hardware

3.2.1   Activity Monitoring

Zamzee activity monitors (See Fig. 2) were used in the system. These battery-operated monitors periodically record activity every 10 seconds (with enough memory for approximately 1 month of data). The activity value recorded is an aggregate of several readings obtained from an integrated 3-axis accelerometer (Analog Devices ADXL345). In addition to excellent battery life (2 weeks), these monitors had a number of highly desirable characteristics, being designed with similar goals in mind to our application. First, they do not have any exterior buttons to configure the device or turn it off. Second, they are housed in a robust plastic packaging with a clip to

attach the device to the user's clothes or shoe. Finally, they have a retractable USB interface to access information through the standard Human Interface Device (HID) protocol. Through the HID protocol, the time-stamped activity could be transferred, along with the unique serial number for the device, and information could be erased to ensure that enough space was available on the device. Although Zamzee provides an application that transfers activity and uploads it to the Zamzee servers, this approach was not compatible with our design. It required users to have individual accounts on the Zamzee website, and would have made the transition of the virtual pet from the activity monitor to the kiosk difficult to synchronize. Thus, we developed a cross-platform device library using the open source HIDAPI library (www.signal11.us/oss/hidapi/). Our device library allows any computer application to use the Zamzee device without being tied to the Zamzee ecosystem. Note, this library is available as free open source software by request.

3.2.2   Kiosk

The kiosk was designed to be a standalone, portable system that could run without requiring staffing or user training. This is similar to requirements of an arcade game or museum exhibit. This required careful consideration and compromises to achieve.

The overall architecture is depicted in Fig. 1. An audio-visual media cart served as the housing for the kiosk, and we attached an LCD television monitor to the bracket. The television height was adjusted such that the center of the television was approximately the average height of the target audience (130cm). A computer was placed inside of the cart's lockable cabinet. A USB extension cable allowed for a single Zamzee device to be plugged into the computer. A second extension cable was used to mount a USB WiFi adapter outside of the metal cabinet, which attenuated the WiFi signal. In addition, a mouse and keyboard were placed in front of the television, and a Microsoft Kinect for Windows was mounted above the television. The Kinect facilitated both speech and gesture interaction with the virtual pet through its built in array microphone and depth camera. The keyboard was used by the player to name their pet, and the mouse was used for goal selection.

A restricted user account was created for the kiosk computer that could only run the virtual pet application, which was automatically started in full screen mode (1920 x 1080 resolution) on reset. This prevented unauthorized uses while the kiosks were deployed, and made them robust to power failure.

Although the television was capable of stereoscopic viewing when paired with polarized glasses, we specifically chose not to use this feature. The benefits of stereoscopic rendering are most evident when near-field depth perception is critical to the application [19], and to our knowledge, no psychological benefits (e.g. co-presence) have been found for similar applications involving interaction with virtual characters, e.g. for public speaking therapy [20]. Furthermore, requiring users to wear 3D glasses presented a logistical challenge to ensure that glasses were always available. As a separate consideration, we did not enable head-tracked fish-tank rendering, for similar lack of known benefit for the application. It is possible that in future incarnations of the virtual pet game, these features may be more useful, particularly if they become the norm for household video entertainment systems. For example, fish-tank rendering may allow for more accurate perception of virtual character gaze, particularly when paired with multi-view rendering in multi-user setups.

3.3   System Software

The virtual pet application was developed using the Unity3D Professional game engine. The user interface to the system is illustrated in Fig. 4. The screen area was divided into three regions, a large main area on the left that occupied 75% of the screen space to present the virtual pet and virtual dog park at life size proportions, and two stacked regions on the right side -- a help panel that continuously played relevant tutorial videos, and an avatar view

Table 2. Pet Tricks

| Trick	Goals Needed	User Action	Pet Action
Sit	1	User says "sit" and holds one hand out in front of their body.	Pet sits until user's hand is put down
Lay Down	2	User says "down" or "lay down" and holds one hand out in front of their body	Pet lays down until user's hand is put down
Speak	3	User says "speak".	Pet speaks (barks) a few times
Fetch	4	User says "fetch" or "ball". A ball appears in their virtual hand in the Kinect view. They can throw the ball by pulling back and then moving their hand forward. The ball follows the trajectory of the hand.	Pet retrieves the ball. If the ball is thrown off the map, the pet immediately returns with it.
Roll Over	5	User says "roll over" and rotates hand in a vertical circle.	Pet rolls over in a speed proportional to hand speed.
Play Dead	6	User says "die" or "play dead" and holds one hand out in front of their body	Pet plays rolls over and plays dead on back until the user's hand is put down
Crawl	7	User says "crawl" and moves both hands back and forth in crawling motion	Pet backs up and then crawls forward in a speed proportional to user's hand speed.
Beg	8	User says "beg" or "give paw" and moves a hand up and down in a handshaking motion	Pet waits for handshake and mimics motion.
Stand Up	9	User says "stand" or "stand up". Crouching ends the trick.	Pet stands on hind legs and waits for a crouching motion to get down.
Spin Around	10	User says "spin" or "spin around" and rotates one hand in a horizontal circle.	Pet spins around in a speed proportional to hand speed.
Moonwalk	11	User says "moonwalk" and then moves backwards and forwards in a moonwalking fashion.	Pet moonwalks in a speed proportional to user's walking speed.
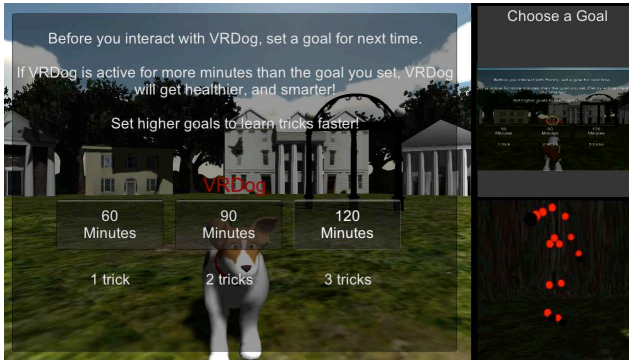
Fig. 4. Virtual Pet Kiosk Interface showing the goal setting screen. The main area on the left is for user interaction. The two panels on the right are for help videos and the Kinect skeleton view.

below the help panel that showed skeleton tracking information (a view from the pet's location).

Without an activity monitor plugged into the USB port below the television, the main area displayed the virtual dog park, and a translucent popup box in the center with the text "Insert your Activity Monitor to Begin". In addition, the video played live-action footage of a person plugging in the activity monitor. This was designed to address any confusion on how to get started with the interaction, given that under typical use, no staff would be present to help players.

When an activity monitor was initially plugged into the USB port, the system would first check the serial number for the device, and retrieve necessary information from an Internet-accessible server created for the virtual pet project. The server stored information in a MySQL relational database. The database contained information such as the mapping between device serial numbers and pets, as well as pet attributes such as name, colors, the current goal, current number of minutes of activity towards the current goal, and total activity accrued. The database was designed to make it easy to swap out a broken or lost activity monitor without losing pet information (only the activity since the last transfer would be lost). The server also stored all activity logs in associated files.

The correct virtual pet would then be instantiated into the virtual dog park facing the user and running an idle animation. New users would be given the opportunity to name their pet (which was

Fig. 3. An obese virtual pet. The pet is also slower at tricks.

prominently displayed above the pet) and to choose a primary and secondary color from a list of 8 options. Their choices were used to determine the color for the pet name, pet collar, pet tag, and user avatar.

The virtual pet was always the same, a virtual dog that had a cartoonish appearance and was of no obvious breed (See Fig. 3). Ideally, we would have had multiple pets to choose from, but this would have required more time and resources than were available for the project.

User avatars were presented as disconnected colored spheres located at each joint. This was purposeful simplification over the use of realistic avatars in an attempt to avoid user focus on this aspect of the system, while providing only the exact information provided by the Kinect for Windows skeleton tracking library, not any derived information, such as joint angles. Our thought was that attempting to render an a more realistic player avatar would have shifted the concept away from direct interaction with the pet to indirect interaction through the player's avatar. Ideally, no avatar would have been displayed, however the Kinect had limitations that the user needed to be aware of, such as ensuring that the correct player was being tracked and that all joints were visible to the Kinect.

Fig. 5. The progression of weight loss by the pet as more exercise is recorded and uploaded.

Although the Kinect for Windows can track up to 6 users simultaneously (full-body skeleton tracking for any two users), only the closest user was tracked and displayed. The purpose of this was to allow for a trivial explanation to users as to who was in control over the virtual pet -- simply the closest person to the Kinect. Alternative schemes that were rejected included raising ones hand to gain control (a popular approach in Kinect video games), and trying to detect who was speaking commands to the virtual pet. These schemes are more challenging to explain to users unfamiliar with the technology and often result in erroneous operation.

Upon instantiation of the pet, the recent activity logs were transferred from the connected activity monitor. The activity readings, which had no apparent dimension and ranged from approximately 10000 (no activity monitor idle) to 20000 (vigorously shaking the monitor), were then mapped from their dimensionless value to "minutes of exercise" by applying a threshold function, with a threshold of 16000. The threshold was established empirically by having volunteers wear the activity monitor and walk at a normal pace. The value 16000 emerged reasonable number for separating exercise from ordinary everyday movement for most volunteers. We determined that mapping scheme yielded similar results to the values obtained from Zamzee.

Once the minutes of activity were downloaded (this only took a few seconds), they were added to the pet's total activity. The total activity was displayed, along with the pet's goal. If the activity met or exceeded the goal amount, which was initialized as 0 minutes so that a goal was always reached the first time, the pet earned a trick award, which was four new tricks to start. Following this, the player selected a new activity goal (See Fig. 4). For example, the goals used in the study were 1 hour, 1.5 hours, and 2 hours of activity, with rewards of 1 trick, 2 tricks, or 3 tricks learned by the pet. After selecting a goal (using the mouse) or if a goal was not met, they could perform tricks with their pet. The trick mode then persisted until the activity monitor was removed from the USB port.

In addition to learning new tricks, pets also increased in physical fitness as more activity was accrued. This was manifested as a thinner, healthier looking pet (See Fig. 5 for a comparison) that could also perform tricks faster. Lastly, as more tricks were performed, the pet became increasingly "happy", manifested as an increasing magnitude and speed of tail wagging.

Tricks were performed with the pet using a multimodal speech and gesture interface. All tricks were instigated using a speech command, for example "sit". Then, a gesture typically modulated the trick in some way. For example, the pet would sit while the user had their hand outstretched, and users needed to make a rotating gesture with their hand to make the pet roll over. A full list of tricks and associated gestures, along with the number of goals needed to earn each trick is listed in Table 2.

The Microsoft Speech Application Program Interface (SAPI) was used for speech recognition, with a command-based recognizer. Command-based recognizers are much more robust than general-purpose natural language recognizers, and are even more appropriate for communicating with a virtual pet. The array microphone within the Kinect enabled hands free and headset free speech recognition.

In practice, a period of relative silence (i.e. low background noise only) needed to precede each command, but otherwise the speech interface was accurate (in a quiet room) in recognizing commands.

Users needed to know which commands to use and which gestures to perform. To assist with this, the interface's help panel had videos of a user performing each trick currently known by the pet playing cyclically while in trick mode (with captions for the speech command to instigate each trick). In addition, a few possible alternatives for each command were included (e.g. "stand" and "stand up").

4 STUDY DESIGN

The goal of the virtual pet experience was to increase the present and future activity of child users. To study the extent to which this goal was reached, we conducted a basic experiment with a treatment and a control condition.

The treatment condition was provided with the virtual pet experience as described in Section 3. The control condition used activity monitors and a goal-setting interface (similar to the treatment group), but featured no interaction with or reference to a virtual pet or virtual environment. The reason for the inclusion of a goal setting interface was to isolate the virtual pet as the treatment, as our theoretical framework would predict that setting and reaching goals alone would result in increased self efficacy and possibly increased physical activity.

The study primary outcome we discuss here was the extent to which the treatment influenced physical activity. A more in-depth analysis on the extent to which the experience may have produced meaningful and long-term behavior change according to our theoretical framework will be discussed in a future article.

4.1 Population and Environment

We partnered with a popular local summer camp (University of Georgia 4-H Center Summer Camp at Rock Eagle), whose focus is on healthy lifestyles, to perform the user study. Children attend the camp for a single week, arriving mid-day on Monday, and departing mid-day on Friday. On average, each week over the summer the camp hosts 600 children from around the region. Children in the camp are divided into a few "tribal groups", who compete with each other in various camp activities. We recruited participants from a single tribal group. Additionally children in tribal groups are divided into cabins, each cabin holding approximately 15 children. We had the capacity to support up to 70 children in the study, which was limited by the number of activity monitors on hand, and we recruited with the intended target of 60 participants for the possibility of the activity monitors being lost or broken.

Under our Institutional Review Board guidelines, to conduct studies involving children a parent must consent and the child must assent to the experiment. These forms were sent to parents a week prior to arrival at the camp. Upon arrival, the forms were checked for completion. Participants from cabins with a high percentage of completed forms were selected for inclusion in the study. To

Table 3. Study demographics

	Control	Treatment
Participants	28	33
Gender		
Female	17	16
Male	11	17
Grade Level		
3rd	0	1
4th	4	2
5th	8	25
6th	16	5
Region		
County 1	24	0
County 2	4	0
County 3	0	33

Table 4. Post experience Survey

Survey Question	Question Type
How much do you think the [virtual pet | pedometer] helped you exercise more that you did before you came to camp?	Likert, 1 (Not Helpful) … 3 (Somewhat Helpful) … 5 (Extremely Helpful)
What did you like the MOST about your [virtual pet | computer/pedometer]?	Free Response
What did you like the LEAST about your [virtual pet | computer/pedometer]?	Free Response
If you could change some things about your [virtual pet | computer/pedometer what would they be?	Free Response

minimize discussion about the study, we opted to assign participants randomly to each condition by cabin, rather than individually.

Table 3 shows the breakdown of the two study groups by population demographics (obtained from the camp records, not from individuals directly). As an unintended consequence of random cabin assignment, two of the individual demographic variables that were recorded co-varied with experimental condition: grade level and region. Grade level was of lesser concern, as ages tend to overlap between 5th and 6th grade, and most grades were represented in each group. Of greater concern, the regions entirely co-varied with condition. Thus, it was impossible to control for region in our analyses. However, neither gender nor grade level were significant factors in results (See Section 5), and would be expected to account at least partially for any variance derived from region.

Two similar kiosks (one had a 55 inch television, the other a 60 inch television, but were otherwise identical) were deployed at the camp in a central assembly building, but located in different hallways to prevent speech recognition interference. Two kiosks were used to allow for increased user throughput during the course of the experiment. These kiosks were to be used exclusively by participants in the treatment condition. Additionally, two desktop computers were deployed in the same building (in a third hallway) for use by participants in the control condition. Participants in the control condition were restricted by the software from using the virtual pet kiosks (nothing would happen when the activity monitor was inserted). Similarly, participants in the treatment group were restricted from using the control condition computers. However, we could not absolutely restrict participants in either group from observing participants in the other, as the kiosks and computers were purposefully located in public areas to promote frequent use. We considered the possibility of dividing the study into two camp weeks, but this would have the additional challenges of differences in non-study-related camp activities and weather patterns between the weeks.

4.2   Measures

The measures for the study primarily consisted of data automatically recording during the course of the experience. This consisted of the activity monitor logs (recorded every time the activity monitor was plugged into one of the study kiosks or computers) as well as interaction logs that included the frequency of visiting the kiosks or computers, a record of goal setting actions, and for the treatment condition, a log of the tricks performed by the pet.

In addition, a short post-experience survey was administered to gather more qualitative data related to the experience. Questions relevant to this analysis are shown in Table 4.

4.3   Procedure

The control and treatment groups were separately gathered near the end of the first day of camp, along with their camp counsellors, to introduce the participants to the study and to distribute the activity

monitors. Participants were given an overview of their responsibilities in study (such as wearing the monitor correctly in their name badge pouch, and not discussing the study with those outside their cabin). They also had an opportunity to ask questions. After this, they were provided a demonstration of the interface they would be using, and used the interface to initialize data collection and set their first activity goals. The demonstration for the treatment group discussed the capabilities of the technology being employed, along with limitations (e.g. speech recognition inaccuracy in noisy environments and the limited field of view of the Kinect camera), and strategies to recognize and address those limitations (e.g. speaking clearly, using the kiosks during quiet times, and ensuring that the full and correct skeleton is being tracked). The various ways that the pet's health, tricks, and happiness could be affected were also discussed. Participants in the treatment condition also entered a name and chose colors for their pets at this time.

Following this, over the next 72-hour period, participants could, at their leisure, use the workstations or kiosks. Counsellors were requested to remind participants to upload their activity and monitor their goals frequently, but participants were not required to do so. After the 72-hour study period, the post-experience survey was administered in a computer lab. At this time, all activity monitors were collected and a final data transfer to upload remaining activity was initiated.

5   RESULTS

An analysis of variance was conducted on the activity data set, with condition and gender used as independent factors. A highly significant effect was found for condition ($F=16.613$, $p<0.001$). Total
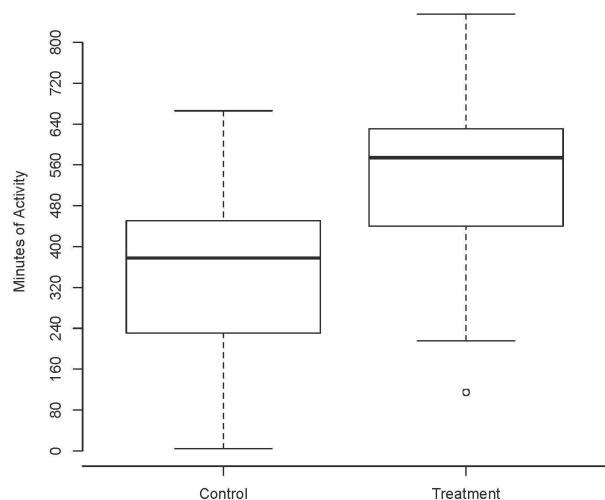

Fig. 6. Total activity over the study period was significantly higher in the treatment group ($p<.001$).

minutes of activity over the 72-hour period were substantially higher in the treatment condition (M=537, SD=165) than in the control condition (M=352, SD=173). A trend for gender was found (F=2.735, p=0.103), with female participants (M=419, SD=177) averaging less minutes of activity than male participants (M=491, SD=204). The box plot of activity data with respect to condition (See Fig. 6) illustrates the large difference between the treatment and control conditions. Additional evidence for this effect can be found in the extremes of the data. Among all study participants, 8 out of 10, including the 7 highest amounts of total activity among all participants in the study were in the treatment group. Similarly, 8 out of 10 of the lowest amounts of total activity were in the control group, including the 3 lowest. The most active participant recorded over 14.26 hours of activity in the 72-hour period.

We suspect that some participants did not wear their activity monitor for the duration of the study, resulting in low activity during what is known to be an active camp with many activities that would, by participation alone have resulted in recorded activity. However, we had no direct way to know if this result was because of inactivity or failing to wear he activity monitor. Excluding the outliers in the dataset did not change the significance of the results.

Fig. 7 shows a time log of activity as it was uploaded to the database at the start of each interaction with the system. Activity appeared to be regular and sustained throughout the study time period, with most activity clustered during waking hours (the study began at approximately 5:00pm). Participants in the treatment condition interacted more often with the virtual pet. The recorded number of interactions with the system was significantly higher (F=10.322, p < 0.01) in the treatment condition (M=13.6, SD=7.1) than the control condition (M=8.3, SD=6.8).

Results from survey questions did not shed substantial light on the reasons for the differences between groups. The average response given to "How much do you think the [virtual pet | pedometer] helped you exercise more that you did before you came to camp?" was 4.18 out of 5 in both conditions, however we note that a higher percentage of participants in the treatment condition rated this item a 5 (54%) than in the treatment condition (42%).

Comments about the systems from both the treatment and control groups were generally positive. Participants in the treatment condition often cited the ability to "teach the dog new tricks" as what they liked most, although some commented that they liked that the dog became skinnier with increased exercise. They also liked the ability to give commands by voice, but this was also what those in treatment condition liked least about the virtual pet, in that speech recognition did not always work (likely because of a noisy environment at the time).

6 Discussion

During the initial design phases of the user study, there were many concerns that the interface would not work with children, that the virtual pet system would not stay operational for the entire study period, that the activity monitors would be lost or broken, that the camp would be too active to find a difference between groups, and that the pet was not a significant motivator for children relative to simple goal setting and providing feedback on their activity. All of these concerns proved to be false.

The interface did work quite well with the children. Many of them performed dozens of tricks with the pet. Participants played "fetch" over 200 times in the study, which was the most interactive of the tricks. This suggests incorporating more interactivity into the system would be additional incentive to be active.

The virtual pet system was highly reliable, though our strategy of deploying two kiosks proved necessary. One of the kiosks was accidentally unplugged during the study, but was returned to operational status within a few hours. The other kiosk remained operational for the duration of the study. One of the reasons for this is that we leveraged well-tested off-the-shelf technology for the majority of the system.
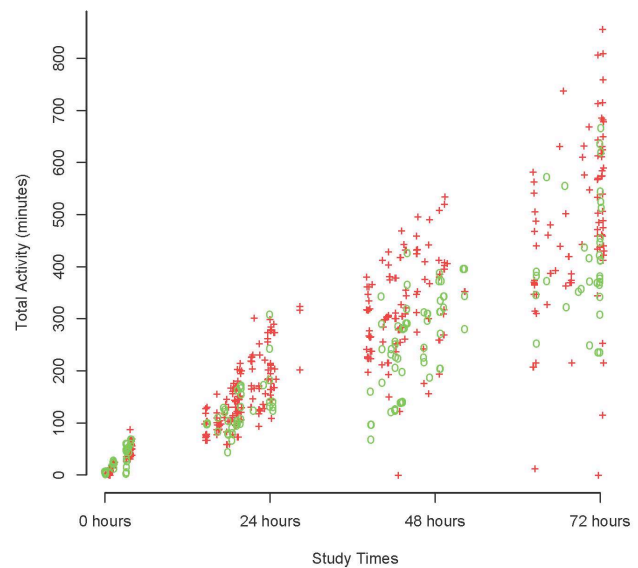

Fig. 7. Accumulated activity upload events over the study period. Green circles are from the control group and red crosses are from the treatment group.

The activity monitors also proved to be reliable and durable. Three of the 61 activity monitors were replaced during the study, a failure rate of 5%. This was below the 10% that we anticipated, given the many opportunities at the camp to damage them.

Our results show a 60% increase in physical activity in the treatment condition relative to the control condition. While we expected an increase, we did not expect that the difference would be so large. First, the camp environment is a place of already high activity. Children are outside for a large portion of the day, and travel between areas of the camp frequently. Thus, we questioned whether the camp environment would cause a ceiling effect on our results. What we found was that the camp environment provided additional possibilities for exercise that were not expected. For example, on the first day of the study, one of the counsellors provided an anecdote to us that she was going to take her cabin to another area of the camp by van, but that the children stopped her that were part of the study and asked if they could walk instead: so that their pets would get the exercise! These anecdotes continued throughout the study.

Taking into account the results from related studies further highlights the magnitude of the increase. Researchers for the Zamzee product claimed a 60% increase in exercise when children used their product relative to a control group when deployed in school environments. Our results show a similar increase (albeit for a shorter amount of time), but without the extrinsic monetary rewards, and with a more rigorous control condition (activity monitor + goal-setting interface). Moreover, the virtual pet system was limited in functionality. In many ways, it was an interactive trophy. The pet could only perform 11 tricks, but it got thinner and faster and wagged its tail more vigorously. It was in the hallway where all of the children pass by on a daily basis. Those who brought the pet to peak physical fitness could be proud of their accomplishment in an environment that promotes such achievement. We believe that these rewards are more likely to promote future, self-motivated physical activity and ultimately healthier weights.

7 Conclusions

Overall, we felt that the study was a resounding success. With minimal game content, relative to far more elaborate entertainment games, the virtual pet succeeded in motivating the treatment group to exercise significantly more than their peers in the control group, who only had activity monitors and the motivation from goal-setting. The

effects did not appear to diminish over time, a concern with most innovative technology solutions. The study did have some internal flaws from not taking into account the organizational structure of the camp, and in participant assignment into conditions. The large magnitude of the difference in activity, though, suggests that the findings are at least partially the result of motivation derived from interaction with the virtual pet experience. Moreover, qualitative feedback from participants and observers suggest that improving the pet's knowledge, behavior, and physical fitness were the likely driving mechanisms behind the behavioral findings.

We were also pleased with the overall logistics of the study. By making the system portable and self-contained, the kiosks could be transported, set-up, and removed easily. Furthermore, the study operated automatically for the majority of the 72-hour period. Researchers were only on-site during the initial orientation and final surveys. This shows the scalability and robustness of the design, traits that are uncommon in virtual reality systems. These are the first steps towards practical, wide scale deployment.

8 FUTURE WORK

By capitalizing on the popularity of virtual pet games for children and adding theory-driven mixed-reality design components, a new weapon in the global battle against obesity has been created. Logical next steps for this work include improving upon the system interface and adding new interaction mechanisms or even entire interfaces (e.g. mobile access). In addition, the study compared a complex system to a simple one, and thus the relative importance of individual components of our system is impossible to determine with this dataset. This will take many more studies, which fortunately is afforded by the camp test-bed.

Our immediate task is to redesign the system interface to be more flexible and reliable towards running a larger, more in-depth study in the camp environment. The speech interface proved to be the largest source of frustration, and will be the initial focus of the redesign. The original design was created without substantial experience working with the target audience, and failed to take into account that children often mumble, do not enunciate, or are very quiet or loud. They tended to speak to the virtual pet as though it were a real pet, and not a computer interface. As a result, speech will move towards being an optional control interface, rather than a required one. Instead, direct gesture commands will be possible. For example, the player could say "sit", but could also simply hold out their hand and make a sitting gesture. Either could work, and the presence of both could help resolve ambiguous cases (much like a real pet might need).

From a content perspective, we will focus on improving the relationship building attributes of the pet. Additional ability to choose and customize the pet could be a powerful motivator to keep the pet healthy. Also, if the pet shows unique behaviors relative to peers' pets, this should further the bond between player and pet. The pet should recognize its owner by face recognition, and not only by the Zamzee serial number. Related to owner recognition, the pet should act "aware" of the user, knowing when it is being looked at, pointed to, and praised. All of this is possible with existing hardware and known algorithms. Ultimately, we hope to add to the growing body of literature on human behavior with embodied conversational agents, which tends to focus on symmetrical human-human interactions, as opposed to the asymmetrical human-animal interactions in the current work.

Finally, while the camp environment provided an excellent test-bed for the concept, it is not an ideal location for eventual deployment. Rather, the system is best suited to be installed in a place that is popular in the community, and part of daily life. Community centers and schools could be ideal venues for the kiosks. Developing a home use or mobile version of the system could also prove powerful, but must maintain the social and environmental motivating factors. We are looking to augment the existing system with other portals to access and interact with the virtual pet, but the mixed-reality kiosk concept will stay central to the overall experience.

REFERENCES

[1] C. Ogden and M. Carroll, "Prevalence of obesity among children and adolescents: United States, trends 1963-1965 through 2007-2008," *NCHS Health E-Stats*, 2010.
[2] R. C. Whitaker, J. A. Wright, M. S. Pepe, K. D. Seidel, and W. H. Dietz, "Predicting obesity in young adulthood from childhood and parental obesity," *New England Journal of Medicine*, vol. 337, pp. 869-873, 1997.
[3] S. R. Daniels, D. K. Arnett, R. H. Eckel, S. S. Gidding, L. L. Hayman, S. Kumanyika, *et al.*, "Overweight in children and adolescents: pathophysiology, consequences, prevention, and treatment," *Circulation*, vol. 111, p. 1999, 2005.
[4] J. L. Baker, L. W. Olsen, and T. I. A. Sørensen, "Childhood Body-Mass Index and the Risk of Coronary Heart Disease in Adulthood," *New England Journal of Medicine*, vol. 357, pp. 2329-2337, 2007.
[5] E. A. Finkelstein, J. G. Trogdon, J. W. Cohen, and W. Dietz, "Annual medical spending attributable to obesity: payer-and service-specific estimates," *Health Affairs*, vol. 28, p. w822, 2009.
[6] A. Must and D. J. Tybor, "Physical activity and sedentary behavior: a review of longitudinal studies of weight and adiposity in youth," *International Journal of Obesity*, vol. 29, pp. S84-S96, 2005.
[7] A. Bandura, Social foundations of thought and action: A social cognitive theory: Prentice-Hall, Inc, 1986.
[8] A. S. Lu, H. Kharrazi, F. Gharghabi, and D. Thompson, "A systematic review of health videogames on childhood obesity prevention and intervention," *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, vol. 2, pp. 131-141, 2013.
[9] S. Göbel, S. Hardy, V. Wendel, F. Mehm, and R. Steinmetz, "Serious games for health: personalized exergames," in *Proceedings of the international conference on Multimedia*, 2010, pp. 1663-1666.
[10] V. B. Unnithan, W. Houser, and B. Fernhall, "Evaluation of the energy cost of playing a dance simulation video game in overweight and non-overweight children and adolescents," *International journal of sports medicine*, vol. 27, pp. 804-809, 2006.
[11] A. J. Daley, "Can exergaming contribute to improving physical activity levels and health outcomes in children?," *Pediatrics*, vol. 124, pp. 763-771, 2009.
[12] J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H. Strub, "Fish'n'Steps: Encouraging physical activity with an interactive computer game," *UbiComp 2006: Ubiquitous Computing*, pp. 261-278, 2006.
[13] J. Pollak, G. Gay, S. Byrne, E. Wagner, D. Retelny, and L. Humphreys, "It's Time to Eat! Using Mobile Games to Promote Healthy Eating," *Pervasive Computing, IEEE*, vol. 9, pp. 21-27, 2010.
[14] S. Cole. (2012, 9/11/2013). *Zamzee Research Rsults*. Available: http://www.hopelab.org/innovative-solutions/zamzee/zamzee-research-results/
[15] K. L. Holmes, P. J. Morris, Z. Abdulla, R. Hackett, and J. M. Rawlings, "Risk factors associated with excess body weight in dogs in the UK," *Journal of Animal Physiology and Animal Nutrition*, vol. 91, pp. 166-167, 2007.
[16] R. F. Kushner, D. J. Blatner, D. E. Jewell, and K. Rudloff, "The PPET Study: People and Pets Exercising Together," *Obesity*, vol. 14, pp. 1762-1770, 2006.
[17] C. Westgarth, J. Heron, A. R. Ness, P. Bundred, R. M. Gaskell, K. Coyne, *et al.*, "Is Childhood Obesity Influenced by Dog Ownership? No Cross-Sectional or Longitudinal Evidence," *Obesity facts*, vol. 5, pp. 833-844, 2012.
[18] J. Fox and J. N. Bailenson, "Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors," *Media Psychology*, vol. 12, pp. 1-25, 2009.
[19] Y.-Y. Yeh and L. D. Silverstein, "Spatial judgments with monoscopic and stereoscopic presentation of perspective displays," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 34, pp. 583-600, 1992.
[20] Y. Ling, W.-P. Brinkman, H. T. Nefs, C. Qu, and I. Heynderickx, "Effects of stereoscopic viewing on presence, anxiety, and cybersickness in a virtual reality environment for public speaking," *Presence: Teleoperators and Virtual Environments*, vol. 21, pp. 254-267, 2012.