"I'm the Jedi!" - A Case Study of User Experience in 3D Tele-immersive Gaming

Wanmin Wu*, Ahsan Arefin*, Zixia Huang
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois, United States
Email: {wwu23, marefin2, zhuang21}@illinois.edu

Pooja Agarwal, Shu Shi, Raoul Rivas, Klara Nahrstedt
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois, United States
Email: {pagarwl, shushi2, trivas, klara}@illinois.edu

Abstract—In this paper, we present the results from a quantitative and qualitative study of distributed gaming in 3D tele-immersive (3DTI) environments. We explore the Quality of Experience (QoE) of users in the new cyber-physical gaming environment. Guided by a theoretical QoE model, we conducted a case study and evaluated the impact of various Quality of Service (QoS) metrics (e.g., end-to-end delay, visual quality, etc.) on 3DTI gaming experience. We also identified a number of non-technical factors that are not captured by the original theoretical model, such as age, social interaction, and physical setup. Our analysis highlights new implications for the next-generation gaming system design, as well as a more comprehensive conceptual framework that captures non-technical influences for user experience in such environments.

Keywords-3D tele-immersion; Gaming; Quality of Experience; Quality of Service;

I. INTRODUCTION

Electronic gaming systems have come a long way from the arcade pioneers in the 1970's to the latest video game consoles (e.g., Microsoft Xbox [7] and Sony Playstation [5]) and computers (e.g., for Massively multi-player online role-playing games or MMORPGs). More recently, with the advent of industrial products (e.g., Nintendo Wii [6], Project Natal [8]) and academic prototypes (e.g., Exertion Interface [15], Kick Ass Kung-Fu [4]), "exergaming" has resurged. Exergaming stands for exercising and gaming, which involves real physical movements of players that are mapped directly onto the characters in the virtual gaming worlds. However, these existing systems are still limited in that they mostly rely on simple graphical avatars as visual feedback for gamers.

The recent emerging 3DTI systems, offering a greater sense of telepresence by immersing players with photorealistic 3D videos, are speculated to become the next-generation gaming technology. Imagine a Star Wars lightsaber game where two geographically dispersed players could see each other in their 3D photorealistic appearances (e.g., using holography), and fight as if they were in front of each other using haptic-equipped lightsabers. In fact, we are getting closer and closer to make this vision a reality as

researchers have reported successful distance art performances and medical consultation with the state-of-the-art 3DTI systems [13][27].

Gaming with the new 3DTI technology can be a revolutionarily new experience for people, which means future 3DTI system designers should re-examine the design principles in order to create more engaging user experience. This becomes our motivation for this paper. The main goal of this research is to explore the user experience in the new cyber-physical gaming environment. We are interested in these research questions: How do people react to the new 3DTI gaming experience? How do the technical QoS metrics (e.g., delay, visual quality) impact QoE in 3DTI gaming? How do the non-technical factors (e.g., age, social interaction) influence QoE?

We developed a 3D tele-immersive gaming environment that is equipped with real-time audio and video communication modules based on the prescribed perceptive thresholds (e.g., on delay, jitter) suggested by the literature. We designed a virtual lightsaber duel game that features interactive visual feedback. In order to systematically understand user experience in this environment, we adopt the theoretical QoS-QoE framework developed by Wu *et al.* [25], where QoE in such Distributed Interactive Multimedia Environments (DIMEs) is modeled as a multi-dimensional construct. We conducted a qualitative and quantitative study of 3DTI gaming that involved over a hundred participants. We present our quantitative results of QoS-QoE evaluations and qualitative analysis of user behaviors during the games. Our analysis highlights new implications for the next-generation gaming system design, as well as a more comprehensive conceptual framework that captures the non-technical influences for user experience in such environments.

The rest of the paper is structured as follows. Sec. II-A gives an overview of the 3DTI gaming environment, and briefly reviews the theoretical model for QoS-QoE evaluations [25]. Sec. III describes the related research. Sec. IV describes our case study, including setup and equipments, game design, participants, and methods. Sec V presents our results and analyses. Sec. VI discusses the theoretical and practical implications for future 3DTI gaming environments. Sec. VII concludes the paper.

---

*These authors contributed equally to this work.

II. OVERVIEW

A. 3DTI Cyber-Physical Gaming Environment

3DTI cyber-physical gaming environment is a multi-party 3D collaborative system, where the 3D videos of users from geographically distributed sites are merged into a shared virtual cyberspace. In the shared domain, each gamer can manipulate his/her virtual size, position, and interact with remote gamers or different graphical objects. 3D photo-realism, distributed communications, and remote interactivity with full-body physical activities are the key characteristics of such environment.

Each site consists of several local computing nodes supporting input and output devices, and an overlay routing node called *service gateway*. Communications to/from remote sites are done through gateways. Input nodes connect sensing devices (e.g, cameras, microphones) to the 3DTI system to capture user's actions in real-time, whereas output nodes connect displays and haptic interfaces to present the joint virtual space to the users.

Figure 1 shows the logical architecture and the corresponding physical setup of a 3DTI gaming environment. For simplicity we show only 3 sites with video and audio input devices. Site-1 and Site-2 send and receive audio and video streams while Site-3 only displays received media. Each 3D camera (C) is connected to an input node that grabs multiple 2D video frames simultaneously, performs 3D reconstruction on the 2D frames, and streams in real time the 3D video frames to the local gateway (G). Microphones (A) are used for capturing audio. The audio is then sent to the local gateway as well. The gateway then disseminates the streams to its local output nodes as well as to remote gateways (G) for remote user(s). Each receiving gateway G feeds the incoming streams to its own output nodes that support speaker and/or rendered visual data on a display (R).

To ensure real-time data delivery, it is important to use an efficient transport layer protocol which provides fast delivery as well as minimal jitter between consecutive packets. In the past, TCP and UDP have been the most widely used protocol for 3DTI systems [17][26]. However, TCP fails to provide low delays or jitters due to its retransmission and conservative congestion control mechanism. UDP provides minimal delays and jitters but applications need to build their own congestion control mechanisms to provide fairness towards concurrent traffic.

We use DCCP, which is a new protocol that provides unreliable but congestion controlled delivery of streams. To adapt to 3DTI semantics, we added (de-)fragmentation, soft real-time scheduling, and packet re-ordering mechanisms to DCCP. We performed experiments between two universities which are about 2000 miles away with 16 streams each of 3 Mbps. The results showed that DCCP could achieve an average one-way delay that was a factor of 3 smaller than TCP, and an average jitter that was half of that by TCP. With a soft timeout set at 120 ms, the loss rate of DCCP rapidly decreased to zero within 45 seconds.

The audio quality highly affects user experience too. Jitters and losses cause audio packets to be unavailable at the receiver at their scheduled play-out times, which has a direct impact on the understanding of speech contents. This is denoted as the listening only speech quality (LOSQ) as it is solely related to the intelligibility of speech heard [19]. Internet latencies also affect audio quality, because a longer audio mouth-to-ear delay (MED) will affect the conversational interactivity. ITU G.114 [11] prescribes that one-way end-to-end delay should be below 400 ms. Finally, video-audio synchronization is also critical. [21] has recommended an in-sync region of a maximum 80 ms.

We have designed an audio component that satisfies the prescribed thresholds. The audio packets are encoded every 20 ms and sent to the other sites via UDP. UDP is more suitable for small-sized audio packets to achieve constant frame rate. For Internet dynamics (jitters and losses), an adaptive play-out scheduling scheme based on previous delay history and a piggybacking algorithm for loss concealment have been developed [19]. For synchronization, we have designed a multi-tier sync framework [9] to deal with the multiple correlated streams, and have achieved proper audio-visual synchronization with an in-sync skew less than 80 ms.

B. QoS-QoE Model

In order to systematically analyze the influences of various QoS metrics on user experience in 3DTI gaming, we adopt Wu *et al.*'s QoS-QoE conceptual framework [25] which is specifically designed for distributed interactive multimedia environments like 3D tele-immersive gaming. According to Wu's model, QoS and QoE are defined as two distinct constructs and related through a causal chain of *environmental influences* (QoS) that impact *cognitive perceptions* and *behavioral consequences* (QoE) of a user.

QoS represents the quality levels and technical environmental influences of the cyber-physical spaces where 3DTI systems are embedded. It has three dimensions: vividness, interactivity and consistency. *Vividness* manifests itself through *breadth* and *depth*, where the former defines the presence of media channel (e.g., visual, auditory, haptic), and the latter defines the resolutions of each channel. *Interactivity* defines (1) *speed* (delay, real-time) which is directly related to the response time between users' action and virtual space reaction (representation), media end-to-end delay, frame rate for continuous playback, jitter, etc., (2) *range* which specifies the degree of freedom to which users can change the mediated environment, and (3) *mapping* which measures the manipulative control of a user interface, i.e., how natural and intuitive the control is. *Consistency* refers to the mutual manifestness of the joint virtual environments presented at all sites.
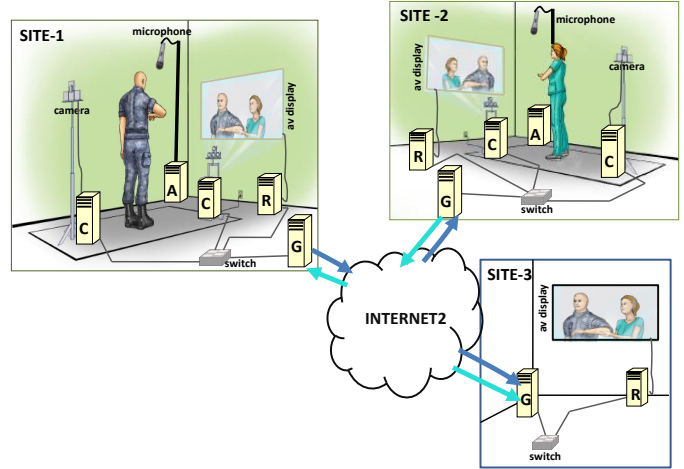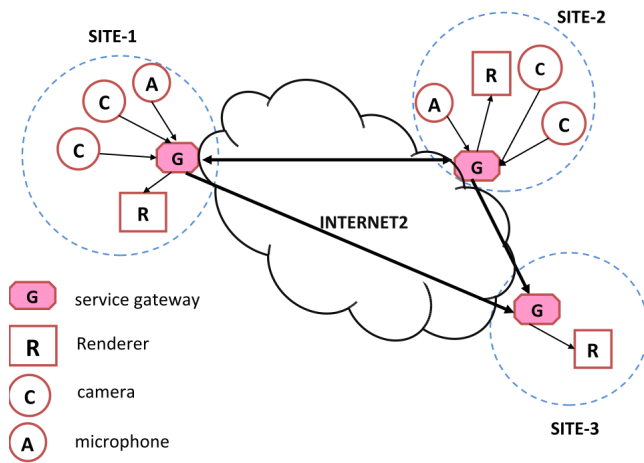
Figure 1. (a) Logical Architecture, and (b) Physical Setup of 3DTI system

QoE represents the experience of users embedded in 3DTI environments, represented by (1) cognitive perceptions, measured by *flow* (psychological experience of a user consisting of metrics including *concentration, enjoyment*), user's perceived *technology acceptance* and *telepresence* experience, and (2) user's behavioral consequences, measured by the *performance gains* of a user in an activity, *technology adoption* degree, and spontaneous *exploratory behavior*.

III. RELATED WORK

To the best of our knowledge, this research is the first attempt to examine gaming experience with the next generation 3DTI technologies. In the literature, there have been some studies on user experience with the current gaming technologies [1], [2], [3], [14], [10], [15], [16], [4], [23], [24]. However, the prior works are limited in three ways: (1) most studies used ad hoc methodologies and lacked a theoretical foundation to characterize the quality of user experience, (2) many studies focused on specific aspects (e.g., gender [18], age [10], network quality [1]), failing to capture a more complete view of the gaming experience, and (3) although experience is strongly context-dependent [12], most studies did not provide any concrete evaluation of the technical environment (e.g., QoS metrics) that can have great influences on user experience.

IV. CASE STUDY OVERVIEW

We are interested in these research questions: What are people's reactions to the new 3DTI gaming experience? How does each QoS metric impact QoE in 3DTI gaming? How does each non-technical factor (e.g., age, social interaction, presence of others) influence QoE?

To best understand the impacts of various factors (e.g., QoS performances, demographics, social aspects) on user experience in 3DTI gamings, we conducted a large-scale case study, which took place during a public event organized in a university in the United States.

A. Setup and Equipments

We set up two 3DTI sites in two separate places in a departmental building. Each was used by one gamer at a time. Gamers could not see or hear each other physically, and had to rely on telecommunication. Each TI site consisted of a 3D camera, a display, a black curtain and several host PCs (refer to Figure 1). A Point Grey Bumblebee 3D camera was set up on a tripod with adjustable height. A black curtain was put up against the camera and behind the participant to facilitate background subtraction. A Philips 42" 3D-WOW-Display was used in each testbed to present the virtual world to the users. Wireless bluetooth headsets were provided for VoIP communication.

B. Game Design

A lightsaber duel game was designed to take full advantage of remote interaction. Figure 2 illustrates how the game is played. Participant 1 in Site-1 puts on a lab coat with red patches and takes a green lightsaber, while Participant 2 in Site-2 puts on a coat with green patches and holds a red lightsaber. Each participant then tries to use the lightsaber to hit the opponent's color patches in the virtual space as much and as fast as possible in order to gain points. A collision detection module is developed in the renderer to automatically detect the hitting. Once a successful hit occurs, the sword and patches turn blueish for an electrifying effect; meanwhile, the game points, represented on a bar of the lightsaber's color, get updated. The lightsaber game is symmetric, meaning that the two players have identical goals and roles in the game.

Figure 2. Virtual Lightsaber Fighting

Figure 3. Subjective QoS-QoE Comparison between Adults and Children (from left to right in descending order of difference)

C. Participants

More than a hundred individuals with no prior experience with 3DTI systems have participated in the study. Fifty participants were able to complete the questionnaires. Among them, ten were female and forty were male. Six were below age 10, thirty-six were between 11 and 20, seven between 21 and 30, and one between 41 and 50. Over 80% of all participants were elementary and middle school children, indicating a possible target population for future 3DTI games. The participants were completely voluntary and not compensated for participation.

D. Methods

We adopt four different subjective measurement approaches to measure the QoS and QoE metrics: questionnaire, interviews, field notes and video taping. Questionnaires were distributed after the games. The questionnaire was designed to quantify the user experience using the QoS-QoE model, and the items therein were developed from the literature (Appendix). Two metrics, performance gains and technology adoption, were not used due to inapplicability. Each item was measured on a seven-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). In addition to questionnaires, interviews helped to explore open observations from the participants and the viewers. Field notes helped us to find out observations from the researchers' points of view. Also, during public experiments, it is very common to miss users' valuable comments and facial expressions. Video taping (with permission) was thus used to gather such data for off-line analysis.

Additionally, we developed a monitoring platform to continuously measure and record the real-world QoS metadata of the 3DTI system in uncontrolled environments. The QoS measurements were taken with time-stamps so that the QoS-QoE correlation could be studied.
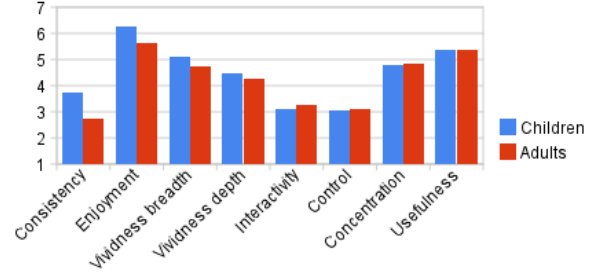
V. FINDINGS AND ANALYSES

A. Age Influence

We are interested in knowing how age impacts user experience. Figure 3 presents a comparison of QoS-QoE rankings by adults (age 21 or above) and children (mostly below 15). The ranks were normalized on the scale of 1 (low/negative) to 7 (high/positive).

From the results, we first note that the children are more tolerant of the inconsistency caused by network congestions with a rank 35.9% higher than adults. This is partly because that inconsistency became less noticeable as they moved their bodies much more vigorously than adults and got fully immersed into the playing.

Second, we observe that the children found the interaction experience much more interesting than the adults with the "Enjoyment" rank 30.4% higher. The children lined up waiting for their turns, enthusiastically cheering for the players. During the game playing, they moved vigorously with excitement.

Finally, adults were more concerned about technology adoption (a behavioral consequence in the QoE construct [25]). An adult participant commented "*This could be a great system for device manufacturing and trouble shooting*". On the other hand, the children were much more concerned about the telepresence and enjoyment. Some of their most frequent comments were "*I am the Jedi!*", "*Does he get hurt?*".

B. Social Aspect Influence

The presence of friends or families significantly promotes engagement of children (both players and observers). Schoolchildren would excitedly acclaim their classmates for defeating the instructors. The players also expressed a much higher level of excitement when there were peers watching.

Children enjoyed playing (competing) much more with their siblings, parents and close friends. *"Let's do this", "Yeah! I'm stabbing her", "Ha, I won again!"*, were some spontaneous comments that reflected the heightened engagement while playing with known peers. Most families and friends requested to play the games multiple times, which

almost never occurred with players who were unknown to each other.

C. Physical Setup Influence

Unlike traditional gaming systems, 3DTI systems require a more complicated physical setup (refer to Figure 1). We find that the placement of different components significantly affects user experience.

First, the relative positions and directions of cameras and displays in a TI environment impacts children's sense of orientation in the games. In 3DTI gaming, gamers have to use their full body to control the virtual photorealistic characters. This involves mapping from the physical 3D coordinate world to a virtual 3D coordinate world through the cameras' fields of view. Moreover, they need to orient themselves by locating or imaging where the remote peer would be in the physical surroundings. In our experiments, the camera and renderer were not aligned on the same plane so that the camera would not obstruct the users to see the screen. With such setup, we find that children had more difficulty orienting themselves than adults. Without being given any instruction, the adult users quickly figured out where to hit in the physical world in order to gain game points. Nevertheless, children tended to follow their experience with 2D game consoles and always acted toward the screens. Therefore, a more intuitive setup for children would be to have a zero angle between the front camera capturing direction and the screen displaying direction. The larger the angle, the more difficult it may be for children to find the correct orientation (refer to Figure 1).

Second, the "sweet area" or activity space where the users can be visible to the cameras should be flexibly configured to accommodate different heights of people (e.g., in families) and different activity patterns. In the lightsaber game, for example, children needed much lower cameras to cover their whole body, and also a larger sweet area to enjoy the game, because they moved much more vigorously than adults and were more immersed into the game.

D. Interactivity and Consistency Influence

Interactivity is the way users interacts with the mediated environments in real time [22]. Among the three factors (*speed*, *range* and *mapping*) that control the interactivity, speed is considered as the most influential metric. In case of TI applications, speed can be measured by the perceived delay from the physical actions to the rendering of the actions. Interactivity is also directly related to consistency. The inevitable network delay from one site to the other results in inconsistency between the remote participants' perceptions.

We monitored the one way delay from the gateway of Site-1 to the gateway of Site-2 during the experiments incurred by our real-time communication protocol based on DCCP. We plot the time-varying delay against different QoE dimensions in Figure 4(a)-(g).

Vividness. If delay increases, users' movements in the real world are reflected with a time offset in the virtual world. In the fast-paced gaming, we found such a noticeable delay affected the perceptive experience with video. Participants of all age group were persistent to rank the video quality as low if the perceived delay increases. Figure 4(a) shows a strong negative correlations between the perceived video quality and the end-to-end delay.

Consistency. Large one way delays create inconsistency, as one player sees his/her action with minimal local delay, while the other player sees it after a long network latency. With the increase of one way delay, even children observed the inconsistent behaviors. Though both parties were fighting, some commented, "*Why isn't he fighting with me?*", "*Why does he think he wins?*". We correlate the delay with such consistency measure, and the resultant graph (Figure 4(b)) shows clear negative correlations.

Flow: Enjoyment. Flow captures the intrinsic psychological experience of users. The intrinsic enjoyment of most users (especially the children) was high even with the existence of large variable delays (Figure 3). The important observation was that some system inefficiencies could be overcome by increasing the impact of flow metrics. Being able to watch one's own 3D appearances in the virtual world interacting with others was something that made the users enjoy the gaming irrespective of delays. Therefore we have not seen a strong correlation between delay and enjoyment as shown in Figure 4(c).

Flow: Concentration. To subjectively measure this QoE flow metric, we asked the users how they were absorbed or involved by the environment and application interaction with the remote parties during the fighting game. We found a strong negative correlation of this metric with delay as shown in Figure 4(d). Children were less concentrated on gaming when there was a large delay. In times of network congestions, some children paused playing after noticing a delay, saying they were confused.

Telepresence. Telepresence or sense of being in the virtual world is also a similar metric that is impacted very little by the delay variance in case of 3DTI gaming for young participants. The correlation graph between delay and telepresence is shown in Figure 4(e). Children felt their existence in the virtual world like a *"warrior"* with the lightsaber and lab coat.

Technology Acceptance. The perceived usefulness and perceived ease of use (or sense of control) are the two belief variables in Technology Acceptance Model (TAM). Perceived control has very little correlation with delay, though perceived usefulness has strong negative correlations (Figure 4(f) and 4(g)). This can be justified as the users
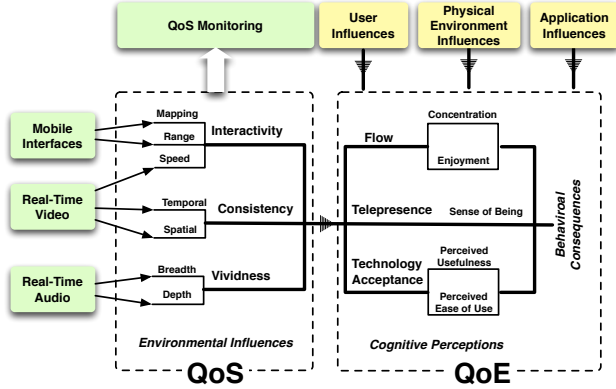
Figure 5. Extensions to Wu's QoS-QoE model

think that reducing the delay would potentially improve the applicability of this system in some other applications like remote training of hardware manufacturing.

Finally, Figure 4(h) shows the overall correlation values between delay and QoS-QoE metrics. As we can see, only video quality, consistency and concentration have higher correlations with delay (interactivity and consistency), whereas enjoyment, telepresence, perceived control and perceived usefulness are not strongly correlated with it.

E. Subjective Correlation Findings

We measured the subjective QoS and QoE perceptions of users, and computed the correlations between the constructs. The results are plotted in Figure 4(i) with only the statistically significant links shown ($p < 0.001$ with two-tail t-test, Table I).

First, we observe that the subjectively perceived interactivity is correlated with flow concentration. This is consistent with our finding on (objective) delay-concentration correlation (Sec. V-D). Second, we note that vividness (video quality) is correlated with telepresence. Higher 3D realism will result in a higher sense of telepresence. This calls for more advanced 3D reconstruction algorithms that produce higher accuracy with smaller real-time delays. Finally, the subjective consistency is also positively related with telepresence. Creating a temporally consistent virtual environment for all remote users is crucial in enhancing user sense of presence in the interaction.

VI. DISCUSSIONS

Although Wu *et al.*'s QoS-QoE model [25] provides us theoretical guidance on which QoS metrics may impact QoE, it focuses only on the technical environmental influences (e.g., response time, visual quality), and fails to account for the important application influences (e.g., game design), user influences (e.g., age, social aspects), and physical environment influences (e.g., device placement). Therefore, we propose to extend the model to incorporate these dimensions, as illustrated in Figure 5.

The 3D photo-realism and interactivity with remote peers with physical exertions are the most critical factors that stimulate people's interest in the new technology. In particular, children - the potential major population for 3DTI games - expressed tremendous excitement with the new gaming experience. Context vividness is less of a concern for them; rather, the 3D photorealistic telepresence in a virtual world with real opponents is the key factor.

In our experiments, we allowed people to communicate verbally over a VoIP module. Additionally, we developed a mobile interface for people to watch the lightsaber duel and freely control their viewpoint from an iPhone. However, the audio and mobile components did not show strong contributions to improving user experience. This reflects the fact that QoS performances are not always as critical as engineers might expect for gaming experience; often times the non-technical aspects become decisive. Voice communication is useful for remote communication; however, it is recommended not to instrument the gamers even with wireless headphones; microphone arrays is a preferred option. Mobile interfaces provide flexible control options to manipulate the 3DTI cyber-environment. Nevertheless, the users experience greater enjoyment in the games if they can control the cyber-environment (i.e., their 3D photorealistic characters) with full body motion. The use of mobile devices should not interfere or obstruct the natural body movement of gaming.

Moreover, social aspects play an important role in the gaming experience. Due to psychological matureness, adults are not as interested in gaming as children, rather they are more concerned about the perceived usefulness of the technology. Children enjoyed the games much more when playing with their parents, siblings, classmates, or instructors. As noted by Schiesel [20], "Paradoxically, at a moment when technology allows designers to create ever more complex and realistic single-player fantasies, the growth in the now $18 billion gaming market is in simple, user-friendly experiences that families and friends can enjoy together." In 3DTI gaming, a great deal of social interaction is interleaved between the physical world and the virtual world. Our study shows that promoting such social interaction with games of moderate difficulty levels can significantly enhance the gaming experience.

Last but not least, unlike the existing gaming consoles, the 3DTI cyber-physical gaming environment also poses new challenges in the physical setup of different components. The capturing and rendering devices should be carefully aligned to facilitate quick orientation of children in the cyber-physical space. Improper arrangement of the physical space can result in confusion, frustration, and loss of interest, hence demotes the gaming experience.
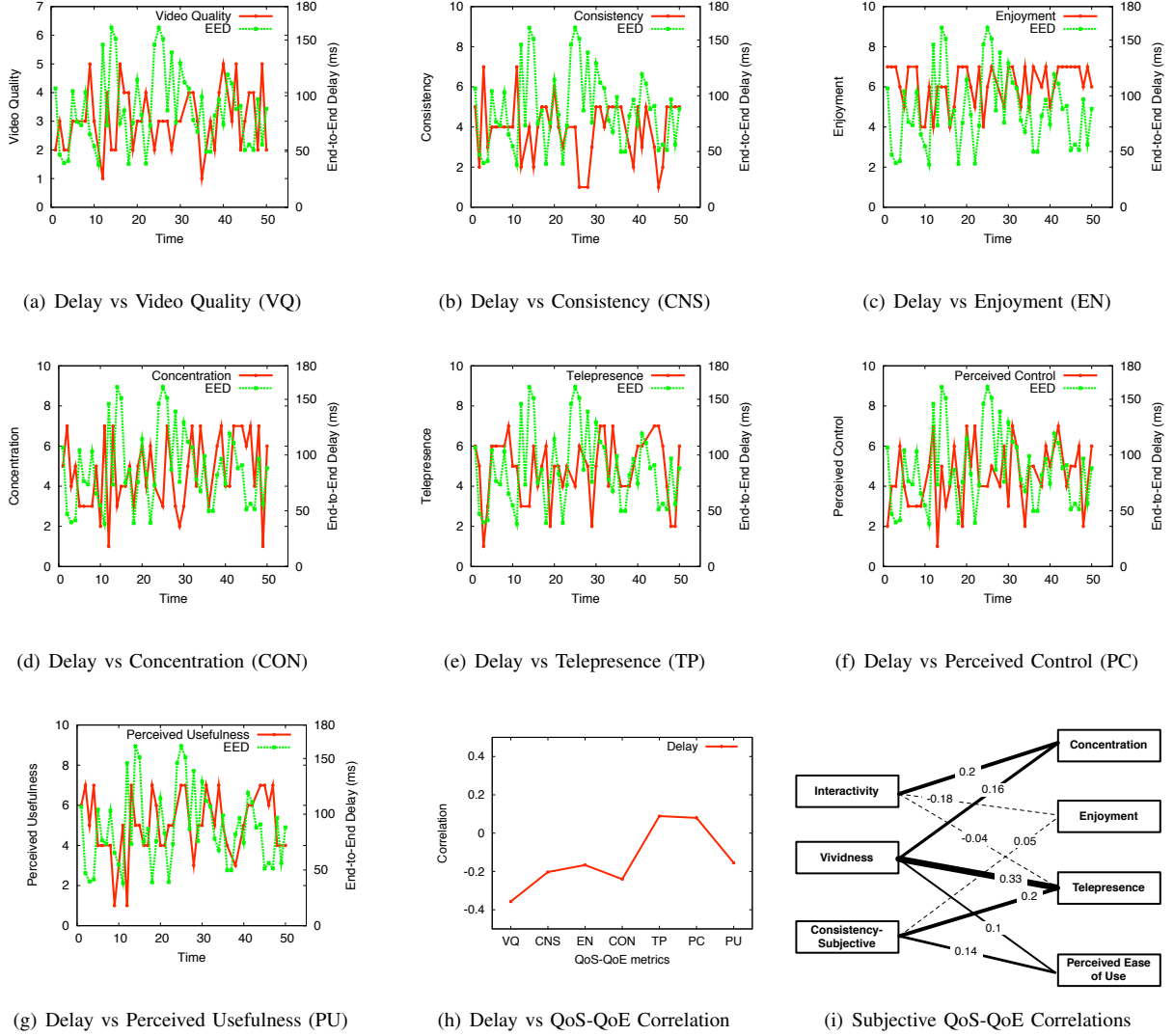
(a) Delay vs Video Quality (VQ)

(b) Delay vs Consistency (CNS)

(c) Delay vs Enjoyment (EN)

(d) Delay vs Concentration (CON)

(e) Delay vs Telepresence (TP)

(f) Delay vs Perceived Control (PC)

(g) Delay vs Perceived Usefulness (PU)

(h) Delay vs QoS-QoE Correlation

(i) Subjective QoS-QoE Correlations

Figure 4.   QoS-QoE Correlation Results

Table I
DATA ANALYSIS (T-VALUE/P-VALUE)

	Concentration	Enjoyment	Teleprescence	Perceived Ease of Use	Perceived Usefulness
Interactivity	$-4.51/\sim 0$	$-9.21/\sim 0$	$-4.75/\sim 0$	$0.09/0.92$	$-5.46/\sim 0$
Vividness	$-1.24/\sim 0$	$-7.34/0.004$	$-1.47/\sim 0$	$-4.02/0.92$	$-2.4/\sim 0$
Consistency	$-3.41/\sim 0$	$-8.21/\sim 0$	$-3.64/\sim 0$	$1.24/0.22$	$-4.39/\sim 0$

VII. CONCLUSION

We have conducted a case study to explore two main questions: How do the technical QoS metrics impact QoE in 3DTI gaming? How do the non-technical factors (e.g., age, social interaction, presence of others) influence QoE? The study reveals new findings that suggest extension and generalization to the existing theoretical QoS-QoE framework [25], and helps to understand user experience with the next-generation "exergaming" technologies.

VIII. ACKNOWLEDGEMENT

REFERENCES

[1] K.-T. Chen et al. How sensitive are online gamers to network quality? *ACM Communications*, 49(11):34–38, 2007.

[2] N. Ducheneaut et al. "Alone Together?" exploring the social dynamics of massively multiplayer online games. In *Proc. ACM CHI'06*, pages 407–416, 2006.

[3] N. Ducheneaut et al. Virtual "third places": A case study of sociability in massively multiplayer games. *CSCW*, 16(1–2):129–166, 2007.

[4] P. Hämäläinen et al. Martial arts in artificial reality. In *Proc. ACM CHI'05*, pages 781–790, 2005.

[5] http://us.playstation.com/index.htm.

[6] http://wii.com/.

[7] http://www.xbox.com/en US/.

[8] http://www.xbox.com/en us/live/projectnatal/.

[9] Z. Huang et al. Tsync: A new synchronization framework for multi-site 3d tele-immersion. In *Proc. ACM NOSSDAV*, June 2010.

[10] W. Ijsselsteijn et al. Digital game design for elderly users. In *Proc. Future Play*, pages 17–22, 2007.

[11] ITU-G.114. One-way transmission time, 2003.

[12] R. Jain. Quality of experience. In *IEEE Multimedia*, volume 11, pages 95–96, 2004.

[13] G. Kurillo et al. Teleimmersive environment for remote medical collaboration. *MMVR17*, 142:148–150, 2009.

[14] S. Lindley et al. Stirring up experience through movement in game play: Effects on engagement and social behaviour. In *Proc. ACM CHI'08*, pages 511–514, 2008.

[15] F. Mueller et al. Exertion interfaces: Sports over a distance for social bonding and fun. In *Proc. ACM CHI'03*, pages 561–568, 2003.

[16] B. Nardi et al. Strangers and friends: Collaborative play in World of Warcraft. In *Proc. ACM CSCW'06*, pages 149–158, 2006.

[17] D. E. Ott et al. Coordinated multi-streaming for 3d tele-immersion. In *Proc. ACM MM'04*, pages 596–603, 2004.

[18] K. I. Rena et al. "We have never-forgetful flowers in our garden:" girls' responses to electronic games. Technical report, University of British Columbia, 1994.

[19] B. Sat et al. The design of a multi-party VoIP conferencing system over the Internet. In *Proc. IEEE ISM'07*, pages 3–10, 2007.

[20] S. Schiesel. In the list of top-selling games, clear evidence of a sea change. In *http://nyti.ms/9rD0Eq*, Accessed 2010.

[21] R. Steinmetz. Human perception of jitter and media synchronation. *IEEE JSAC*, 14(1):61–72, 1996.

[22] J. Steuer. Defining virtual reality: Dimensions determining telepresence. In *Journal of Comm.*, pages 73–93, 1992.

[23] C. Szentgyorgyi et al. Renegade gaming: practices surrounding social use of the nintendo ds handheld gaming system. In *Proc. ACM CHI'08*, pages 1463–1472, 2008.

[24] A. Voida et al. Wii all play: The console game as a computational meeting place. In *Proc. ACM CHI'09*, pages 1559–1568, 2009.

[25] W. Wu et al. Quality of experience in distributed interactive multimedia environments: Toward a theoretical framework. In *Proc. ACM MM'09*, pages 481–490, 2009.

[26] Z. Yang et al. TEEVE: The next generation architecture for tele-immersive environments. In *Proc. IEEE ISM*, 2005.

[27] Z. Yang et al. Collaborative dancing in tele-immersive environment. In *Proc. ACM MM'06*, pages 723–726, 2006.

APPENDIX

Please circle a score from the scale 1 (strongly disagree) to 7 (strongly agree) below which most closely corresponds with how you perceive while playing the games.

(a) *Vividness breadth*: It is important for me to have both audio and video when interacting with a remote partner.

(b) *Vividness depth*: The image quality was sufficient for effective interaction.

(c) *Interactivity Speed*: The delay between my physical actions and the showing of myself on the screen is significant. *(reverse-scaled)*

(d) *Enjoyment*: I found the interaction experience interesting.

(e) *Telepresence*: During the interaction, my body was in the room, but my mind focus was inside the virtual space with the remote person.

(f) *Consistency*: During the interactions of our session, there were specific moments when I could tell that my partner and I were not seeing the same thing on our screens. *(reverse-scaled)*

(g) *Concentration*: I was absorbed by the experience of interacting with my remote partner - it held my attention fully.

(h) *Perceived Ease of Control*: At some point I felt confused - the system didn't behave the way I expected. *(reverse-scaled)*

(i) *Perceived Usefulness*: I believe the system will be useful for remote collaboration.

(j) *Perceived Usefulness*: I could imagine that the virtual interaction environment has the potential of nurturing creativity - people can do things that are physicall impossible but digitally possible.