Evaluating User Performance in 3D Stereo and Motion Enabled Video Games

Arun Kulshreshth
Department of EECS
University of Central Florida
4000 Central Florida Blvd
Orlando, FL 32816, USA
arunkul@knights.ucf.edu

Jonas Schild
Entertainment Computing Group
University of Duisburg-Essen
Forsthausweg 2
47057 Duisburg, Germany
jonas.schild@uni-due.de

Joseph J. LaViola Jr.
Department of EECS
University of Central Florida
4000 Central Florida Blvd
Orlando, FL 32816, USA
jjl@eecs.ucf.edu

ABSTRACT

We present a study that investigates user performance benefits of playing video games using 3D motion controllers in 3D stereoscopic vision in comparison to monoscopic viewing. Using the PlayStation 3 game console coupled with the PlayStation Move Controller, we explored five different games that combine 3D stereo and 3D spatial interaction. For each game, quantitative and qualitative measures were taken to determine if users performed better and learned faster in the experimental group (3D stereo display) than in the control group (2D display). A game expertise pre-questionnaire was used to classify participants into beginners and expert game player categories to analyze a possible impact on performance differences. The results show two cases where the 3D stereo display did help participants perform significantly better than with a 2D display. For the first time, we can report a positive effect on gaming performance based on stereoscopic vision, although reserved to isolated tasks and depending on game expertise. We discuss the reasons behind these findings and provide recommendations for game designers who want to make use of 3D stereoscopic vision and 3D motion controllers to enhance game experiences.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Evaluation/methodology;
K.8.0 [**Personal Computing**]: Games

Keywords

3D Stereo, video games, PlayStation 3, Move Controller, 3D interaction, user performance & experience.

General Terms

Experimentation, Measurement, Performance

1 INTRODUCTION

3D stereo has been limited to research labs and commercial setups (like IMAX theaters) but recent advances in technology have made it possible at the consumer level. The electronics industry is now pushing this technology by releasing devices (TVs, blu-ray players, and game consoles) which are 3D compatible and the entertainment industry is promoting it by releasing 3D blu-ray movies and especially 3D enabled games. As the technology has started to become more available to consumers, game designers and hardware manufacturers have started to take notice. However, it is unclear how 3D stereo affects gameplay and the user experience. Are there any measurable benefits when playing games using stereoscopic vision? In other words, do gamers gain a performance advantage when using a 3D stereo display and, if so, why? By understanding these performance benefits and the reasoning behind them, we hope to gain insights into ways that games can be made more enjoyable and help users to play them more effectively. This motivation is especially important in the serious games domain which targets education-based games, training in military simulations, obesity reduction and physical activity.

A previous study [9] for PC games showed that playing games in 3D stereo does not provide any significant performance benefits than with using a 2D display. However, this study used a traditional game controller (the Xbox 360 controller) as the interaction device and the games used were not designed with 3D stereo in mind. It has been shown that an increase in body movement imposed, or allowed, by the game controller results in an increase in players engagement level [2, 8]. In the cognitive science literature, it has been shown that there is a connection between actions and depth perception in that motor actions affect our perception of 3D space and objects [19]. Thus, allowing observers to physically act can drastically change the way they perceive the third dimension. Research has also shown that 3D stereo can be beneficial to user performance in certain, isolated tasks in the context of virtual reality and 3D user interfaces [1, 6, 16] using a 6 degree of freedom (6 DOF) input device. Based on these studies, we hypothesize that coupling 3D stereo with 3D spatial interaction using motion controllers in video games could lead to better user performance than with a 2D display and a motion controller.

In this paper, we present a study investigating whether user performance is enhanced when using 3D stereo over a traditional 2D monitor coupled with a 3D spatial interaction device in modern 3D stereo games. We made use of the PlayStation 3 game console, the PlayStation Move controller, and five carefully chosen game titles as a representative sample of modern motion enabled games. These games all had 3D tasks that, in our judgment, may potentially benefit from 3D stereo. To evaluate player performance, we collected both quantitative data based on the tasks associated with each game and qualitative data based on post-questionnaires to gauge user perception of their performance. We used a between subjects design where the control group played the games using a 2D monitor and the experimental group played with a 3D stereo display with both groups using the PlayStation Move controller as the input device.

2 RELATED WORK

Stereoscopic 3D technology has been around for decades and been found to be beneficial depending on the task involved [1, 6, 12, 16]. Hubona et al. found that 3D stereo in a game is a more effective depth cue than shadows in a user's perception of 3D, based on accuracy and speed with which users completed gaming tasks [6]. Stereo has also been found helpful in playing simple games where a user is manipulating a single object at a time [3]. Litwiller and LaViola explored benefits of 3D stereo in modern PC based games using the Nvidia 3D Vision Kit. No significant advantage was found in user performance over a 2D display [9]. Another study evaluated game performance with a shooter game on autostereoscopic displays [13] but found no differences in performance for stereo vs. monoscopic vision modes. The same study further revealed that the 3D display mode evoked significantly higher positive emotions and stronger feelings of presence than the 2D mode and was highly preferred by a large majority of the participants. An increased engagement and preference for stereoscopic games was also confirmed in [9] and [14]. The latter further found effects varying strongly across different games and target groups. Stereo evoked higher immersion and presence, especially in males, and affected attention in a way that indicates a more natural, less self-reflective gameplay. Stereoscopic 3D viewing has also been found to improve performance in medical applications (e.g., robotic-assisted laparoscopic surgeries on bench models [12] or endoscopic tasks [20]).

The interplay between 3D stereo and interaction technique has led to conflicting results. In one study, the interaction technique was found to be significant while stereo was not [10]. However, this finding has been somewhat contradicted by Teather and Stuerzlinger, who presented different positioning techniques that were dependent on the input devices used. They found that stereo was beneficial for accuracy in the tasks they presented to users, but not for speed [16].

The interplay between 3D stereo and motion cues has been studied in very simple tasks. Ware and Mitchell found roughly an order of magnitude increase in the size of a graph that can be read at similar performance levels when 3D stereo viewing is available along with motion depth cues [18]. Merritt et al. [11] studied effects of motion parallax and stereoscopic 3D in a task to touch, in sequence, ten target sites embedded in a complex wire maze. They found a large significant advantage for the 3D stereoscopic display condition vs. the 2D condition and a smaller significant advantage for the motion-parallax vs. the static condition.

While 3D stereo has shown some positive benefits depending on the task, it also has shown to cause negative symptoms as well, such as eyestrain, headache, dizziness, and nausea [5]. Thus, as a secondary objective of our work, we wanted to see if these symptoms exist while playing motion controlled 3D stereoscopic video games on a 3D TV. We believe our work is the first to explore whether any performance benefits exist when using 3D stereo coupled with a 3D motion controller in complex video game style environments.

3 SELECTING THE GAMES

We required a gaming environment that natively supported 3D stereo and 3D spatial interaction. At the time of this work, the only system that supported both these features was the PlayStation 3. We were able to find 16 games that supported both 3D stereo as well as the PlayStation Move controller and we examined them to see if playing them in 3D stereo provides any performance benefits. Some of these games had all their tasks in 2D so playing them in 3D stereo would not provide any benefit. Some games make use of the PlayStation Move controller for just a few tasks just to label the game as Move compatible. We removed all such games and we were left with just eight games. Out of those eight games, we

removed three more games which we felt did not use 3D efficiently or had poor interface controls deterring any performance benefit. We were left with five games, Hustle Kings, Pain, The Fight: Lights out, Tumble, and Virtua Tennis 4, that could potentially provide performance benefits in a 3D stereo environment.

Hustle Kings is a pool table game which uses the Move controller as a cue stick used to hit the cue ball. The game displays an aiming line while adjusting the cue stick before taking a shot which indicates where the ball is going to hit. We disabled the aiming line so we could better judge how users performed with the coupling of 3D stereo and 3D spatial interaction.

Pain has a level "PainMotion: Skeet X3" which supports both 3D stereo as well as the Move controller. In this level, the player has to destroy all the incoming objects thrown at him or her by throwing a bomb before being hit by those objects. If the bomb misses any incoming object then the player will get hit and loses health and eventually dies. The Move controller is used to aim in the direction that the player wants to throw the bomb. We expected that the 3D stereo display and the Move controller together would let players perform better.

The Fight: Lights Out is a boxing game in which two Move controllers are used as two hands of the player. The player had to fight against an CPU controlled opponent. We thought that 3D stereo would help support better aiming when throwing punches at an opponent.

Tumble is a game which involves manipulation of 3D blocks of different shapes, materials, and sizes. This game involves many 3D selection and manipulation tasks. We chose a level called "Zone 2: Variety box" in which the player has to choose different objects and fit as many as possible on a table. The Move controller is used to pick and manipulate the objects. We felt that judgment of depth is critical when placing objects accurately so 3D stereo could be beneficial in this game.

Virtua Tennis 4 is a first person tennis game in which the Move controller is used as a tennis racquet. While playing this game, knowledge of distance of the tennis ball is necessary to time the racquet swing to hit the ball accurately. We thought that 3D stereo could be helpful in playing this game by achieving better ball hit accuracy. Moreover depth information could be used to hit the ball in different directions while playing.

4 USER STUDY

We conducted a usability experiment with five PlayStation 3 3D games (as discussed in the previous section) where participants played each game in either a 2D or 3D viewing mode using the PlayStation Move Controller. We examined both quantitative metrics based on each game's goals and tasks and qualitative metrics based on whether participants preferred playing the games in 3D and whether they perceived any benefits. Based on previous findings in related work and our analysis of the games, we hypothesize that users will gain a performance advantage when using 3D stereo coupled with a 3D interaction device over a 2D display coupled with the same 3D interaction device. In addition, we felt players would prefer playing games in 3D stereo coupled with the 3D interaction device (Move Controller) because it provides a more engaging user experience.

4.1 Participants and Equipment

Fifty participants (38 males and 12 females ranging in age from 18 to 34 with a mean age of 23.04) were recruited from the University of Central Florida. A modified version of Terlecki and Newcombe's Video Game Experience survey [17] was used as a pre-questionnaire in which they answered questions about their previous gaming experience. The survey was modified to include questions related to

Figure 1: The experimental setup used a Samsung 50 inch DLP 3D HDTV, PlayStation 3 game console, PlayStation Move controller, and a Mitsubishi 3D adapter (to convert the HDMI 1.4a 3D signal to a DLP 3D signal). These are all commodity hardware components.

previous experience with the PlayStation 3, the Move Controller, and the stereo 3D games used for the study. Of the 50 participants, 18 were ranked as beginners (10 in stereo group and 8 in non-stereo group), 23 as intermediate (9 in stereo group and 14 in non-stereo group), and 9 as advanced (6 in stereo group and 3 in non-stereo group). Since there were only a few advanced users, we decided to combine intermediate and advanced categories into one category called expert users. The experiment duration ranged from 75 minutes to 90 minutes depending on how long participants took to complete the tasks presented to them in the games and how much time was spent on the questionnaire. All participants were paid $10 for their time.

The 3D setup (see Figure 1) used for the study consisted of a PlayStation 3 with PlayStation Eye camera, PlayStation Move Controllers, Mitsubishi 3D adapter (to convert HDMI 1.4a signal to DLP 3D signal), Samsung 50 inch DLP 3D HDTV, and Optoma (DLP Link) active shutter glasses. For the 2D display condition, the Mitsubishi 3D adapter was not used and PlayStation 3 rendered graphics in 2D.

4.2 Experimental Task

The participants were given tasks to play through levels of the five games that were selected for this study. For each game, they were presented with a task specific to that game and a goal for completing each task. Participants played these games in random order (counter-balanced Latin Squares design) with three attempts for each game.

Hustle Kings: The participants played "Free Play" with no opponent. Their task was to pocket as many balls (with no constraint on the ball color) as possible in six shots (including the first hit). The table was reset after each attempt (consisting of 6 shots).

Pain: In this game participants played "PainMotion: Skeet X3" and their task was to destroy as many objects (thrown at them) as possible by throwing a bomb. The better they play the longer they survive in the game.

The Fight: Lights Out: The participants played "First Fight" using two Move controllers. Their task was to fight and defeat the opponent.

Tumble: The participants played "Zone 2: Variety Box" where the task was to put as many blocks as possible on a table in the game. The size of the table was limited so they had to cleverly arrange the objects on the table in order to stack other objects on

Table 1: Summary of metrics for each game. The metrics are used to quantify how users in the 2D and 3D display groups performed.

Game	Metric
Hustle Kings	Number of balls pocketed
Pain	Time
	Accuracy of throws
The Fight: Lights Out	Calorie Count
	Accuracy of punches
Tumble	Time taken
	Number of blocks used
Virtua Tennis 4	Match won or lost

top of them. The maximum time limit for each attempt was ten minutes. The time limit of 10 minutes was determined during pilot testing with this game. The moderator kept track of time taken by participants for each attempt.

Virtua Tennis 4: The participants had to play "Motion Play" using the move controller as a tennis racquet. Participants played three matches against a randomly chosen opponent. The moderator kept track of score and time taken by participants for each attempt.

4.3 Design and Procedure

Our study design was based, in part, on the study done by Litwiller and LaViola [9]. We chose a between subjects design, to avoid any effects of learning on user performance, where the independent variable was display mode (2D display or 3D stereo display) and the dependent variables were the various scoring metrics used in each game. We wanted some additional information about the use of 3D stereo in video games for those who played the games in the 2D display condition. Thus, we chose to have those participants who were in the 2D display condition, pick one game to try in 3D stereo to gather their reactions.

In order to group the participants into expertise levels based on the pre-questionnaire data, we scored the questionnaire by assigning points to each question. Particular questions were given more points based on how the results fit within the context of our experimental setup. For example, participants who owned gaming consoles and have been playing regularly were considered to have a higher expertise level. We then used the raw scores from adding up the points for each question to group the participants into the appropriate category. Both the quantitative and qualitative data was explored collectively as well as according to the two groupings (beginners and experts).

4.3.1 Quantitative and Qualitative Metrics

For each game, we tracked quantitative data that we felt was a good indication of how well the user performed. Quantitative metrics are summarized in Table 1.

In Hustle Kings, we kept a record of the number of balls pocketed in each run consisting of 6 shots. In Pain, player survival time and bomb throwing accuracy (calculated from number of hits and throws reported by the game) were tracked as performance metrics. In The Fight, calories burned and punch accuracy (reported by the game) were tracked as performance metrics. In Tumble, the number of blocks used and level completion time with a maximum limit of 10 minutes per trial were tracked. In Virtua Tennis 4, match outcome (lost or won) was the only performance metric.

For qualitative data, all participants filled out an immersion questionnaire [7] (see Table 2) upon completion of all trials of each game. The questionnaire was modified to include two questions about controllers to determine the helpfulness of the Move con-

Table 2: Post-game Questionnaire. Participants answered these questions on a 7 point Likert scale after playing each game. We used this data for qualitative analysis.

	Postgame Questions
Q1	To what extent did the game hold your attention?
Q2	How much effort did you put into playing the game?
Q3	Did you feel you were trying your best?
Q4	To what extent did you lose track of time?
Q5	Did you feel the urge to see what was happening around you?
Q6	To what extent you enjoyed playing the game, rather than something you were just doing?
Q7	To what extent did you find the game challenging?
Q8	How well do you think you performed in the game?
Q9	To what extent did you feel emotionally attached to the game?
Q10	To what extent did you enjoy the graphics and the imagery?
Q11	How much would you say you enjoyed playing the game?
Q12	Would you like to play the game again?
Q13	Was use of Move Controller helpful in playing the game?
Q14	To what extent your arm was tired after playing the game?
Q15	DualShock 3 would be a better choice to play this game?

Table 3: 3D Stereo Questionnaire. Participants responded to statements 1-4 on a 7 point Likert scale. Questions 5-10 were multiple choice and open ended questions to gauge the users perception of the effects of 3D stereo. In question 11, each symptom had a 7 point Likert scale to indicate the extent of each symptom ranging from not at all to very much so.

	3D Stereo Questions
Q1	3D stereo improved the overall experience of the game.
Q2	I would choose to play in 3D stereo over normal viewing.
Q3	I felt that stereo enhanced the sense of engagement I felt.
Q4	3D stereo is a necessity for my future game experiences.
Q5	Did 3D stereo help you perform better in the games?
Q6	Which games did it help you in?
Q7	How did it help you in those games?
Q8	Did 3D stereo decrease your performance in the games?
Q9	Which games did it decrease your performance in?
Q10	How did it decrease your performance in those games?
Q11	Did you feel any symptoms from viewing the games in stereo (eye strain, headache, dizziness, nausea)?

troller and the traditional controller (DualShock 3). Another question was included to determine if the player experienced arm fatigue from using the Move controller. Responses were measured on a 7 point Likert scale (1 = most negative response, 7 = most positive response). Upon completion of all experimental tasks participants were given a survey to determine how 3D stereo affected their gaming experience (see Table 3) and whether they preferred to play the games in 3D stereo and if 3D stereo helped or hurt their performance.

4.3.2 Procedure

The experiment began with the participant seated in front of the TV and the moderator seated aside. Participants were given a standard consent form that explained the study. They were then given a pre-questionnaire that focused on their gaming expertise. Participants were then presented with the games in random order (Latin Squares design). Half the participants played the games in 2D display mode (control group) and half played in 3D stereo (experimental group). The moderator would present the game and

Table 4: Two-way ANOVA analysis for Hustle Kings. No significance was found.

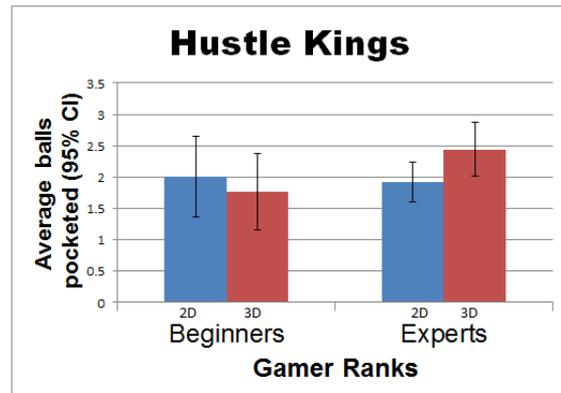| Source	Average Balls Pocketed
DM	$F_{1,46} = 1.491, p = 0.228$
EXP	$F_{1,46} = 0.348, p = 0.558$
EXP $\times$ DM	$F_{1,46} = 2.374, p = 0.130$

Figure 2: Hustle Kings: Differences in average number of balls pocketed between the 2D and 3D groups in the two gamer categories. Expert gamers did better in 3D mode.

give instructions to the participant as to what they needed to accomplish in the game and what their goals were. They were also instructed on how to use the PlayStation Move controller. During the experiment, the moderator recorded quantitative data using scores from the games and a stopwatch for timing information. After each game, the participant filled out a post-questionnaire with questions about their experiences with the game. If the participants played the five games in the 2D display group, they then selected one game to play in 3D stereo. All participants were given a final post-questionnaire about their experiences with the 3D stereo display.

5 RESULTS AND ANALYSIS

We broke up the participants in each display group (3D and 2D group) into beginners (8 participants in the 2D display group, 10 participants in the 3D stereo group) and expert gamers (17 participants in the 2D display group, 15 participants in the 3D stereo group). To analyze the performance data, a two-way ANOVA was conducted that examined the effect of game-play expertise (EXP), beginner or expert, and the display mode (DM) on the user performance (see Table 1 for metrics used for each game). We did a post-hoc analysis using independent sample t-tests. We used Holm's sequential Bonferroni adjustment to correct for type I errors [4] and the Kolmogorov-Smirnov test to make sure our data is parametric. We also wanted to see whether there was learning taking place in the form of game play improvement. We looked at the improvement in the performance measures for each game from the first user run to their last run using a repeated measures ANOVA. Finally we wanted to look at the participant's perception of their performance through the post questionnaires. To analyze this Likert scale data, we used the Mann-Whitney test. For all of our statistical measures, we used $\alpha = 0.05$.

5.1 Hustle Kings

No statistically significant differences were found between overall mean performance scores of the two display mode groups (see Ta-

Table 5: Two-way ANOVA analysis for Pain. Difference due to game play expertise was significant.

Source	Time	Accuracy
DM	$F_{1,46} = 1.702, p = 0.199$	$F_{1,46} = 2.251, p = 0.140$
EXP	$F_{1,46} = 17.109, p < 0.01$	$F_{1,46} = 32.936, p < 0.01$
DG × EXP	$F_{1,46} = 0.313, p = 0.579$	$F_{1,46} = 0.000, p = 0.992$

ble 4 and Figure 2). However, when we looked at the individual runs for the expert gamers, we did find a significant difference ($t_{30} = -2.79, p < 0.01$) between the balls pocketed in their last attempt, with the 3D stereo group pocketing an average of 2.93 balls ($\sigma = 1.33$) compared to an average of 1.65 balls ($\sigma = 1.27$) for the 2D display group. In terms of score improvement, beginners had no significant score improvements from the first attempt to the last attempt for both the 3D display ($F_{2,8} = 1.964, p = 0.169$) and the 2D display ($F_{2,6} = 0.467, p = 0.637$) groups. For expert gamers, the 3D stereo group ($F_{2,13} = 3.530, p < 0.05$) significantly improved the average number of balls pocketed from 2.00 in the first run to 2.93 in the third run (46.5% improvement). No significant score improvements were found for the 2D display group ($F_{2,15} = 0.888, p = 0.421$). From the qualitative data, beginners in the 2D display group ($\bar{x} = 6.62, \sigma = 0.517$) enjoyed the graphics and imagery significantly more ($Z = -2.563, p < 0.05$) than the 3D stereo group ($\bar{x} = 5.4, \sigma = 1.173$). There were no significant differences in any of the questions for expert gamers. Overall, participants thought that the DualShock 3 controller would not be a good choice for this game ($\bar{x} = 3.06, \sigma = 2.20$).

5.2 Pain

Table 5 shows the 2-way ANOVA analysis for completion time and accuracy. Expert gamers performed significantly better than beginners in terms of time ($t_{48} = -4.029, p < 0.025$) and accuracy ($t_{48} = -5.609, p < 0.025$). For both beginners and expert gamers, the 3D stereo group performed slightly better than the 2D display group (see Figure 3), but the differences were not significant. For score improvement, beginners showed no significant improvement in either the 2D display group ($F_{2,6} = 0.008, p = 0.992$) or the 3D stereo group ($F_{2,8} = 1.444, p = 0.262$). For expert gamers, we did not find any significant improvements from their first attempt to their last attempt in the 2D display group ($F_{2,15} = 0.513, p = 0.604$) or the 3D stereo group ($F_{2,13} = 1.066, p = 0.358$). However, we did find significant improvements from the first attempt to the second attempt in the 3D stereo group ($F_{1,14} = 5.202, p < 0.05$) with no such significant differences in the 2D display group ($F_{1,16} = 1.546, p = 0.232$). The 3D stereo group improved their accuracy from 58.15% ($\sigma = 13.06$) in their first attempt to 61.73% ($\sigma = 12.03$) in their second attempt which is a 6.15% improvement. For the questionnaire data, people thought that the DualShock 3 controller would not have been a good choice for this game ($\bar{x} = 2.80, \sigma = 2.27$). For beginners, participants in the 2D display group ($\bar{x} = 6.50, \sigma = 0.755$) put significantly more effort ($Z = -2.002, p < 0.05$) into playing the game than the 3D stereo group ($\bar{x} = 5.5, \sigma = 1.08$) while the expert gamers did not show any significance for effort between the 2D and 3D stereo groups ($Z = -1.659, p = 0.097$). All other Likert scale questions between the two groups were not significant.

5.3 The Fight: Lights Out

No statistically significant differences were found (see Table 6 and Figure 4) based on display group or game play expertise. In terms of score improvement, there was no significant improvement for beginners from their first to the last attempt in the 2D display group ($F_{2,6} = 1.110, p = 0.357$). However, we did find significant
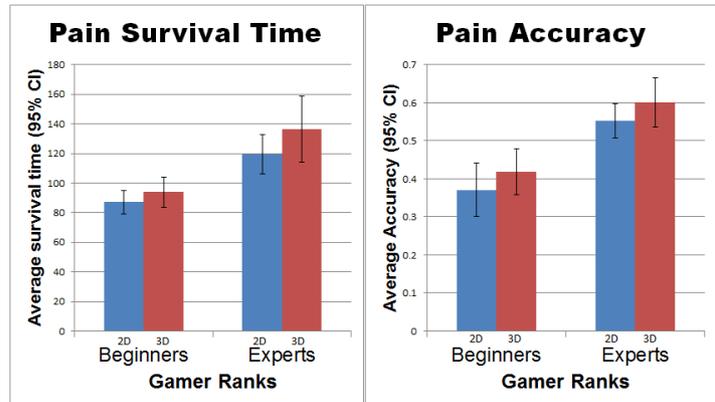

Figure 3: Pain: Differences in survival time and accuracy between the 2D and 3D groups in the two gamer categories. Players survived slightly longer in 3D stereo in both gamer ranks. People had slightly more hit accuracy in 3D stereo compared to the 2D display group.

Table 6: Two-way ANOVA analysis for The Fight. No significance was found.

Source	Calories	Accuracy
DM	$F_{1,46} = 0.230, p = 0.634$	$F_{1,46} = 0.58, p = 0.811$
EXP	$F_{1,46} = 1.599, p = 0.212$	$F_{1,46} = 0.320, p = 0.574$
DG × EXP	$F_{1,46} = 0.273, p = 0.604$	$F_{1,46} = 0.033, p = 0.857$

improvements in the 3D stereo group ($F_{2,8} = 4.870, p < 0.05$). The beginner 3D stereo group improved their accuracy from 66.4% ($\sigma = 12.08$) in their first attempt to 74.0% ($\sigma = 5.37$) in their last attempt, a 11.44% improvement. In the case of expert gamers, there were significant improvements in accuracy for both the 2D display group ($F_{2,15} = 11.662, p < 0.05$) and the 3D stereo group ($F_{2,13} = 5.511, p < 0.05$) from their first to last attempt. The 2D display group improved their accuracy from 64.29% ($\sigma = 7.74$) to 73.05% ($\sigma = 9.90$), a 13.62% improvement and the 3D stereo group improved their accuracy from 67.53% ($\sigma = 9.24$) to 73.80% ($\sigma = 8.97$), a 9.28% improvement. Both groups felt their arms got tired after playing this game ($\bar{x} = 5.58, \sigma = 1.72$). For beginners, the game held their attention significantly more ($Z = -1.954, p < 0.05$) with the 2D display group ($\bar{x} = 7.0, \sigma = 0$) group than with the 3D stereo group ($\bar{x} = 6.3, \sigma = 1.05$).

5.4 Tumble

Table 7 shows the results from the 2-way ANOVA analysis for Tumble. For beginners, the 3D stereo group ($\bar{x} = 15.333, \sigma = 2.504$) performed significantly better ($t_{16} = -2.628, p < 0.025$) than the 2D display group ($\bar{x} = 12.375, \sigma = 2.19$) for average number of blocks used. For expert gamers, no significant differences between average completion times ($t_{30} = -0.233, p = 0.818$) or average number of blocks used ($t_{30} = -0.306, p = 0.762$) was found between groups (see Figure 5 for plots).

For score improvement, beginners showed no significant improvements from their first attempt to last attempt in the number of blocks used for either the 3D stereo group ($F_{2,8} = 0.507, p = 0.611$) or the 2D display group ($F_{2,6} = 1.661, p = 0.225$). However, for expert gamers, we found significant improvements for blocks used in the 2D display group ($F_{2,15} = 5.759, p < 0.05$) but not in the 3D stereo group ($F_{2,13} = 0.781, p = 0.468$). The 2D display group improved the number of blocks used from 14.64 ($\sigma = 4.51$) in their first attempt to 18.94 ($\sigma = 3.36$) in their last attempt, a 29.37%
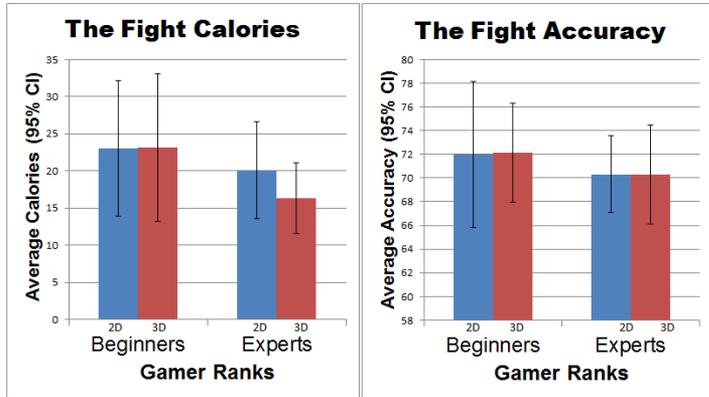
Figure 4: The Fight: Differences in calories burned and hit accuracy between the 2D and 3D groups in the two gamer categories. Expert players burned more calories in 2D display group. Beginners were more accurate than expert gamers.

Table 7: Two-way ANOVA analysis for Tumble. Significance in the number of blocks used.

Source	Time	Blocks used
DM	$F_{1,46} = 1.125, p = 0.294$	$F_{1,46} = 4.106, p < 0.05$
EXP	$F_{1,46} = 1.497, p = 0.227$	$F_{1,46} = 13.706, p < 0.01$
DG $\times$ EXP	$F_{1,46} = 0.601, p = 0.442$	$F_{1,46} = 2.682, p = 0.108$

improvement.

When we analyzed the qualitative data we found significant differences for many questions. For beginners, participants felt they performed significantly better ($Z = -2.093, p < 0.05$) in the 3D stereo group ($\bar{x} = 5.900, \sigma = 1.100$) than in the 2D display group ($\bar{x} = 4.250, \sigma = 1.752$). Beginners in the 2D display group ($\bar{x} = 2.875, \sigma = 2.167$) felt the DualShock 3 controller would be a significantly better choice ($Z = -2.438, p < 0.05$) for this game than beginners in the 3D stereo group ($\bar{x} = 1.100, \sigma = 0.316$). For expert gamers, Tumble held significantly more attention ($Z = -2.723, p < 0.05$) of the 3D stereo group ($\bar{x} = 6.800, \sigma = 0.414$) than the 2D display group ($\bar{x} = 5.823, \sigma = 1.236$). The 3D stereo group ($\bar{x} = 1.733, \sigma = 1.579$) felt significantly less distracted ($Z = -2.676, p < 0.05$) than the 2D group ($\bar{x} = 3.705, \sigma = 2.114$) and did feel the need to look around to see what was happening around them. The 3D stereo group ($\bar{x} = 6.466, \sigma = 0.990$) enjoyed the game significantly more ($Z = -1.976, p < 0.05$) than the 2D display group ($\bar{x} = 5.647, \sigma = 1.411$) as well. When asked whether the participants would play the game again, the 3D stereo group ($\bar{x} = 6.266, \sigma = 1.387$) showed significantly more interest ($Z = -2.660, p < 0.05$) than the 2D display group ($\bar{x} = 4.764, \sigma = 1.953$).

5.5 Virtua Tennis 4

In this case, the average number of wins is not normally distributed so we used ordinal logistic regression analysis, with display mode (DM) and gaming expertise (EXP) as predictors, for between subject effects and Friedman test for learning effects. No statistically significant differences were found (see Table 8 for Wald's-χ^2 test results). No statistical significance was found between any group (beginners 2D vs 3D stereo and expert 2D vs 3D stereo) in terms of score improvement. For beginners, participants in the 2D display group ($\bar{x} = 6.00, \sigma = 2.07$) thought they performed significantly better ($Z = -2.155, p < 0.05$) than the participants in the 3D
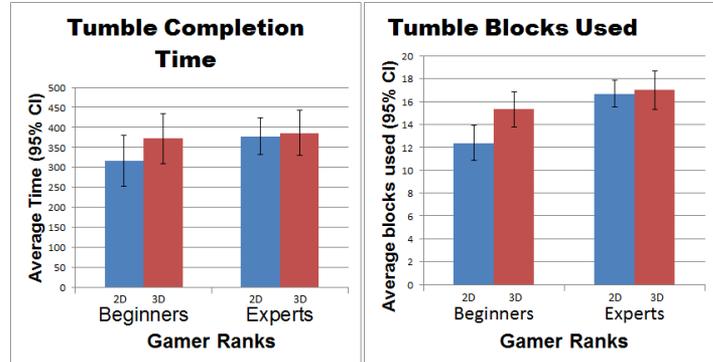

Figure 5: Tumble: Differences in number of blocks used between the 2D and 3D groups in the two gamer categories. Beginners performed significantly better in 3D stereo, while there was no significant difference (between the 2D and 3D display groups) for expert gamers.

Table 8: Ordinal logistic regression analysis for Virtua Tennis 4. No significance was found.

Source	Average number of wins
DM	$\chi^2(1) = 2.098, p = 0.147$
EXP	$\chi^2(1) = 1.792, p = 0.181$
DG $\times$ EXP	$\chi^2(1) = 2.118, p = 0.146$

stereo group($\bar{x} = 4.1, \sigma = 2.23$). For expert gamers, we did not find any statistical significance in the qualitative data.

5.6 Stereoscopic 3D Questions

Out of the 25 participants in the 2D display group, one chose to play Hustle Kings, one chose to play Pain, five chose to play The Fight, six chose Tumble, and 12 chose to play Virtua Tennis 4. The participants who played Hustle Kings and Pain thought that 3D stereo helped them. All five participants who played The Fight thought that 3D stereo helped them. Five out of six participants who played Tumble thought that 3D stereo helped them while eight out of 12 people who played tennis thought that 3D stereo helped them.

Out of the 25 participants from the 3D stereo group that played all the games in stereo, 19 participants thought that it gave them an advantage in at least one of the games, 12 participants thought that it decreased their performance in at least one of the games and three participants thought that it did not help nor decrease their performance in any way. No participants in this group thought that 3D stereo decreased their performance in Tumble.

All the participants filled out a questionnaire about their 3D stereo experience, responding to questions Q1-Q5 on a 7 point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree) (see Table 3). Participants agreed that 3D stereo improved their overall gaming experience ($\bar{x} = 6.00, \sigma = 1.44$), they would choose to play video games in 3D stereo over the 2D display ($\bar{x} = 5.76, \sigma = 1.90$), and that it enhanced the sense of engagement they felt ($\bar{x} = 5.68, \sigma = 1.89$). However, some participants did not think it was a necessity for their future game experiences ($\bar{x} = 4.72, \sigma = 2.10$). None of participants felt any significant cybersickness symptoms such as eye strain ($\bar{x} = 1.90, \sigma = 1.40$), headache ($\bar{x} = 1.38, \sigma = 0.77$), dizziness ($\bar{x} = 1.56, \sigma = 1.28$), or nausea ($\bar{x} = 1.08, \sigma = 0.34$). When we divided the data between the 2D display group and 3D stereo group, we found that the 2D display group ($\bar{x} = 6.48, \sigma = 0.871$) felt that 3D stereo significantly improved their overall ex-

perience ($Z = -2.125, p < 0.05$) compared to the 3D stereo group ($\bar{x} = 5.52, \sigma = 1.73$). When broken down based on gamer ranks, there was no significant difference ($Z = -0.786, p = 0.432$) between groups. However, for expert gamers, the 2D display group ($\bar{x} = 6.47, \sigma = 0.943$) felt 3D stereo improved their overall experience significantly more ($Z = -2.029, p < 0.05$) than the 3D stereo group ($\bar{x} = 5.4, \sigma = 1.63$).

6 DISCUSSION

From our quantitative data analysis, we can see that 3D stereo provided significant performance advantages for expert gamers for the last attempt in Hustle Kings and for beginners for Tumble in general. The other games tested showed no significant performance benefits compared with a 2D display. When participants interacted with only one object at a time with a more or less static background environment (e.g., aiming a cue ball or putting blocks on a table in 3D space) significant performance benefits occurred for 3D stereo over the 2D display. However, no significant user performance benefits were found in tasks where the scene was complex (e.g., a fight scene with player moving around) or dynamic (e.g., many incoming objects in Pain or tracking a moving ball in Tennis). Note that similar results were found in Litwiller and LaViola [9] in terms of games with complex and dynamic scenes not showing user performance benefits in 3D stereo compared to a 2D display. Additionally with some games (e.g., The Fight and Virtua Tennis 4), users need to move around with the 3D glasses which can cause the 3D glasses to flicker because of signal loss between the 3D sync signal from the TV and the glasses. This flicker could also cause distraction and affect user performance. We believe that this result supports prior findings [1, 10, 16] in the virtual reality and 3D user interface communities that 3D stereo can provide user performance benefits for isolated object position and manipulation tasks in static scenes.

Another interesting finding from the quantitative results is based on performance between beginners and expert gamers in both the 2D display and 3D stereo groups. For example, in Tumble, beginners showed significant performance benefits with 3D stereo compared to the 2D display but not expert gamers while the opposite is true for Hustle Kings. For Tumble, we believe different depth cues coupled with game experience is one of the reasons for this result. In Tumble, shadows are present under the blocks as an additional depth cue indicating their position and orientation in 3D space. Shadows have been shown to be helpful in 3D tasks [6] and are a common depth cue in video games. We theorize when presented with two depth cues (e.g., 3D stereo and shadows), expert gamers would be more used to using shadows as a tool to judge depth since it is common in their game play experience, while beginners indicated that 3D stereo served as a better depth cue than shadows. Thus, all things being equal, we believe beginners made better use of 3D stereo than the expert gamers because the expert gamers focused on shadows, a depth cue that was common to both conditions. However, more work is needed to verify this postulate.

For Hustle Kings, we believe the controller itself played a role in user performance between beginners and expert gamers. It has been shown that interaction devices significantly affect user performance [10]. Interacting with the pool cue in Hustle Kings was sensitive using the PlayStation Move Controller, making it challenging to do fine grained manipulation. We observed many cases where beginners had difficulty controlling the cue stick while the expert gamers tended to have more control and more patience with the interface. Coupled with 3D stereo, the expert gamers were able to achieve better performance in their last attempt with this game.

While examining the learning effects (e.g., score improvement across attempts in each game), we noticed that 3D stereo helped participants improve their scores in games which provided an advantage from added depth perception. The Fight, Pain, and Hustle Kings

are notable examples. What this result shows is that the 3D stereo group, especially for the expert gamers, were able to catch up to the performance levels of their 2D display counterparts. For Tumble and Virtua Tennis 4, performance improvement across runs was not significant. In Tumble, the 3D stereo group already started at a higher level than the 2D display group and did not improve much from attempt to attempt. However, the 2D display expert gamer group improved their performance with repeated attempts and caught up with 3D stereo group. It is also interesting to note that none of the beginners in the 2D display group were able to significantly improve their scores in any of the games tested.

User fatigue could also have been an important factor during these experiments. The PlayStation Move controller requires user motion while playing which can possibly cause arm fatigue, depending on the kind of motion. This may have suppressed learning effects to some extent because after each trial the user's arm fatigue possibly increases; in some cases reducing performance in the next trial. Pain is an example in which we find significant improvements from first to second trial but there were no improvements from second to third trial. Surprisingly, 3D stereo fatigue did not appear to play a role in the experiment as we did not notice any significant side effects (headache, eye strain, dizziness, and nausea) of 3D viewing in any of the participants during game play.

In the qualitative data we found some significant differences in the two user groups (2D vs 3D stereo). We noticed that 3D stereo is perceived to be more enjoyable and immersive than 2D viewing only for the games which provide an overall advantage using 3D stereo (e.g., Tumble), but no significant differences were found in the other games we tested. In general, almost all participants were not familiar with the games we used for our study, so most of them were excited to play using the PlayStation Move controller with 3D stereo acting as a secondary factor in their game play experience. This may be the reason that, qualitatively, they perceived similar game play experiences, no matter what group they were in. Most people liked the 3D stereo game play experience but some users felt they were so accustomed to playing on a 2D display that the 3D effect distracted them.

7 CONCLUSION AND FUTURE WORK

We have presented a study exploring the effects of 3D stereo on user performance in 3D motion controlled video games. For the first time, we observed a positive impact of 3D stereo on gaming performance, which seems to be related to 3D interaction. However, our results reveal that performance in 3D interaction gaming does not automatically benefit from 3D stereoscopic vision. Interestingly, 3D stereo can specifically provide a significant performance advantage over 2D vision in rather isolated tasks, when users are manipulating one object at a time and when a scene is more or less static. In simple scenes impact of 3D stereo on performance is much greater than in complex games where many dynamic factors (camera perspective, enemy behavior, and other animated elements) around the interacting object influence the course of the game. A third important finding is that game expertise has the potential to nullify this effect, as observed in the Tumble game. A possible reason is that gamers may have learned to rely on other cues than binocular disparity (e.g., on shadows and lighting). Hence, beginners are more open to using new visual cues and thus benefit more from using 3D stereoscopic vision.

Additionally, our qualitative data indicates that 3D stereo is perceived to be more enjoyable and immersive than 2D viewing only for the games which provide an advantage in 3D stereo. This outcome contradicts previous findings, which reported preference for 3D stereo although no advantages in performance were found [9, 13]. These results lead to our conclusion that games need to be particularly designed to allow a benefit in performance from stereo-

scopic vision. As part of such a design, using a 3D motion controller as a game controller can have a positive impact. A starting point in game design, based on the game Tumble, could be to isolate depth precision tasks. The stereoscopic effect could be used alongside other depth cues to create game conflicts and for balancing tasks, an approach also described in [15]. Based on the above findings and on our observations, we recommend game designers to

- Utilize relatively simple scenes or static environments where interaction is focused on isolated tasks to provide user performance benefits with 3D stereo. This approach can help to avoid user distraction.

- Try to emphasize the stereo effect, showing how to use it in gameplay, especially expert users who may not take it into account.

- Provide a way to control the sensitivity of the controller to make it a more enjoyable user experience.

- Avoid requiring a lot of user motion in front of the display to avoid any sync signal loss issues with active 3D stereo glasses and to reduce geometric errors when leaving the sweet spot for 3D effect. Alternatively provide tracked stereo or RF signal based sync.

In the future, we plan to continue to explore how 3D stereo affects user performance in video games. Based on this work, we want to specifically explore 3D stereo's relative strength as a depth cue in terms of the other 3D depths cues that are available for gamers with varying experience levels. One depth cue that is especially interesting is motion parallax, which requires head tracking. We plan to explore how to best fit the 3D tasks that support improved user performance into existing game genres.

8 ACKNOWLEDGMENTS

9 REFERENCES

[1] K. W. Arthur, K. S. Booth, and C. Ware. Evaluating 3d task performance for fish tank virtual worlds. *ACM Transactions on Information Systems*, 11,3:239–265, 1993.

[2] N. Bianchi-Berthouze, W. W. Kim, and D. Patel. Does body movement engage you more in digital game play? and why? In *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*, volume 4738, pages 102–113. 2007.

[3] M. Fujimoto and Y. Ishibashi. The effect of stereoscopic viewing of a virtual space on a networked game using haptic media. In *Proceedings of the ACM SIGCHI international Conference on Advances in Computer Entertainment*, pages 317–320, New York, NY, 2005. ACM.

[4] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

[5] P. A. Howarth. Potential hazards of viewing 3-d stereoscopic television, cinema and computer games: a review. *Ophthalmic & Physiological Optics : The Journal of the British College of Ophthalmic Opticians (Optometrists)*, 31(2):111–122, 2011.

[6] G. S. Hubona, P. N. Wheeler, G. W. Shirah, and M. Brandt. The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization. *ACM Transactions on Computer-Human Interaction*, 6,3:214–242, 1999.

[7] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66,9:641–661, 2008.

[8] S. Lindley, J. Le Couteur, and N. L. Berthouze. Stirring up experience through movement in game play: effects on engagement and social behaviour. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 511–514, New York, NY, USA, 2008. ACM.

[9] T. Litwiller and J. J. LaViola Jr. Evaluating the benefits of 3d stereo in modern video games. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, pages 2345–2354, New York, NY, USA, 2011. ACM.

[10] R. P. McMahan, D. Gorton, J. Gresock, W. McConnell, and D. A. Bowman. Separating the effects of level of immersion and 3d interaction techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 108–111, New York, NY, 2006. ACM.

[11] J. Merritt, R. Cole, and C. Ikehara. A rapid-sequential-positioning task for evaluating motion parallax and stereoscopic 3d cues in teleoperator displays. In *Systems, Man, and Cybernetics, 1991. 'Decision Aiding for Complex Systems, Conference Proceedings., 1991 IEEE International Conference on*, number 2, pages 1041 –1046, oct 1991.

[12] Y. Munz, K. Moorthy, A. Dosis, J. D. Hernandez, S. Bann, F. Bello, S. Martin, A. Darzi, and T. Rockall. The benefits of stereoscopic vision in robotic-assisted performance on bench models. *Surgical Endoscopy*, pages 611–616, 2004.

[13] R. J. Rajae-Joordens. Measuring experiences in gaming and tv applications. In *Probing Experience*, volume 8 of *Philips Re*, pages 77–90. Springer Netherlands, 2008.

[14] J. Schild, J. J. LaViola Jr., and M. Masuch. Understanding stereoscopic game experience. In *To appear in the Proceedings of the 2012 annual conference on Human factors in computing systems*, 2012.

[15] J. Schild and M. Masuch. Fundamentals of stereoscopic 3d game design. In *ICEC, International Federation For Information Processing*, pages 155–160, 2011.

[16] R. J. Teather and W. Stuerzlinger. Guidelines for 3d positioning techniques. In *Proceedings of the 2007 Conference on Future Play*, pages 61–68, New York, NY, 2007. ACM.

[17] M. Terlecki and N. Newcombe. How important is the digital divide? the relation of computer and videogame usage gender differences in mental rotation ability. *Sex Roles*, 53:433–441, 2005.

[18] C. Ware and P. Mitchell. Reevaluating stereo and motion cues for visualizing graphs in three dimensions. In *Proceedings of Applied Perception in Graphics and Visualization*, pages 51–58, New York, NY, 2005. ACM.

[19] M. Wexler and J. J. van Boxtel. Depth perception by the active observer. *Trends in Cognitive Sciences*, 9(9):431–438, 2005.

[20] Y. Yamauchi and K. Shinohara. Effect of binocular stereopsis on surgical manipulation performance and fatigue when using a stereoscopic endoscope. In *Medicine Meets Virtual Reality 13: The Magical Next Becomes the Medical Now, IOS Press*, volume 111, pages 611–614. feb 2005.