Altering Gameplay Behavior using Stereoscopic 3D Vision-Based Video Game Design

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

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

Maic Masuch
University of Duisburg-Essen
Entertainment Computing
Forsthausweg 2, 47057
Duisburg, Germany
maic.masuch@uni-due.de

ABSTRACT

We explore the potential of stereoscopic 3D (S3D) vision in offering distinct gameplay using an S3D-specific game called Deepress3D. Our game utilizes established S3D design principles for optimizing GUI design, visual comfort and game mechanics which rely on depth perception in time-pressured spatial conflicts. The game collects detailed S3D player metrics and allows players to choose between different, evenly matched strategies. We conducted a between subjects study comparing S3D and monoscopic versions of Deepress3D that examined player behavior and performance and measured user-reported data on presence, simulator sickness, and game experience. Confirming previous results, stereo users reported higher spatial presence. More importantly, for the first time, our game metrics indicate that S3D vision can measurably change player behavior depending on actual game content and level design, without necessarily affecting performance or emotional experience. These findings indicate the potential for optimizing applications for stereo users distinguishing them as a distinct group in HCI research.

Author Keywords

Stereoscopic 3D; user experience; gameplay metrics; player behavior; game design; spatial presence; simulator sickness

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Experiencing media content with Stereoscopic 3D vision (S3D) is often perceived in two ways. The S3D effect can initially draw attention and fascinate viewers but many doubt if there is an actual benefit to experiencing movies or games with S3D. One important aspect of the user experience with S3D vision is how it affects user interaction. For example, do players score higher playing a game with S3D vision than with monoscopic vision? Previous work has found only sporadic performance benefit in commercial games, e.g., with isolated 3D interaction [7]. However, besides the question of whether users are "better" with S3D, the question of whether S3D changes *what* users do has not been considered before. Can S3D influence how users behave when using a system?

We explore this question in the domain of digital games and try to understand how to create games specifically for use with S3D vision in a meaningful way. Recent work supports this approach by proposing S3D-specific guidelines for game design (e.g., [9, 17]), tools for affective S3D GUI design [15] and prototypes that demonstrate depth illusions that are easier to resolve in stereo [13, 23].

This paper presents the first holistic approach to S3D game design by incorporating previous findings and exploring if change in behavior and meaning from S3D can be created intentionally using an S3D-specific game. Deepress3D is (a) created around S3D guidelines, (b) is representative as a fun and usable game, (c) supports multiple strategies in gameplay and (d) provides detailed in-game metrics of player behavior. We conducted a between subjects study and analyzed how players experience and play (i.e., perform and act in) the game differently in S3D and in monoscopic viewing modes.

RELATED WORK

Since the success of the movie Avatar in 2009 and the subsequent hype of S3D technology in the media, a significant amount of work has looked at possible benefits of S3D vision in games, first focusing on existing commercial games.

To better understand overall affective player experience when using S3D vision, Schild et al. previously examined three existing video games by comparing both S3D and monoscopic modes using self reporting questionnaires and a psychophysiological device [16]. The results showed that S3D increased experienced immersion, spatial presence, and simulator sickness. These effects differed strongly between the different games tested for both genders and more effects appeared in games with large amounts of depth motion than in a side-scrolling game. In another work, Takatalo et al. [20] evaluated experiencing a racing game with different parallax settings and monoscopic viewing. Their results showed higher presence in one of the parallax conditions and did not find any effect in emotional factors such as fun or enjoyment when compared to the monoscopic viewing condition.

Regarding player performance benefits from S3D vision, two studies examined several existing games (e.g., racing, first and third person shooters, and sports) [12, 8]. Although participants preferred playing the games in S3D, doing so did not provide any discernible advantage in performance compared to using monoscopic vision. Kulshreshth et al. examined user performance with existing games that included motion control (e.g., Playstation Move) [7]. They found two cases, depending on player expertise and when interacting with isolated tasks, where S3D helped to improve gameplay.

In summary, S3D with existing games only sporadically causes measurable impact on user experience (perhaps more with depth-motion views) or player performance (e.g., in isolated depth-based tasks). Inline with these findings, some researchers argue, that those games were not specifically designed with S3D as their focal point and present design guidelines for creating S3D-specific games [9, 17], e.g., proposing depth-based game mechanics, depth-promoting camera effects, and S3D-specific graphical user interfaces (GUIs).

Very recent work has consequently explored S3D-specific game content. Schild et al. experimented with live-feed S3D video and mixed-reality depth mechanics in gameplay but provided no evaluation besides reporting an initial fascination among users [18]. Zerebecki et al. created a top-down shooter offering an obstacle course of abstract blocks [23]. Most monoscopic depth cues were removed, further minimizing perspective effects using a very narrow depth of field. A within-subjects study showed that the players who started their play-through in monoscopic mode did not complete the game on their first run. In a similar approach, Rivett and Holliman reported on a top-down shooter with similarly reduced monoscopic depth cues, where the player has to control a space ship through spatially aligned hoops [13]. A within-subjects study showed significantly higher scores in the S3D mode. Hence, a performance impact is indeed possible using strong depth illusions resolvable through S3D vision as game mechanics. But the actual reason for such performance differences was not analyzed in game metrics which would have allowed transfer to other game projects. It is further questionable whether these effects with rather abstract graphics hold valid if reduced to a subtle level in a visually more vivid game environment as commonly used and expected in games.

The latter two studies did also not show any differences in user engagement between the two vision modes with S3D-specific content. In attempting affective impact, Schild et al. found that S3D-specific GUI design can influence emotional user experience and should be integrated with the spatial and fictional aspects of a game [15].

As of yet, there has been no holistic attempt to create and evaluate an S3D-specific game which integrates previous guidelines and combines reported findings for effective S3D applications (e.g., depth motion, depth delusions, GUI design). Creating a custom game enables us to conduct detailed logging of game metrics. Hence, in addition to measuring user scores and affective experience, we can for the first time examine impact on player decisions, by offering a game design which supports multiple balanced gameplay strategies.

DESIGNING A STEREOSCOPIC 3D GAME

Stereoscopic 3D Game Design Principles

To design a video game that makes explicit use of S3D vision, we made use of several underlying principles that have been developed in the literature. These principles provide a set of guidelines that provide suggestions for how S3D can be used in a game in order to create a noticeable benefit in both user performance and experience when compared with a monoscopic equivalent. These principles include:

- Provide a comfortable experience through a choice of perceptual settings, content and technology in order to avoid visual discomfort [5, 17, 19].

- Using depth perception in gameplay contains S3D-specific opportunities for game design [17] which can be actually applied and are potentially fun to play [13, 18, 23].

- Content with a lot of depth animation has the largest potential to specially influence affective user experience (e.g., immersion and presence) through S3D but also causes more simulator sickness [9, 16].

- Static scenes and focused tasks have the largest potential to increase user performance from S3D and seem not to cause simulator sickness [7].

- Graphical user interfaces need to be adjusted, (e.g., by displaying information implicitly through objects in the game world or by spatially integrating information) [15].

Deepress3D

Based on these principles we designed Deepress3D (see Figure 1), a custom-created stereoscopic game. The player has to fly and fight with a space ship through a spatial obstacle course while being time-pressured from auto-scrolling. In order to avoid obstacles the level design offers multiple strategies which supports variety in player behavior. The game features an emotional story setting and is playable from two perspectives: side-view and behind-view. The game was implemented using the professional Havok Vision game engine.

The story tells the loss of an important person. The ship resembles the player's future self, getting through the resulting sentimental chaos as quickly as possible. The game consists of three levels with different design, backgrounds and music tracks, subsequently themed with sorrow, aggression, and relaxation. Additional emotional statements are given through 3D text included with the game scenery (cf. Figure 1d).

The gameplay offers a conflict where the player has to manage a limited energy resource to quickly choose between navigating around obstacles or fighting enemies. To surpass obstacles, the player can move the ship in two dimensions across the screen as well as in depth by switching between three discrete depth layers (i.e., foreground, center, and background). Changing depth layers is often necessary as the current layer may be blocked by obstacles (i.e., static hills, moving boxes, and aggressive enemy space ships). The player ship can shoot and when hit by a bullet an enemy ship explodes.

To balance player choices between moving in 2D, in 3D, or fighting enemies, all such player activity and damage through

| (a) | (b) | (c) | (d) |

Figure 1: In Deepress3D, the player has to share energy between gameplay choices, e.g., changing depth, flying out of scope or shooting enemies. (a+b) Section 5 in level 1 features a labyrinth of deadly static and moving boxes where stereo players used the fast speed more often than mono players. (c) In level 2 of Deepress3D, the player is primarily confronted with shooting enemy ships in an aggressive setting. (d) Level 3 creates a relaxed setting, with positive story moments. Deepress3D supports comparable gameplay for two perspectives: (a+c) side-scrolling and (b+d) behind-the-ship perspectives (anaglyph stereo images).

collisions draws from a common energy reservoir, resembled as a green bar below the ship. If the energy is empty, the ship explodes and re-spawns at the beginning of the current section. Energy can be increased by collecting rings distributed across the levels. Energy decreases when colliding with obstacles, firing the cannon, changing depth layers, or while leaving a central area called the personal scope.

The personal scope is a spatial rectangular frame which assures that players use similar space across all three depth layers and both perspectives. The player may temporarily leave the scope to overcome obstacles. However, staying out of scope will drown energy until the persona ship is destroyed. The border of the scope is shown graphically when close to the border or during switching depth layers. The player can adjust speed of game progression by moving to the right border or deep border in ego perspective (faster auto-scrolling) or to the left/near border of the scope (slower auto-scrolling).

For analyzing the impact from depth animation ([16]), two different perspectives were developed: a side-scrolling viewpoint with only few depth animations (see Figure 1a) and a depth-scrolling viewpoint behind the ship (see Figure 1b).

Stereoscopic 3D Specific Features

In order to provide the most compelling S3D experience in Deepress3D, we make use of three different methods based on prior work to enhance S3D in the game. These include using a S3D-specific GUI, using depth cue-deluding, and providing an optimal parallax budget.

S3D Specific GUI. Using the S3D Game GUI design space [15], we used an intra-diegetic item attached to an object of the game world: the amount of energy in the reservoir is displayed as a colored bar attached to the bottom of the player's ship. As part of the player ship the energy bar is animated and can be occluded through other world objects. It also saves the player from switching between the depth layer of the ship and that of the screen. This concept further fits well as the energy bar directly refers to the ship and its status. A timer and progress indicator were also implemented according to the design space. In contrast to the energy bar, we used abstract numbers attached to the view. This fits our concept as the autoscrolling perspective directly represents the progression of the game. To minimize change between depth

layers, we rendered the digital numbers of time and progress always at the depth layer of the player ship. But to express attachment to the view, they are displayed all the time, never occluded by game objects closer to the camera, and always with the same 2D size and position in the view. Only the parallax of the digits change with that of the player ship to reduced depth jumps when looking from the ship to the digits and back. This works in both side-view and behind-view.

Depth Cue-Deluding. The graphical design of the landscapes and obstacles is aimed to sublty reduce impact of monocular depth cues. The textures of the hills and boxes show reduced contrast in pattern frequency, a known depth cue. The enemy models are similar to the player ship as they are not textured and look rather abstract. The lighting does not feature any projected shadows or dynamic light sources. We added subtle animation effects to the hills. Three of the four hills rotate very slightly around the z-axis in varying speed and direction. The hills also grow and shrink very slowly in a looped animation. These effects have no impact on gameplay but influence depth perception: in the side-view the rotation confuses monoscopic cues of motion parallax. The grow/shrink-animations distort relative size cues. In contrast to [23], Deepress3D does not use orthographic projection which would have reduced the overall perception of depth too much in our opinion. In summary, the graphics of Deepress3D use subtle delusions but no real depth illusions. These effects occur in both monoscopic and stereoscopic vision.

Optimal Parallax Budget. One significant problem with existing stereoscopic games is that the depth budget has to stay inside a certain comfortable range. Typically the player can adjust this range towards less or more depth in, for example, both Nvidia 3D Vision and Nintendo 3DS using a slider. The Nvidia solution offers advanced players the ability to further adjust convergence and driver settings for each game. Games which involve different scenes and dynamic perspectives require the player to re-adjust the stereo settings during gameplay to always gain the optimum between depth and comfort.

Deepress3D addresses these problems using two approaches. First, the concept of the game supports keeping the camera view from a constant perspective minimizing changes of the depth budget. Secondly, the game automatically creates an

optimal setting depending of the interpupillary distance (IPD) of the user and the typical screen distance during gameplay. Our calculations of the comfortable viewing setup follows the approach by Shibata et al. [19] which according to Banks et al. [1] represents the current best guess of the zone of comfort. For most active game objects within the range of the player ship, the parallaxes (near/far) for average settings for IPD (.063 m) and viewing distance (.65 m) are set to 0 m/.012 m for the side-view and .07 m/.017 m for the behind-view. None of the visible objects exceed CVR limits (-.0023 m/.0023 m) as proposed by [19].

USER EXPERIENCE AND PERFORMANCE STUDY

Design and Hypotheses

The primary interest of our study is the impact of stereoscopy on how the game is played and experienced. Thus, we use a between-subjects design, grouping the participants into a control group (monoscopic 3D vision) and an experimental group (stereoscopic 3D vision). To gain additional information on the impact of the game perspective, each participant played the game first in the side-scrolling view and then in the behind-the-ship view. This mixed design provides clear results on comparisons of stereo versus mono primarily for the side-scrolling perspective.

The fixed order of the perspectives supports a judgment on how the behind-view with a large amount of depth animation changes effects for each viewing condition. The fixed order was necessary as the behind-view lends more overview on level architecture than the side-view. In pretests, using behind-view first gave users a performance advantage than if they used side-view first. Using side-view after behind-view was further experienced as a downgrade (large overview first, then reduction to small cutout).

Based on previously reported findings (cf. Related Work) and our adjustments to create an S3D-specific game, we formulated the following five hypotheses for the study:

H1 Game experience will be more immersive in the stereo vision group [16].

H2 Presence will be improved by stereoscopy [16, 20].

H3 Simulator sickness will not be increased by stereo over mono vision (due to the fixed perspective, UI adjustments and individual settings).

H4 Game performance will be improved by stereoscopy (supported by the depth deluding).

H5 Stereo player will behave differently from mono players (as provided by multiple gameplay strategies).

Participants and Apparatus

The study involved 44 subjects of which two data sets were invalid and were removed together with two other data sets to keep the two vision groups homogeneous regarding demographic data (job, education, sex, age and prior expertise). For the remaining 40 participants (23 male, 17 female), their ages ranged from 19 to 30 years with a mean of 23.1 years and a median age of 24 years. The group included 32 students and

8 employees. The participants had all ranges of game expertise: nearly half of the subjects either never play games (eight subjects) or for 1 to 5 hours per week (9 subjects). Nine subjects play for 6 to 10 hours; ten play for 11-20 hours and four reported playing 21 to 30 hours per week. Only one subject had never experienced stereo 3D before. Eight subjects reportedly used S3D 1 to 3 times. The majority had used S3D 4 to 9 times (18 subjects) or more than 10 times (13 subjects).

Deepress3D was installed on a testing computer positioned behind a separation wall. The lighting setting was day-light indoor shadow without artificial lights interfering with the shutter glasses. The game ran on a Windows 7 desktop PC, with a GeForce GTX 275 graphics card and an Acer GD245HQ 24" display at 120 Hz with 1080p resolution. Nvidia 3D Vision was used to generate side-by-side full-resolution stereo. Since there was an impact on the subjective feel and comfort by wearing the 3D Vision glasses, both groups had to wear them during the sessions. In monoscopic display, the open shutters of the glasses provide an image slightly darker than without glasses but minimally brighter than the stereo version and similar in color tone (altered towards yellow). During play, the experimenter sat apart from the participants without being able to see the screen and was not involved but possibly available to answer questions.

Metrics

We used the GEQ for game experience [2], the MEC-SPQ for spatial presence [21], and the SSQ for simulator sickness [6]. In addition, game experience reaction towards fun was assessed through nine custom 6-level Likert-items ranging from -3 to +3 in a forced-choice design without a neutral option.

General usability of the game was measured using the SUS [3] as a variation where terms relating to a "system" were adopted as relating to a "game" and likewise the term "use" was exchanged with "play". This approach has been applied in other studies [4, 10] and proven valid as summarized by [14]. The SUS consists of ten items combined in a total usability score. The Overall SUS score can range from 0 to 100 and is directly comparable with other results [14].

Game performance was measured in detail through in-game logging. We implemented logging of performance related statistics (overall play-time duration, number of deaths, kill ratio) as well as a protocol of single events (shooting, changing depth layers, entering zones within or outside the personality scope, flying faster or slower, dying, respawning, collecting items, colliding with enemies, enemy bullets, boxes). The game logging collected summarized statistics for the overall game and for the separate levels.

Procedure

Each session started with a welcoming statement and a short introduction to the procedure. Participants were informed about possible health-related issues concerning epilepsy and game addiction, as well as about possible symptoms of visual fatigue to which they had to agree as well. Without knowing, the participants were then randomly selected for one of the two groups by their own choice for one of two types of chocolate bar. Besides the candy, the student participants received

"study points" they needed to earn in order to conduct their own experiments in their courses. The study was approved by the departments' ethics board. We measured the stereo group participants' interpupillar distances and the distance between user and display in a comfortable seating position to calculate individual stereo settings. The Nvidia medical test image was used to test stereo abilities among participants of the experiential group. All participants passed the test.

The game introduction print-out was handed over and the participant could read through the document. As hero of the game, the player navigates a glider through the three game levels as fast as possible. The description introduced all elements of the game, possible player procedures and consequences of failure. It also announced a playable tutorial. After this introduction, the participants filled out the first questionnaire about demographic data and SSQ (T0=Baseline). During this time, the experimenter prepared the configuration data of the game resulting in individual stereo settings or monoscopic rendering and the logging. The participants geared up with the shutter glasses and took the gamepad. The experimenter then started the tutorial of Deepress3D. The tutorial was a level of the game played in side-view which introduced each game element separately. While receiving explanation and support by the experimenter, the player could control the ship and get used to the game.

After the tutorial, the experimenter started the first game session and moved away from the game computer. The participants played through all three levels of the game with side-view. This session was followed by the first questionnaire block (T1). During the questionnaire, the experimenter prepared the behind-view and logging of the second session. The second session offered the participants the tutorial again to get used to the slightly changed controls. The participants played through the full game in behind-view, again followed by a questionnaire (T2). At the end of the trial the experimenter thanked the participant and handed out the study-points. The average duration for a full trial was about 90 minutes.

RESULTS

We tested the resulting data for normal distribution using Shapiro-Wilk-tests and analyzed effect sizes using Cohen's d. For analysis of variance between groups, we applied a Two-Samples independent T-Test for normally distributed data and a Mann-Whitney U-Test elsewhere. Within each group we analyzed differences between data points (T0, T1 and T2) using a dependent samples T-test for normally distributed data and Wilcoxon-Signed-Rank test elsewhere. The data of the fun-related questions was analyzed on a Likert scale through descriptive analysis based on medians.

Questionnaire Results

Game Experience and Fun

Game experience was measured through the seven GEQ dimensions Immersion, Flow, Competence, Tension, Challenge, Positive Affect, and Negative Affect, as extracted from the 33 items following [11]. On a 1-to-5-range the GEQ scales Immersion, Flow, Competence, Challenge, and Positive Affect scored mediocre (means between 2.33 and 3.67)
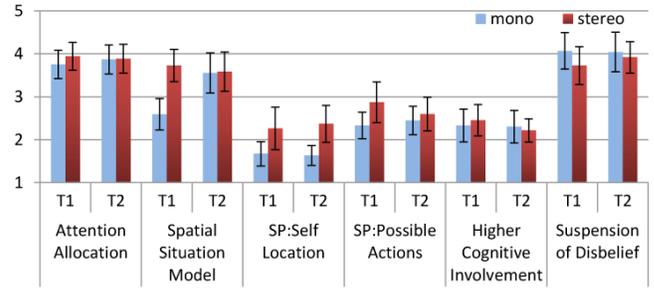

Figure 2: The means of the six spatial presence dimensions for the side-scrolling gameplay (T1) and the behind-view (T2). Error bars show 95% confidence interval (CI).

for both vision groups and at both data points. Tension and Negative Affect were rather low (means below 2.33). Overall, none of the described differences between the two viewing modes were statistically significant as where no differences between the two data points T1 and T2.

In the nine fun-related items ranging from -3 to +3, both groups reported the game a fun (MD=2) and comprehensible (MD=2) experience without stating large emotions towards either positive or negative among the majority of users. They thought of the game as rather simple (MD=1) at T1 but at rather complex at T2 (MD=-1). Differences between the groups were very slight and not statistically relevant. At T1, the mono group had a higher appealing median (MD=2) than the stereo group (MD=1). At T2, the stereo group had a higher appealing median (MD=2) than the mono group (MD=1). While the mono group reported the game at T1 as rather difficult (MD=-1), the stereo group described it as rather easy (MD=1).

Spatial Presence

Overall, presence was significantly increased in the stereo group over the mono group especially with the side-scrolling perspective but also in the behind-view. After T1, the mean values for all presence dimensions except for Suspension of Disbelief were higher in the stereo group than in the mono group (see Figure 2), with significant effects in Spatial Situation Model ($t_{40} = -4.517$, $p < .001$, Cohen's $d = 1.43$), SP:Self Location ($t_{31.371} = -2.180$, $p < .05$, $d = .68$) and a medium-sized trend in SP:Possible Actions ($t_{33.67} = -1.996$, $p = .54$, $d = .62$). After T2, the means of Attention Allocation, Spatial Situation Model, SP:Self Location, SP:Possible Actions were higher in the stereo group than in the mono group. Only for SP:Self Location this effect was significant and large ($Z = -2.69$, $p < .01$, $d = .99$). The means of Higher Cognitive Involvement and Suspension of Disbelief were higher in the mono group at T2.

When comparing T1 vs. T2 for the mono group, Spatial Situation Model was significantly higher with the behind-view (T2) than with the side-view ($t_{18} = -4.320$, $p < .001$, $d = 1.34$). A minimally higher mean of Attention Allocation is reported at T2 than at T1 ($Z = -1.964$, $p < .05$, $d = .31$). In the stereo group no differences between T1 and T2 were significant.
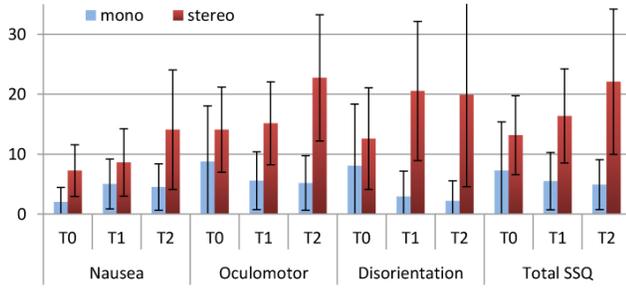
Figure 3: The means of the four SSQ at the three data points: before having played anything (T0=baseline), after the side-scrolling gameplay (T1), and after the behind-view (T2). The groups judged simulator sickness different already at baseline (error bars show 95% CI).

Usability
The means of the SUS range between 73.8 and 76.6, which is above average and scores on a percentile rank higher than between 67% and 77% of other SUS-results [14]. The mean scores of the SUS were slightly higher for mono than stereo at T1 and slightly lower for mono than stereo at T2 but these effects were not measurable ($d < .2$). Respectively we found no significant differences between the vision groups at both data points, the game was usable significantly above average in all conditions and for both groups.

Simulator Sickness
The SSQ means of the stereo group (see Figure 3) were higher than the means of the mono group in all subscales and for all three measures. As these differences between mono and stereo were already significant for the baseline (T0) in Nausea ($Z = -2.331$, $p < .05$, $d = .59$) and Total ($Z = -2.157$, $p < .05$, $d = .38$), the impact of the viewing mode has to be analyzed in a comparison of how the means develop within mono and stereo groups respectively. A non-parametric Friedman-Test reports significant differences between baseline, T1 and T2 only in the stereo group for the Total SSQ dimension ($\chi^2 = 7.529$, $df = 2$, $p < .05$). Pairwise comparisons of significance showed no significant effects between Baseline and T1 but only from T1 to T2 in the stereo group, where three of the four SSQ scores increased significantly from the side-scrolling perspective to the behind-view perspective: Nausea ($Z = -1.980, p < .05, d = .43$), Oculomotor ($Z = -2.509, p < .05, d = .63$), and Total SSQ ($Z = -2.006, p < .05, d = .41$) but not Disorientation.

Performance-Based Results
We logged the duration of play and counted number and reason of deaths for the separate levels. We calculated a one-way repeated measures ANOVA to compare the two vision groups (mono/stereo) across the three levels of the two data points/perspectives (T1=side, T2=behind). The results show that the levels differed significantly concerning completion time ($F(1.803, 48.693) = 8.166, p < .01, \eta^2 = .232$) and number of deaths ($F(2, 54) = 19.128, p < .001, \eta^2 = .415$). Also the two data points T1 and T2 had an impact on completion time ($F(1, 27) = 17.411, p < .001, \eta^2 = .392$).
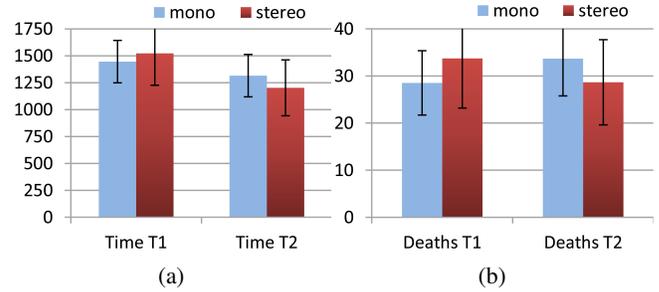

Figure 4: The mean total (a) play durations (time in seconds) and (b) number of deaths did not differ significantly between the stereo and the mono groups (error bars show 95% CI).

Regarded independently of these differences, vision had no impact in both completion time ($F(1, 27) = .118, p = .73$) and number of deaths ($F(1, 27) = .209, p = .65$). Only in interaction with level and data point factors, the vision group had a significant impact on both completion time ($F(2, 54) = 3.608, p < .05, \eta^2 = .118$) and number of deaths ($F(2, 54) = 3.684, p < .03, \eta^2 = .120$).

As depicted in Figure 4a, the results of play durations for the side-view indicate similar means in the stereo group as in the mono group (m_{mono}=1446 sec, m_{stereo}=1523 sec, $d < .2$) and for the behind-view slightly lower means in the stereo group than in the mono group (m_{mono}= 1316 sec, m_{stereo}=1203 sec, $d = .23$). None of these differences were significant. Similarly, the number of deaths (see Figure 4b) was slightly higher in the stereo group at T1 (m_{mono}=28.53, m_{stereo}=33.68, $d = .29$) and slightly lower in the stereo group at T2 (m_{mono}=33.74, m_{stereo}=28.67, $d = .28$). Again, these differences were not statistically significant. No significant differences were found concerning completion times and death rate for the particular levels.

Game Metrics and Behavior Results Between Groups
The logs of the game metrics allowed further analysis of the playing behavior logged in certain events: firing a shot, killing an enemy, collecting an item, colliding with an obstacle, increasing or decreasing speed, going to the edge of the scope or beyond, and changing a depth layer. We analyzed these results using a one-way repeated measures ANOVA. As with the performance, there were only slight effects in interaction with the within-subject factors.

Vision group (mono/stereo) in interaction with data point (perspective) had significant impact on how much time proportionally was used at slow speed ($F(1, 27) = 5.563, p < .05, \eta^2 = .171$). In interaction with Level and data point (perspective), the stereo group showed differences in the number of layer changes ($F(1.632, 44.952) = 5.275, p < .05, \eta^2 = .163$) and how much time was spent in the foreground or right layer ($F(2, 54) = 3.659, p < .05, \eta^2 = .120$), at normal speed ($F(2, 54) = 4.169, p < .05, \eta^2 = .134$), and within the scope ($F(2, 54) = 3.789, p < .05, \eta^2 = .123$).

Side-scrolling Perspective
The behavior of the groups differed in several aspects in the side perspective and for the three separate levels. As Figure 5
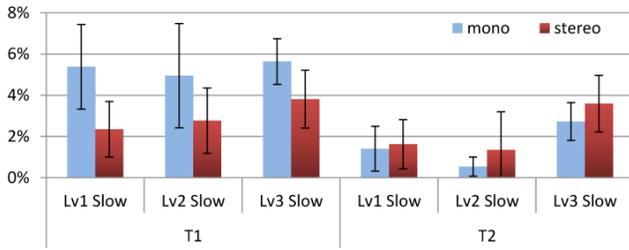
Figure 5: The duration per level spent at slow speed differed significantly between the groups in level 1 and 3 of T1.
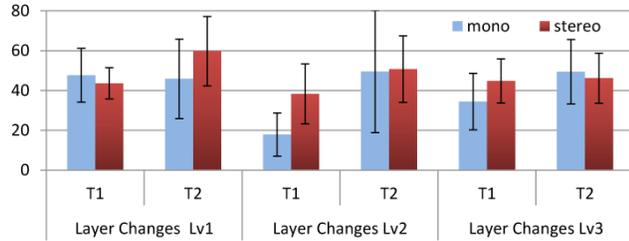

Figure 6: The number of layer changes differed between levels, data points and vision groups. Error bars show 95% CI.

illustrates, stereo players used significantly less time to slow down than mono players ($Z = -2.614, p < .01, d = .47$). This effect applied strongly in level 1 ($Z = -2.804, p < .01, d = .93$), slightly in level 3 ($Z = -2.709, p < .01, d = .44$), but not in level 2. This effect also applied when calculating the relative time spent at slow speed per total play duration.

Additionally with the side-view, our stereo players in general changed depth layers more often than mono players ($d = .59$, see Figure 6) but this effect was large and statistically significant only in level 2 ($Z = -2.363, p < .05, d = .87$), not significantly but with a medium effect in level 3 ($d = .54$).

Looking at how the players distributed their play time across the three levels (left part of Figure 7), mono players spent 4% more of their playtime in level 1 than stereo players ($t_{34} = 2.319, p < .05, d = .78$) while stereo players spent 5% more of their time in level 2 than mono players ($t_{29.672} = 2.303, p < .05, d = .76$). There was no difference in level 3. Similarly, as depicted in the right part of Figure 7, the proportional amount of deaths was distributed more but insignificantly towards level 1 for our mono group ($d = .61$) and likely more towards level 2 for stereo players ($t_{34} = 1.905, p = .65, d = .64$).

As of usage of the depth layers (see Figure 8), stereo players spent a longer portion of their playtime in the central depth layer than mono players did in level 1 ($t_{34} = 2.370, p < .05, d = .79$). Stereo players also played proportionally longer time on the background depth layer at level 3 ($Z = -2.202, p < .05, d = .77$).

We found more differences between the stereo and the mono groups in means which had a noticeable effect size but which were not statistically significant: The stereo group fired less shots (see Figure 9) with the side-scrolling game ($d = .72$). Our stereo group also killed less enemies (see Figure 9) ($d =$
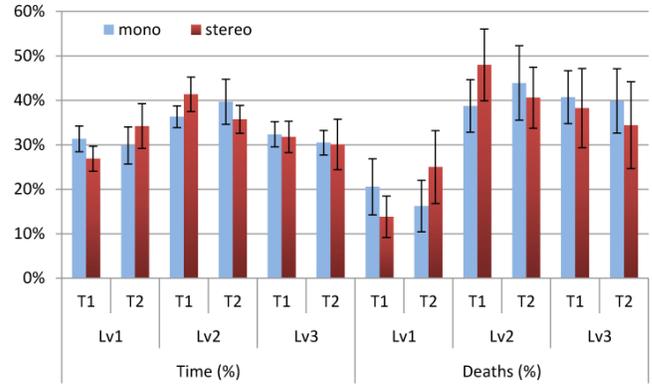

Figure 7: The time and number of deaths in percentage distributed across the three levels, data points (T1 and T2) and vision groups (mono and stereo). For example, at T1 (normal perspective) the stereo group spent 27% of their time in level 1, 41% of their time in level 2, and 32% of their time in level 3 (100% in total). Error bars show 95% CI.
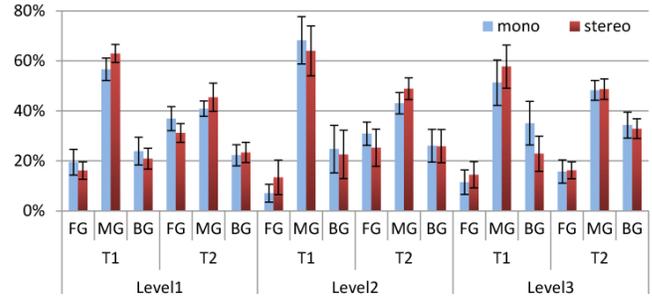

Figure 8: The players used the three depth layers FG (foreground at T1/right layer at T2), MG (middle or center in both perspectives), and BG (background at T1/left layer at T2) differently across data points, levels and vision group. The data shows percentage of each level's playtime per data point and vision group. Error bars show 95% CI.

.57). Effectively, the killing rate (kills per shot) was similar between both groups at T1 (m_{mono}=.19, sd=.08, m_{stereo}=.20, sd=.11) and at T2 (m_{mono}=.12, sd=.05, m_{stereo}=.13, sd=.06).

Behind-view Gameplay
The behind-the-ship perspective gameplay showed only few significant differences in gameplay between the vision groups. Overall, stereo players died proportionally more frequently from changing layers and hence losing too much energy than mono players ($Z = -2.048, p < .05, d = .69$). In absolute numbers they died much less from colliding with enemies than the mono group ($Z = -2.351, p < .05, d = .71$).

In total they spent a significantly shorter part of their time at the left or right layers than the mono group ($t_{38} = 2.140, p < .05, d = .68$)(see Figure 8). This was visible as a trend in level 2 ($t_{38} = 1.981, p = .055, d = .63$). In level 2 the stereo group spent less absolute time in the left depth layer than the mono group ($Z = -2.262, p < .05, d = .51$) and played proportionally more of their time in the central depth layer ($t_{38} = -1.981, p = .055, d = .63$). In level 3 the
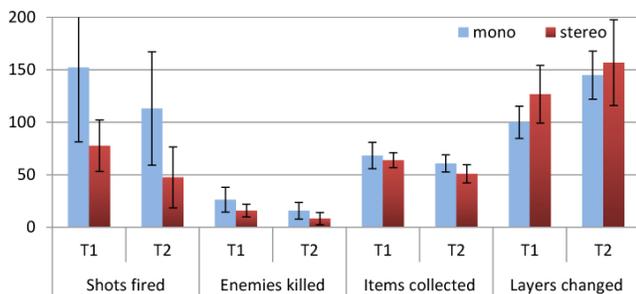
Figure 9: The player behavior is described through mean number of shots fired, enemies killed, items collected and depth layers changed for both data points: T1 (side-scrolling) and T2 (behind-view). Error bars show 95% CI.

stereo group died less in the central depth layer than the mono group ($Z = -2.422$, $p < .05$, $d = .66$). In level 3, the stereo players died proportionally and significantly less at the centered depth layer ($t_{37} = 2.303$, $p < .05$, $d = .74$). In level 1, stereo player spent proportionally less time in comparison to their total time for this level in the left depth layer ($t_{38} = 1.993$, $p = .053$, $d = .63$). Similar to T1 our stereo group fired less shots than the mono group, visible tentatively as a medium sized trend ($Z = -1.952$, $p = .051$, $d = .72$). Our stereo group also killed less enemies than the mono group ($d = .51$, see Figure 9) and collected less items ($d = .55$).

Player Comments
Half of the users (20) gave comments, seven from the mono group and thirteen from the stereo group. Positive comments from multiple users in both groups found the game fun, a nice idea, challenging, dynamic and rich in variety with good level design and visually appealing. They praised the user interface as well-designed, easy and clear, offering good and sensitive controls. One mono player noted the graphical frame of the personality scope helpful. One stereo player also approved of the tutorial, two praised sound design and one liked the landscape in stereo. Two comments in each group judged the behind-view easier. Two stereo players found the behind-view more fun than the side-scrolling perspective.

Negative comments reported it difficult to judge distance between layers, with four comments in the mono group and two comments in the stereo group. Three across both groups noted minor bugs with collision detection. Two subjects from each group noted difficulties for the behind-view caused by occlusion of foreground hills, solvable through semi-transparency. One stereo subject wanted to look around the hills using head-tracking. One mono subject stated the game would be easier in stereo. One stereo subject wanted the game to be faster. One stereo subject recognized the depth cue conflict with the GUI rendered deeper than an occluded nearer hill. Another stereo subject noted the energy bar hard to read in the behind-view when flying deeper to increase speed.

DISCUSSION

Game Experience
The reported differences in game experience were of small effect and non-significant. Our stereo group had a slightly more immersive, flowing, tension and challenging experience than the mono group for the side-scrolling game but these results are not likely reliable for other groups. Even if the statistical relevance could be increased by a larger group, the impact of stereoscopy on the experience on affect measured through GEQ and the fun-related items would probably remain subtle. We have to reject our first hypothesis H1. These results contradict previous findings described in [16], where immersion was clearly raised by the use of stereo in two of the three tested games. There are two possible reasons for this difference. One, the game type and implementation of Deepress3D does not benefit from stereo concerning the tested items, or two, the within-subject design applied in the former study has influenced the outcome in a way that the effect in immersion has been amplified due to the emphasis on the comparison of the two vision groups in such a design.

Presence
The feeling of spatial presence was significantly increased in the stereo group over the mono group especially with the side-scrolling perspective but also in the behind-view. We can therefore accept our second hypothesis H2: presence will be improved by stereoscopy. The results for the behind-view are valid under the condition that the user has played the side-view before. Looking at the within-subjects effects in the mono-group (indicating a better Spatial Situation Model at T2 in comparison to T1), it seems as providing an alternative spatial perspective with monoscopic vision contributes similarly to the imagination of a spatial model as adding stereoscopy to the side-view perspective. A potential benefit from S3D in a higher presence could arguably be compensated with mono vision by offering multiple perspectives or views to a scene. This cue can be related to discussions which compare stereo vision and motion parallax concerning their impact on user performance [22].

Simulator Sickness
Concerning the third hypothesis on simulator sickness, the data provides contradicting findings. The results clearly show more symptoms in the stereo group than with mono users at all data points which would argue against hypothesis H3. However, this difference seems to be the case already at the baseline. In comparison to the baseline, playing the game in the side-perspective caused no significant impact on simulator sickness in both vision groups. This effect points in favor of our hypothesis at least for this game mode. For the stereo group, the behind-view session significantly increased overall symptoms in comparison to after the first game task which contradicts the hypothesis.

One important question is why the symptoms were already higher at baseline in the stereo group. Possibly influential differences in the procedure were the medical test image and measuring interocular and viewing distances only conducted with the stereo group. Also knowing that they were playing in stereo could theoretically pose additional psychological pressure onto experiential subjects. On the other hand, these testing and measuring procedures were very short (< 1 minute in total) and followed by the game introduction in which all participants read a sheet of paper for about five minutes. This

should have nullified any possible symptoms caused by a still image shown at very low parallax levels for a few seconds. Still, the subjective sensitivity towards any such symptoms might have been influenced due to this difference in procedures. An analysis of correlation of any other factor with SSQ results at baseline has brought no significant results.

In conclusion, we argue to interpret the SSQ in relative change to the baseline. Hence, stereo vision has no significant impact on simulator sickness with the side-scrolling view, but there was an impact with the behind-view. Given a combination of individual S3D settings and a constant perspective with an UI optimized to reduce parallax changes in vision, the hypothesis H3 can be accepted only without a constant change in depth animation. As we included the behind-view in our hypothesis, the results show an impact on simulator sickness in case of a perspective with a lot of depth animation only in S3D vision, the hypothesis H3 is rejected.

Game Performance

The primary performance goal was to complete the game as quickly as possible. The overall completion times did not differ significantly between the mono and the stereo group and there was no overall effect. The effects found were small and not significant for any specific levels or data points. In conclusion, a clear contribution from stereo to performance in the actual design of the game Deepress3D was not found, the hypothesis H4 is therefore rejected.

Gameplay Behavior

The results show that mono players used more time to slow down the scrolling by moving the ship to the backward edge of the scope. Stereo players had a tendency to change more often their depth layers, especially in level 2, the first level with enemies. Another interesting effect is the lower mean for the number of shots fired. While the effect was of medium size for both data points it was not significant. However, this score was not tracked for the separate levels which could have brought more insight in level 2. The usage of depth layers was different but these effects stayed subtle and varied strongly between levels and data points.

One problem of the behavioral data is the feature of respawn: When the ship explodes, its position is set back to a fixed respawn point of which there are five in each level. A cumulative number of deaths and respawns increases the number of events in a certain level section and adds noise to the data.

In a more detailed analysis we traced the logs of all participants and counted events according to their progress in the game (0-100%). We did this separately for the two vision groups and data points and processed the data according to the following procedure: The absolute counts of events in each group were normalized to events per participant by dividing through the number of participants in each group. We then counted the number of respawns for each of the fifteen sections (five per level) per group and data point. As the ship might have exploded in an early or late position of a section without having played through a large part of a section, we took half of the respawns as an indicator for the number of

repeated play-throughs and added one for the initial play-through to generate a number of total approximated play-throughs. For each level section, we counted the events and divided them by the amount of approximated play-throughs. We also looked at the sections and identified parts of the level design which contribute to the described differences.

As an example, section 2 in level 2 is the first section with a lot of of shooting enemies. In the side-view, stereo players fired 5.1 shots, performed 3.3 layer changes, died 2.1 times of colliding with enemy ships and 0.3 times by staying out of scope. Mono players fired 10.8 shots, performed 1.5 layer changes, died 0.8 times of colliding with enemy ships and 0.1 times by staying out of scope. Stereo players fired half the shots, changed layers more than twice as much, and died three times a much from leaving the scope.

A second example, section 5 in level 1, features a labyrinth of moving boxes (see Figure 1). In this section stereo players performed 6.5 layer changes, changed 0.3 times to slow speed and 5 times to fast speed and died 0.3 times from box collisions. Mono players changed layers for 5.2 times, changed 1 time per run to slow speed, 3.6 times to fast speed, and died 1.41 times from box collisions.

These and other examples in the data indicate that stereo and mono players deal differently with very specific game mechanics and situations in level design, confirming our fifth hypothesis. Many players in the mono group seemingly tried to fight the enemies by shooting them and stereo players preferred to move around them by changing more between depth layers. One possible interpretation is that, in contrast to boxes and hill obstacles, an enemy can shoot, change a layer or fly towards the player. An enemy's reaction is unpredictable as the visual design does not provide any hint regarding a likely behavior. The stereo might invite players more to utilize the spatial environment when a quick decision becomes necessary in the case of an approaching enemy. However, other parts of the game did not reveal such differences. By identifying such regions and modifying the level design, a game designer could easily create more advantages for either stereo or monoscopic players.

Opportunities and Challenges

The effects found in Deepress3D have to be compared against the expectations during the creation process. As designers of the game Deepress3D, we would have intuitively expected a visible impact on user performance from playing the game in stereo. This expectation was also put up from external people who looked at the game and tried it out. Everyone who saw the game found it much clearer and easier in stereo; likewise, one monoscopic player had commented that the game would probably be a lot easier in stereo. These expectations have proven wrong in the current state of the game. The game was developed with prior experience in commercial game development which has resulted in awarded commercial products. The level of expertise in stereoscopy can be validated through the number of peer-reviewed publications over the course of several years. Given this background, the discrepancy between subjective intuitive expectations and measured results

is considerable. We therefore argue that it is objectively difficult to control a specific effect on user performance and experience mas intended.

Still, it appears theoretically possible to control such an effect. Additional development iterations would support stereoscopic player behavior and may create a different experience. Such progress with more games will presumably help to identify more generalizable factors in game design that impact player behavior and potentially user experience.

CONCLUSION

We provide the first evidence that stereoscopic 3D vision can change how users play a video game, regarding the decisions they make and the actions they perform during gameplay. This conclusion is grounded in the design and evaluation of Deepress3D, a custom-created S3D-specific game which tasks its player to fly and fight through a spatial obstacle course. In comparison to previous research, Deepress3D allows for multiple evenly balanced strategies in gameplay, incorporates design guidelines and principles from previous research on S3D vision in video games, and is successfully evaluated as a fun and highly usable experience in both vision modes. These advantages make Deepress3D a representative and valid game case for both monoscopic and stereoscopic vision while collecting detailed gameplay metrics.

Our analysis further confirms previous findings that S3D vision increases experience of spatial presence but also indicates simulator sickness occurring in combination with depth-animation. Other than intuitively expected, such effects do not necessarily lead to a difference in emotional experience or overall performance, emphasizing the challenge of creating benefits from S3D vision. However, knowing that certain settings of spatial conflicts lead to choosing different gameplay strategies with S3D vision highlights the potential for testing gameplay specifically with stereo players.

Furthermore, these findings indicate the need for distinguishing stereo users as a distinct group in HCI research. Future attempts should explore how differences in behavior can yield benefits in S3D user experience. Besides stereo, an impact on user behavior should be expected and examined more generally with other technologies which increase perceived visual output fidelity.

REFERENCES

1. Banks, M. S., Read, J. C. A., Allison, R. S., and Watt, S. J. Stereoscopy and the Human Visual System. *SMPTE Motion Imaging Journal 121*, 4 (June 2012), 24–43.

2. Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology 45*, 4 (2009), 624–634.

3. Brooke, J. SUS: A quick and dirty usability scale. In *Usability evaluation in industry*, P. W. Jordan, B. Weerdmeester, A. Thomas, and I. L. Mclelland, Eds. Taylor and Francis, London, 1996.

4. Emmerich, K. *Helping Friends or Fighting Foes The Influence of Collaboration and Competition on Player Experience in Digital Games*. Master's thesis, University of Duisburg-Essen, 2012.

5. Howarth, P. 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 (Mar. 2011), 111–22.

6. Kennedy, R., Lane, N., and Berbaum, K. Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology 3*, 3 (1993), 203–220.

7. Kulshreshth, A., Schild, J., and LaViola, J. J. Evaluating user performance in 3D stereo and motion enabled video games. In *Proceedings of the International Conference on the Foundations of Digital Games - FDG '12*, ACM Press (New York, New York, USA, 2012), 33.

8. LaViola Jr, J., and Litwiller, T. Evaluating the Benefits of 3D Stereo in Modern Video Games. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, ACM (2011), 2345–2354.

9. Mahoney, N., Oikonomou, A., and Wilson, D. Stereoscopic 3D in video games: A review of current design practices and challenges. In *2011 16th International Conference on Computer Games (CGAMES)*, IEEE (July 2011), 148–155.

10. Nacke, L., Schild, J., and Niesenhaus, J. Gameplay experience testing with playability and usability surveysAn experimental pilot study. In *Playability and Player Experience: Proceedings of the Fun and Games 2010 Workshop*, K. Calvi, L.; Gualeni; S.; Nuijten, K.; Nacke, L.; Poels, Ed., Breda: NHTV Expertise Series (2010), 31–45.

11. Nacke, L. E. *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment*. PhD thesis, Blekinge Institute of Technology, 2009.

12. Rajae-Joordens, R. J. Measuring experiences in gaming and tv applications. In *Probing Experience*, J. Westerink, M. Ouwerkerk, T. Overbeek, W. Pasveer, and B. Ruyter, Eds., vol. 8 of *Philips Research*. Springer Netherlands, 2008, 77–90.

13. Rivett, J., and Holliman, N. S. Stereoscopic game design and evaluation. In *IS&T/SPIE Electronic Imaging*, N. S. Holliman and G. E. Favalora, Eds. (Mar. 2013), 864813–864813–13.

14. Sauro, J., and Lewis, J. R. *Quantifying the User Experience: Practical Statistics for User Research*. Morgan Kaufmann, 2012.

15. Schild, J., Bölicke, L., LaViola Jr, J., and Masuch, M. Creating and Analyzing Stereoscopic 3D Graphical User Interfaces in Digital Games. In *Proceedings of the 2013 ACM annual conference on Human Factors in Computing Systems - CHI '13*, ACM (New York, USA, 2013).

16. Schild, J., LaViola, J., and Masuch, M. Understanding User Experience in Stereoscopic 3D Games. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, ACM Press (New York, New York, USA, 2012), 89.

17. Schild, J., and Masuch, M. Fundamentals of Stereoscopic 3D Game Design. In *Proceedings of ICEC 2011*, Ifip (2011), 155–160.

18. Schild, J., Seele, S., and Masuch, M. YouDash3D: Exploring Stereoscopic 3D Gaming for 3D Movie Theaters. In *Proceedings of SPIE Vol. 8288: Stereoscopic displays and applications XXIII (2012)*, SPIE (2012).

19. Shibata, T., Kim, J., Hoffman, D. M., and Banks, M. S. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision 11*, 8 (Jan. 2011), 1–29.

20. Takatalo, J., Kawai, T., Kaistinen, J., Nyman, G., and Häkkinen, J. User Experience in 3D Stereoscopic Games. *Media Psychology 14*, 4 (Nov. 2011), 387–414.

21. Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner, T., and Jäncke, P. MEC spatial presence questionnaire (MEC-SPQ). *Report to the European Community, Project Presence: MEC (IST-2001-37661)* (2004).

22. Ware, C., and Mitchell, P. Reevaluating stereo and motion cues for visualizing graphs in three dimensions. In *Proceedings of the 2nd symposium on Applied perception in graphics and visualization - APGV '05*, ACM Press (New York, New York, USA, Aug. 2005), 51.

23. Zerebecki, C., Stanfield, B., Hogue, A., Kapralos, B., and Collins, K. S3D depth-axis interaction for video games: performance and engagement. In *IS&T/SPIE Electronic Imaging*, N. S. Holliman and G. E. Favalora, Eds. (Mar. 2013), 86481B–86481B–8.