EGVE Symposium (2008)
B. Mohler and R. van Liere (Editors)

Display Devices for Virtual Environments: Impact on Performance, Workload, and Simulator Sickness

J. Conradi and T. Alexander

Research Institute for Communication, Information Processing and Ergonomics (FGAN/FKIE)
Wachtberg, Germany

Abstract

Usability and, thus, success of Virtual Environments (VE) systems are closely related to the type of display used. Applicable VE-displays range from simple desktop monitors with low immersion to high-end, immersive HMDs. It is often inferred that more sophisticated displays always produce higher performance. In this paper this opinion is critically questioned. To estimate effectiveness and usability of the display measures of human performance, subjective workload, and simulator sickness serve as critical criteria. The effect of three different displays (desktop monitor, projection wall, HMD) with varying degrees of immersion on each of the criteria was analyzed empirically. In the experiment n=30 participants performed an exploration task and memorized waypoints in a virtual village. The VE itself was created by integrating a commercial game engine as rendering system into our laboratory VE system. The results of the experiment showed no significant differences between displays for performance and simulator sickness. However, workload was rated lowest for desktop monitor and highest for HMD. Furthermore, the amount of simulator sickness decreased significantly between trial sessions. The results show that more immersive displays are not always the best choice per-se. The selection of display has to consider requirements of the application and user characteristics, especially training and VE-experience.

Category and Subject Descriptor: I.3.6: Computing Methodology, Computer Graphics, Methodology and Techniques

1. Introduction

Virtual Environments (VE) subsume new methods and technologies for presenting and experiencing synthetic, computer-generated scenes in an intuitive, natural, and vivid way. This requires the utilization of new technologies and devices for presentation and user interaction. Because of the dominance of the visual modality the characteristics and capabilities of visual displays critically influence the overall VE system usability and performance.

There are various ways to classify visual displays. They are either based on technical specifications or on subjective, more human-oriented factors. Both correlate with each other. Whereas technical specifications (e.g., brightness, refresh rate, field of view) can be determined and quantified relatively easily, subjective factors (e.g., image vividness, immersion) are vague. However, they have to be considered for a more extensive analysis of the Human-VE system performance. For such an analysis it is important to include objective measures of performance, but also subjective measures like human workload and the amount of negative side-effects like e.g. cyber sickness.

The work presented in this paper addresses these issues and analyzes the effects based on experimental data. The target application was a navigating and experiencing application for gaining orientation knowledge in an urban environment. Three different visual displays with varying degrees of immersion were compared to each other.

2. Related Work

Based on their technical specifications displays provide different degrees of immersion. Full or high-immersion systems provide a completely simulated field-of-view. They often include additional, non-visual displays as well to enhance the comprehensive experience of the VE. Technical solutions for such systems are Head-Mounted Displays (HMDs) or projection rooms with multiple projection planes. Instead, a low-immersion system provides just a partial field-of-view. Larger parts of the real environment can still be perceived. Projection tables, projection walls or desktop monitors are typical examples for such displays. Based on this terminology, each VE display can be assigned to a degree of immersion [BKLP04]. Immersion can bee seen as a physiological state, in which one perceives oneself enveloped by a VE and thus, experiences the feeling of “being there” [WS98]. The perceived sense of immersion differs for displays as well as for individuals.

One of the close interrelations between technical specification and immersion is the field of regard (FOR). It defines the angular extend of visual information that can be perceived, including eye movement and head rotation. It refers to the amount of physical space surrounding the user in which virtual images are displayed. A more common term, the field of view (FOV), refers to the maximum number of degrees of the visual angle that can be perceived with *a stable eye* on the display [BKLP04]. Consequently, the FOV is always the same or smaller than the FOR. In our analysis we will refer to the FOR. The proportion of synthetic and real visual information serves as a main input criterion for immersion: The larger the synthetic part, the higher the degree of immersion is.

There has been few research on the dependence of performance on immersion. According to Ruddle [RPJ99], the dimension of the field of view affects the size perception in a VE. This finding is based on a survey of user behaviour and performance with a HMD compared to desktop display. Participants using HMD navigated faster through a building and developed a more accurate sense for straight-line distances. In particular, they took advantage of the natural interaction interface provided by the HMD and head-tracking.

Tan et al. [TGSP06] describe another attempt for quantifying the effect of display size. They compared the performance of users working with a large monoscopic projection wall to users working on a standard desktop monitor. Their results show higher performance in mental rotation tasks, 3D navigation, mental map formation, and memorising with the large display. In an empirical study Patrick et al. [PCS*00] compared the influence of different monoscopic displays on creating mental maps of a virtual environment. No significant differences were found.

It has to be highlighted that performance depends strongly on the special application. Commonly used measures of performance include error rate, accuracy, and time to complete a task. It is often considered that both correlate positively with each other, so that high immersion always results into higher performance. As shown in the previous section this is not always the case.

A further relevant criterion for estimating usability is workload, which can be considered as subjective task difficulty as well. Whereas the previous measures of performance are objective, workload reflects the subjective aspect of human performance. Though objective measures are constant, workload varies between users. One way to measure workload is a subjective rating scale. Though they are characterized by large variance, they are still the most comprehensive way [PK88].

For practical usability the appearance of negative physiological side-effects like e.g. nausea, headache, or disorientation have to be considered as well. These effects are subsumed by the term “cyber sickness”. Cyber sickness is a special form of simulator sickness, which occurs with all kind of virtual simulation. The symptoms are alike, but the intensity of these symptoms varies between simulator and cyber sickness [SKD97]. Therefore, cyber sickness is often measured and analysed by means of simulator sickness. Simulator sickness is influenced by characteristics of user, system, and task [KGL*96]. In this study the system characteristic “display system” and the user characteristic “experience” are investigated. The display system varies between the different displays with differing degrees of immersion. The experience is applied by the session schedule and, thus, the number of VE experiences. Past studies show that the type of display causes negative side effects because of the degree of immersion [AGM*05].

Although it would be desirable to measure simulator sickness during the VE experience, no reliable method of measurement for side effects is available to date. Instead, subsequent questionnaires are used. A commonly used questionnaire is the “Simulator Sickness Questionnaire” (SSQ) by Kennedy et al. [KLB*93]. The amount of subjective sickness is measured by rating 16 items of individual state. [SKD97].

Nowadays, VEs are used for many applications and new operational areas. One trend in this is the impact of gaming technology. By utilizing rendering capabilities of games, the quality and vividness of the displayed virtual scene is increased. Moreover, available tools and examples for scene generation make an application even for scientific purposes reasonable [AGM*05]. Gaming software, which allows this kind of modification, are for example Half-Life 2 [Val08], Far Cry [Ubi08] and Quake 3 [Ids08].

3. Method

The overall objective of this analysis was to assess the effect of different types of stereoscopic displays on practical usability. Our target application was transferring waypoint knowledge in a synthetic, village-like scene. The main criteria for usability were performance, workload/subjective task difficulty, and simulator sickness. The measures were taken subsequently to the VE sessions.

The selected displays offer different degrees of immersion (desktop monitor, projection wall, HMD). The main difference is the field of regard. For experimental design a between-factors design was chosen to minimize training effects. Consequently, each participant used one display only.

3.1. Participants

Originally, 37 participants volunteered to take part in the experiment. However, 4 of them were excluded due to vision deficiencies (checked before experimental trials). 3 of them refused a further participation due to exceeding simulator sickness during first experimental session. Consequently, n=30 completed the experiment successfully. The participants were aged 31.0 ± 7.0 (Mean ± SD). 27 of them were male. Gaming experience was also recorded by subjective rating. A first analysis showed significant effects of gaming experience neither on performance nor simulator sickness.

Each participant was assigned to one of three groups. Each group performed the experiment with one display only. The total experiment consisted of four experimental sessions. Each of them took about 30 minutes. The sessions were separated by at least one day.

3.2. Task

Because the target application was obtaining navigation and orientation knowledge in an urban environment, the main task of the experiment was to explore the virtual scene. The participants were familiarized to the exocentric map of the urban environment. This was done by presenting a conventional paper map and testing general location knowledge. Afterwards the participants moved through the virtual scene. They explored it and memorised the position and colour of the added 8 flags.

Each session in the VE lasted 12 min and the subjects navigated at a maximum speed of 12 km/h through the VE. This seems to be too fast for a pedestrian in a reality. But it was found suitable under the VE conditions. However, compared to the speed of a car or the common velocity of a player in a gaming-surrounding, this speed is quite moderate.

After each session, the subject's task was to place the locations and colors of the flags into an exocentric map of the urban environment. Participants were advised to fulfill this thoroughly and efficiently.

3.3. Virtual Environment

The virtual environment was based on the gaming-software engine Quake III Arena [Ids08]. The level resembles a village with the dimensions 115m x 105m. Please see the notes in the acknowledgement.

Within the village, 8 flags have been inserted. Depending on the session and task, their location and color varied. No additional moving objects were inserted in the VE. Figure 1 shows an example screen of the experiment setting.

Figure 1: *Scene of the virtual environment*

Under all conditions rotation and translation were controlled by the participant, no special exploration path was given.

3.4. Independent Variables

The independent variables of the experiment were VE displays and user experience. User experience was determined by the number of performed sessions. The experiment included 4 sessions, requiring a subsequent repeated measurement / within-factors analysis.

The second factor, VE displays, included 3 different displays. The same input device was used for all displays. As handheld, cordless, easily applicable device a gamepad (Logitec Rugmblepad) was selected.

Each participant performed the experiment with a single display only, so a between-factors analysis was fulfilled.

DESKTOP MONITOR (DM) (Figure **2**). A standard desktop monitor was chosen as low-immersion display. The 21", 120 Hz desktop monitor was low cost and easily available. For displaying the stereoscopic image, Crystal Eye shutter glasses were used [Rea08]. The frame rate of the monitor and the shutter glasses were synchronised by a wireless infrared-system. This resulted in a 60Hz stereo presentation.

The participants were in a pre-defined position relatively to the monitor. The resulting field of regard stretched 53 degrees horizontally and 41 degrees vertically, respectively. The monitor provided a small-screen, small-field-of-regard display. The gamepad was used for navigation through the VE. Consequently, the monitor provided low immersion in comparison to the other used devices.

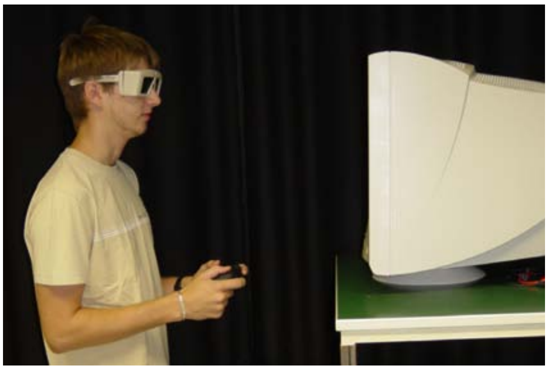

Figure 2: *Desktop Monitor*

PROJECTION WALL (PW) (Figure 3). The second display was a large-scale projection wall with a projection area of 170 x 130 cm. Because of the constant viewpoint of the participants, this resulted in a FOR of 93 x 77 degrees. The stereoscopic presentation applied polarized filters in front of the projection system and light-weight glasses with polarized glass. As input device for translation and rotation the hand-held gamepad was used.

Compared to the monitor, the projection wall had a medium FOR, with the real environment still visible. The PW had medium immersion in relation to the other displays in the setup.

Figure 3: *Projection wall*

HMD (Figure 4). Third device was the NVIS nVisor SX as Head-Mounted Display [Nvi08]. It consists of two miniaturized displays for each eye, thus enabling stereoscopic visualization. Each Monitor had a solution of 1280x1024, and a field of view of 47° x 37°. The display was equipped with blinders masking the peripheral view of the real environment.

Figure 4: *HMD NVIS nVisor SX [Nvi08]*

Markers of a passive, optical infra-red-tracking system were attached to the HMD. By means of a six-camera tracking-system (AR-Track [Art08]) the direction of the head was recorded and the corresponding images were presented. Consequently, view direction was controlled by the head orientation. This resulted into an overall FOR of 360 degrees. So, controlling navigation included both, head orientation for orientation control and gamepad input for positional control. The HMD caused the highest degree of immersion of the used devices.

3.5. Dependent Variables

The selected dependent variables allow a quantification of the effect of display and user experience on usability of a VE system. Usability was determined in three categories: Performance, workload/subjective task difficulty, and simulator sickness. As said before, these measures were taken after the VE session.

PERFORMANCE: The task during the VE session, was to memorise location and colour of eight flags in a virtual village. The individual performance in memorising was determined after each session. Therefore, the participants drew locations and colours of the memorized flags on a map. The location error between real and memorized position were calculated and served as one dependent variable for performance. As a second measure of performance, the time required to complete the mapping was logged.

WORKLOAD AND SUBJECTIVE TASK DIFFICULTY: A subjective, two-level rating scale was used to determine workload. The terminology used for linking values to meanings refers to subjective task difficulty. A rating scale with the minimum of 0 and the maximum of 100 was used. The participants assessed the difficulty of the task, whereas the lowest level ('0') was 'very easy' and the highest ('100') was 'very difficult' [PK88].

SIMULATOR SICKNESS: To minimize effects of habituation to the real environment, cyber or simulator sickness was determined immediately after completing a VE session. For the experiment the SSQ by Kennedy et al. [KLB*93] was used. The scale ranges from a minimum of 0, i.,e. no perceived symptom, to a maximum of 235.6, i.e. all symptoms are "strong". According to Stanney [SKD97], simulators can be categorised based on the SSQ-rating into categories between 'no symptoms' ('SSQ=0') and 'bad simulator' (SSQ>20).

4. Results

For determining differences between the displays and, thus, effects of immersion on usability, methods of inferential statistics were applied. At first, Kolmogoroff-Smirnoff-Test for the distribution of each measure was carried out. For normal distributed measures, different analyses-of-variance (ANOVA) were carried out. In case of significant effects, subsequent post-hoc tests were performed. In case of non-normal distributed data, appropriate tests were used instead. The results of these analyses are presented in this chapter.

4.1. Performance

Performance was determined by the temporal and accuracy measures described in the previous chapter. The one-way ANOVA showed no significant differences between the displays. Neither time ($p=0.78$) nor accuracy ($p=0.23$) varied. This can also bee seen in Figure 5.

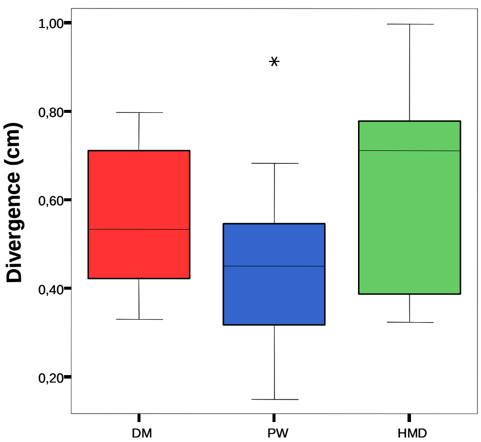

Figure 5: *Boxplot of the divergence of given and mapped positions of flags.*

For the experience, i.e. performance between sessions, no effects were found ($p>0,05$).

4.2. Workload and subjective task difficulty

ANOVA showed significant differences between displays for workload and subjective task difficulty ($p<0,01$). A post-hoc Scheffé-test specified significant differences between monitor and HMD. No significant differences were detected for the projection wall.

The mean rating for the monitor was $\bar{x}_{DM}=22.3$, which resembles 'quite easy' in the questionnaire. The projection wall caused intermediate workload because the subjective task difficulty was rated medium. With $\bar{x}_{PW}=36.35$ the display was rated as 'rather easy than difficult'. The HMD, providing the largest field of regard, resulted into the highest workload. Subjective task difficulty was rated as 'neither easy nor difficult' ($\bar{x}_{HMD}=48.7$, see Figure **6).**

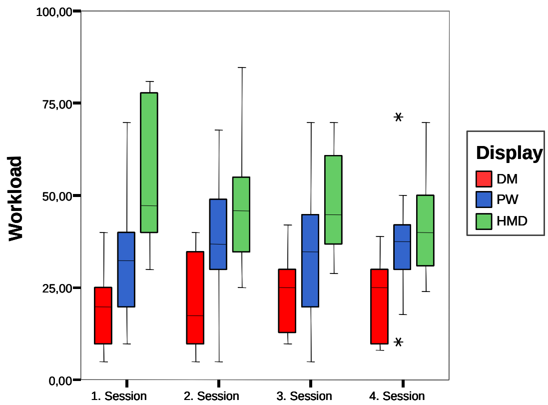

Figure 6: *Boxplot of the level of workload*

ANOVA revealed no significant differences between sessions ($p>0.05$). Consequently, there was no effect of experience on workload.

4.3. Simulator sickness

Ratings of simulator sickness were not normal distributed. Therefore a Wilcoxon-test for matched pairs was carried out. The test found significant differences between the SSQ rating before and after each session ($p<0.01$). A significant effect of display on simulator sickness was not found ($p>0.05$).

A Friedman-test for 4 matched samples revealed significant differences between sessions ($p<0.03$, see figure 7). This is caused by increasing experience from session to session. The median of the size of the simulator sickness was $\bar{x}_1=13.1$ for the first session and $\bar{x}_2=5.6$, $\bar{x}_3=7.4$ and $\bar{x}_4=3.7$ for the subsequent sessions. According to Stanney's categorisation [SKD97], the rating of the VE lowered from 'significant symptoms' (session1), and 'minimal symptoms' (session 2 and 3) to 'negligible symptoms' (session 4).

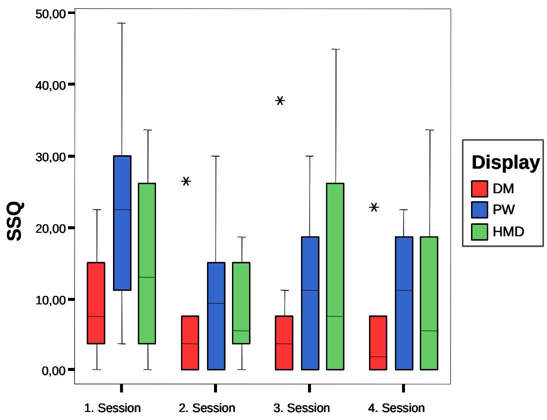

Figure 7: *SSQ-data according to display and session*

5. Conclusion

The results show no differences for orientation performance between the different displays. This is an important finding because it contradicts the general opinion that more immersion always results into higher performance. Especially orientation tasks and gaining navigation and orientation knowledge is independent from immersion. For some applications, simple VE displays with low-cost desktop monitors only can be as effective as high-tech immersive systems.

With regard to workload and subjective task difficulty, the HMD was rated as more difficult than the desktop monitor. This can be contributed to the 'closed' HMD which isolates the user from real world. It is simply more difficult to interact with the interaction device (i.e. gamepad) if there is no visual feedback. An alternative, more natural navigation mode for immersive displays would help to overcome this problem.

Interestingly no effect of immersion on simulator sickness was found. Because of prior findings and the higher workload it was expected that simulator sickness is higher for immersive than non-immersive systems. However, by remembering the short exposition to the VE this finding has to be put into perspective. All measured values are low and have a large variance. This results into a large amount of random noise which covers eventual small effects. For a statistical analysis either larger sample sizes or higher SSQ-ratings would be necessary. Therefore, the explanatory power of the finding is relatively low.

In contrast to immersion, experience contributes to simulator sickness. This is explained by the adaptation to the VE.

As a result it is summarized that the simple expectation "the more immersion the better a VE system is" cannot be generalized. For the navigation and orientation tasks as described in the paper, simple systems with low immersion can result into the same performance as high-tech systems. Moreover, an exposition of 12 minutes is still a safe way to prevent simulator sickness for most users. However, it has to be considered that a large proportion of participants experienced extreme levels of simulator sickness even after shorter times and independent of display. This highlights the importance of safety issues to prevent further negative side-effects and enhance usability.

6. Acknowledgement

Furthermore, we would like to thank Mr. Tristian Barrett, Sydney, Australia, for sharing the virtual village used for our experiments.

7. References

[AGM*05] Alexander, T., Goldberg, S., Magee, L., Rasmussen, L., Delleman, N., Borgvall, J., Lif, P., Smith, E., Cohn, J. (2005): Compendium on National Research Activities in Augmented, Mixed and Virtual Environments. RTO-TR-121 Part I. Neuilly-Sur-Seine: NATO RTA.

[Art08] AR-Tracking (2008): System. Internet page: http://ar-tracking.eu/Products.8.0.html. Visited: Jan 08.

[BKLP04] Bowman, D. A., Kruijff, E., LaViola, J. J., Poupyrev, I. (2004): 3D User Interfaces. Theory and Practice. Wokingham: Addison-Wesley.

[CGBL98] Chance, S. S., Gaunet, F., Beall, A. C., Loomis, J. M. (1998): Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. Presence: Teleoperators and Virtual Environments Nr.7.

[Ids08] ID-SOFTWARE (2008): Quake-III Arena. Internet page: http://www.idsoftware.com/games/ quake/quake3-arena/. Visited: Jan 08.

[JF97] Jul, S.; Furnas, G.W. (1997): Navigation in Electronic Worlds, SIGCHI Bulletin, Vol. 29, No. 4, A CHI 97 Workshop, October 1997.

[KGL*96] Knerr, B.W., Goldberg, S.L., Lampton, D.R., Singer, M.J., & Witmer, B.G. (1996). Training in virtual reality: Human performance, training transfer, and side effects. Presented at the 1996 IMAGE Conference, 23-28 June 1996, Scottsdale, AZ.

[KLB*93] Kennedy, R. S., Lane, N.E., Berbaum, K.S., Lilienthal, M. G. (1993): Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology Nr.3.

[Nvi08] NVIS (2008): nVIsio SX. Internet page: http://www.nvisinc.com/product.php?id=1. Visited: Jan 08

[PCS*00] Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J. A., Verratti, T. & Chiselko, G. (2000): Using a large projection screen as an alternative to head-mounted displays for virtual environments. CHI 2000 Conference Proceedings. ACM Press, New York.

[PK88] Pitrella, F. D. & Käppler, W.-D. (1988): Identification and evaluation of scale design principles in the development of the Sequential Judgement, extended range Scale. Wachtberg: Forschungsinstitut für Anthropotechnik, Report Nr. 80.

[Rea08] RealD (2008): Crystal Eyes. Internet page: http://www.reald-corporate.com/scientific/crystaleyes .asp. Visited: Jan 08:

[RPJ99] Ruddle, R. A., Payne, S. J., Jones, D. M. (1999): Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays? Presence: Teleoperators and Virtual Environments Nr.8.

[SKD97] Stanney, K.M., Kennedy, R.S., Drexler, J. M. (1997): Cybersickness is not simulator sickness. Proceedings of the 41st Annual Human Factors and Ergonomics Society Meeting (pp. 1138-1142). Albuquerque, NM, September 22-26.

[SLU*96] Slater, M.; Linakis, M.; Usoh, M.; Kooper, R. (1996): Immersion, Presence, and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess. In: Green, M. (Hrsg.): Proceedings: ACM Symposium on Virtual Reality Software and Technology (VRST 96), p. 163-172.

[TGSP06] Tan, D. S., Gergle, D., Scupelli, P., Pausch, R.(2006): Physically Large Displays Improve Performance on Spatial Tasks. ACM Transactions on Computer-Human Interaction, 13(1), pp. 71-99.

[Ubi08] Ubisoft (2008): Farcry: Internet page: http://farcry.de.ubi.com/. Visited 01/2008.

[Val08] Valve (2008): Half-Life. Internet page: http://orange.half-life2.com. Visited 01/2008.

[WS98] Wittmer, B.; Singer, M. (1998): Measuring presence in Virtual environments: A presence questionnaire. Presence: Teleoperators and Virtual Environments, 7, (3).