Spatial Presence in Dynamic Environments: Tracking the Visual Attention Allocation During Exposure

Bartholomäus Wissmath[12], David Weibel[12], Daniel Stricker[12], and Fred W. Mast[1]
[1]University of Bern, Switzerland, [2]Swiss Universitary Institute of Distance Education
{wissmath@psy.unibe.ch}

Abstract

Attention allocation towards the mediated environment is assumed to be a necessary precondition to experience presence. In presence research, however, the potential of visual attention theories or eye-tracking techniques have not been exploited so far. In this study, participants (N= 44) ride on a virtual roller coaster simulation. We compare participants scoring high versus low on presence. In addition, we manipulated the degree of presence (low vs. high). During the ride, the eye movements were captured. In addition, we assessed subjective ex-post presence judgments. We found high sensations of presence to be associated with more fixations, shorter fixation durations, smaller saccade amplitudes, and decreased saccade velocity. We discuss individual scan patterns and reflect on the possible implications of our findings for presence theory and VR-design.

Keywords---**Presence, Visual Attention, Attention Allocation, Scene Perception, Eye-Tracking**.

1. Introduction

Presence includes a perceptual illusion of non-mediation triggered by technical interfaces [1]. Thereby, mediated contents are perceived as real and one's self-awareness is immersed into this other world [2]. Sadowski and Stanney describe presence as "a sense of belief that one has left the real world and is now 'present' in the virtual environment" (p. 791) [3]. Presence in virtual environments implies the departure from the physical environment and the arrival in the mediated environment [3][4][5]. There is a plethora of different presence concepts (e.g. social presence, self-presence, or environmental presence). In this study, however, we will restrict ourselves to spatial presence being the core form of presence. Since the importance of attentional processes was emphasized in previous publications [5][6], we pursue the aim to empirically investigate visual attention allocation in the context of spatial presence.

Draper, Kaber, and Usher [2] introduced an attentional resource model of telepresence in the context of tele-operation. The model distinguishes task-relevant and distracting information across immediate and mediated environments. Thereby, they assume that the probability to experience telepresence is increased when more attentional resources are allocated to the mediated environment than to the immediate environment.

Recently, a comprehensive spatial presence model was introduced [7], which distinguishes two critical steps: In the first step, the focus of attention must be allocated towards the mediated environment and the user has to establish a mental representation of this environment. Then, the second important step is that the media users no longer locate themselves in the immediate environment but rather feel present in the mediated environment. Thus, the model suggests that specific attentional processes are required to experience presence. Thereby, the attention allocation towards the mediated environment may be media-induced (involuntary) or user-directed (controlled). The former results from media characteristics such as high pictorial realism, whereas the latter is associated with user characteristics such as interests and motivation. The authors state that in interactive or/and immersive media, a continuous sensory input captures and maintains the involuntary attention whereas in non-interactive media such as books, the controlled attention processes are central.

Despite the fact that the relevance of attentional processes in the context of spatial presence is evident, there still is scarce empirical research focusing on the attentional processes.

Similar to presence, attention is a complex concept including various sub-dimensions. William James suggested two categories: passive vs. active attention [8]. These two categories have persisted although the modern terms are bottom-up and top-down (the spatial presence model introduced above includes this dimension). In addition, several forms of attention have been proposed: *attentional orientation* (directing the attention to a particular stimulus), *selective attention* (focusing on one particular stimulus instead of another), *divided attention* (distributing the attentional resources over two or more different stimuli), and *sustained attention* (attending to a stimulus over a period of time) (cf. [9]).

Attentional processes have been investigated in different modalities. However, most research has been done in the visual domain, and in *scene perception* in particular. Due to the fact that in the human eye, only a small region of the retina (ie, the fovea) provides high quality visual information, we move our eyes about three time each second. These rapid eye movements are termed saccades whereas the periods of relative gaze stability are termed fixations. Fixations can be seen as a deictic pointer to entities in the environment. Moreover they may act as primary origin for the coordinate systems of vision, motor control and cognition.

Recent work in the field of scene perception has focused on the question whether the eye movements are controlled bottom-up (i.e. based on stimulus characteristics such as contrast) or top-down (i.e. based on memory or cognitive processes) [9].

A central tool in the visual attention research is eye tracking since eye movements are an overt behavioral manifestation of the allocation of attention in a particular scene. Thus, eye movements serve as a window into the operation of the attentional system. Moreover, eye movements provide an unobtrusive and sensitive index of the ongoing visual and cognitive processing [9].

Although Bailenson and Yee [10] introduced head movements as a proxy for gaze to assess attention allocation, the potential of eye movement tracking has not yet been realized in presence research. We think that tracking the eye movements enables us to address several unsettled issues: It is assumed that attention allocation is a prerequisite for spatial presence. However, it is obvious that paying close (visual) attention to a mediated environment only increases the probability that presence emerges. For example, a screener using an x-ray device to check suitcases will usually attend the screen very closely. However, it is unlikely that he or she feels located inside the suitcases. This raises the question to what extent presence is influenced by the visual scene perception.

We also investigate whether strong sensations of presence are associated with specific patterns of eye movements. Such patterns could be of major interest from several perspectives. First, the role of visual attention in spatial presence could be clarified (e.g. is it controlled bottom-up or top-down?). Second, identifying the eye movement patterns triggering presence could have not only theoretical but also practical implications (e.g. for VR-designers). Third, since eye tracking is unobtrusive and highly reliable, the identification of specific "presence scan patterns" could form the basis for a new indicator of spatial presence. Fourth, scene perception in mediated environments could help to understand scene perception in natural environments even better. Fifth, the eye movement patterns could bear light on the cognitive processes during presence experiences.

2. Method

2.1. Design

We used a between subjects design comparing a high presence and a low presence group. The mediated environment used as a stimulus was a virtual roller coaster simulation. We chose the auditory channel to manipulate presence since haptic or visual manipulations would have directly influenced the eye movements. We want to point out that our two conditions were visually identical – the only difference between high and low presence conditions was the sound (present vs. absent).

Prior to the experimental condition, participants once rode on a virtual roller coaster simulation. Thereby, the participants of the high presence condition had no sound effects, whereas participants of the low presence condition had sound effects. This pre-trail made our presence manipulation more salient since participants in the high presence condition experienced a sensory richer environment than during the first ride, whereas for participants in the low presence condition experienced the opposite manipulation.

2.2. Material

We used a commercially available rollercoaster simulation [11]. To rule out any differences between the trails (e.g. due to different real-time image rendering or viewing angles), we generated a high-resolution video clip displaying a ride on the track "Plutonium" (cf. figure 1). We switched off speed displays and chose good weather conditions (ie, sunny day). To present the ride we used a 46'' LCD television.

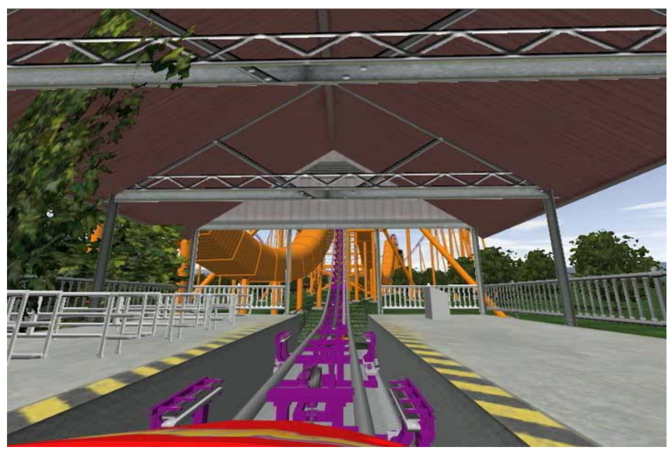

Figure 1 Screencap of the rollercoaster simulation

2.2. Participants

Forty-four undergraduate students enrolled in Psychology volunteered to participate in this investigation. Mean age was 22.14 years (*SD* = 4.06). Among those, 40 were female. They received an extra credit for their participation and could end the experiment at any time.

2.3. Measurement

To track the eye movements we used the eye link II device [12]. This is a head mounted video-based eye tracker using infrared light to monitor the pupil–corneal reflection. The average accuracy is high (usually < 0.5°). The device allows for wearing glasses and head movements up to 30°. We tracked the subject's dominant eye with a sampling rate of 500 Hz. Thereby, we captured fixation location, fixation duration, saccadic amplitude, and saccade velocity. Eye movements smaller than one degree of visual angle within two measures were integrated as one fixation, whereas all eye movements greater than one degree within two measures were counted as saccades.

To assess the subjective sensations of spatial presence we used the MEC spatial presence (MEC-SPQ) [13] questionnaire. This instrument assesses nine constructs associated with spatial presence including traits (e.g. absorption, imagery skills) and spatial presence states (e.g. spatial situation model, possible actions). Since we focus on the process we included the most relevant spatial presence dimensions for our study. Those consist of attention allocation and self location. Both dimensions are represented by eight items. According to the authors, the scales' reliability is high (*Cronbach's alpha* = .93). As suggested, we used 5-point Likert scales ranging from 1 ('I do not agree at all') to 5 ('I fully agree'). These questionnaire data and the demographics were collected on a computer.

2.3. Procedure

After informed consent was obtained, the participants were seated in front of the LCD screen. Then, we put the head mounted eye tracker on the participant's head and calibrated the eye tracking system. Before first the roller coaster ride started, participants were told that they could enjoy the following presentation without any task. When the ride had ended, we removed the eye tracker and participants answered the questionnaires. Before the second trail started, we anew mounted the eye tracker on the participant's head and again calibrated the system. After the second ride, the tracker was removed and the questionnaires were filled out. We then debriefed and thanked the participants. The whole experiment lasted about 20 minutes.

3. Result

In a first step, we calculated the reliabilities of the two MEC-SPQ scales. Self location and attention allocation were both highly reliable (*Cronbach's alpha* = .97, *Cronbach's alpha* = .96 respectively). The means and the standard deviations of the scores can be seen in figure 2.

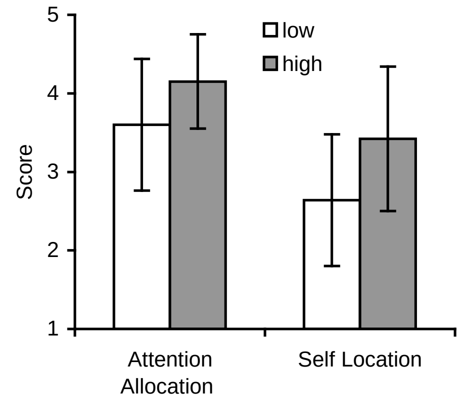

Figure 2 Means (+-*SD*) of the two presence indicators in high vs. low presence conditions

Then, we performed a manipulation check. As intended, lower scores of self location (*M* = 2.64; *SD* = .84) and attention allocation (*M* = 3.60; *SD* = .84) in the low presence condition resulted than in the high presence condition, self location (*M* = 3.42; *SD* = .92) and attention allocation (*M* = 4.15; *SD* = .60). The corresponding *t* tests turned out significant, $t\ (42) = 2.91$; $p < .01$, and $t\ (42) = 2.55$; $p = .01$, respectively. Inter-individual differences in both dimensions were high (cp. Figure 2). We calculated median splits and found that 65 % percent of the participants being in the high presence condition provided subjective presence ratings (ie. self location) above the median. Congruently, 35 % of the subjects in the high presence condition provided self-location scores below the median. Therefore we used median split categorisation in the following analyses.

We compared the low and high presence groups (ie. regarding self location) in terms of total amount of fixations, average fixation duration, saccade amplitude, and saccade velocity. High presence was associated with more fixations, shorter fixation durations, smaller saccade amplitudes, and decreased saccade velocity. Table 1 presence the corresponding *t* tests. All of those were significant.

Table 1
Eye Movement Patterns in Low vs. High Presence Conditions

	Mean		SD				
Indicator	low	high	low	high	*df*	*t*	*d'*
Fixation amount	275.95	303.46	34.20	58.70	6502	25.06***	.56
Fixation duration	472.38	435.95	326.85	309.95	7231	4.96***	.12
Saccade amplitude	3.56	3.36	3.02	2.77	7145	2.98**	.07
Saccade Velocity	95.80	93.10	48.53	50.24	7399	2.36*	.06

Note. *d'* = Cohens effect size *d*. *** < .05. ** < .01. * < .001

Besides those objective overall indicators, we also considered the aggregated scan-paths over all participants. We found no clearly visible difference in the overall patterns. On an individual level, however, eye data give an objective account of the viewing behaviour. The figures 3 and 4 show fixations and saccades of two participants. The first participant experienced strong sensations of spatial presence, whereas the second one scored very low on spatial presence. The patterns show that during high presence the attention allocation is more focussed to the centre of the picture compared to low presence. In addition, the figures illustrate that the saccade amplitude is greater during low presence.

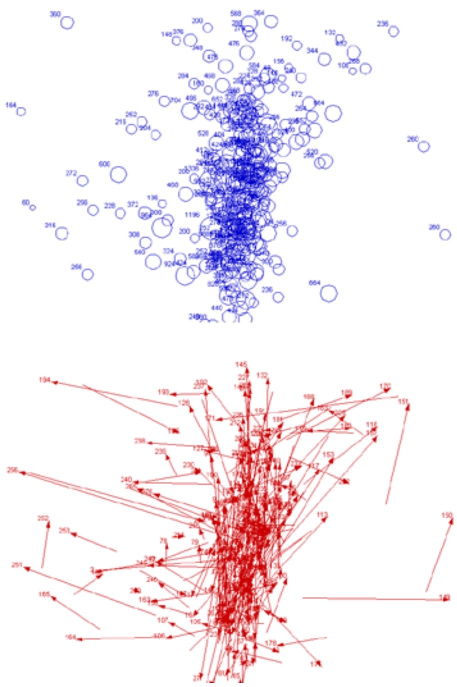

Figure 3 Fixations (blue) and saccades (red) of a participant in the high presence condition scoring high on self location

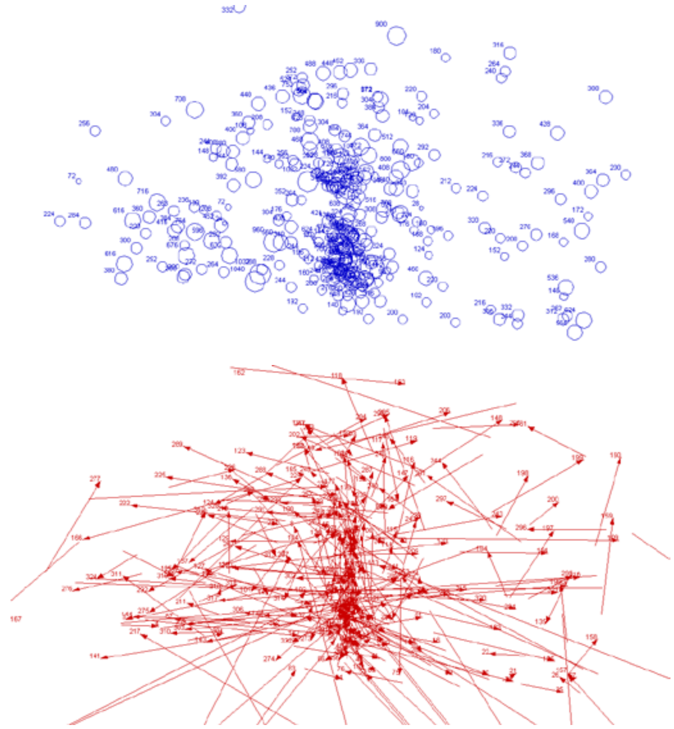

Figure 4 Fixations (blue) and saccades (red) of a participant in the low presence condition scoring low on self location

4. Discussion

Most noteworthy, our findings corroborate central assumptions of spatial presence theories. The strongest effect – a greater amount of fixations during strong spatial presence compared to low spatial presence – can be seen as empirical evidence for the assumption that presence increases with attention allocation towards the environment if the total amount of fixations is interpreted as an indicator of attention allocation. That is, the more often particular object or environment is fixated, the more attention is allocated to it. Notice that the high and low presence conditions used exactly the same visual input. Therefore, any differences in eye movements cannot be attributed to changes in the visual input stimulation.

The average duration of the fixations was shorter during high presence compared to low presence. Thus, it seems that in dynamic environments, high presence is associated with a more focused and active viewing behavior (ie, in terms of fixation duration rather than saccade velocity). Participants experiencing high presence seem to closely follow the optical flow pattern: individual scan patterns suggest that participants experiencing high presence follow respectively anticipate the course of the track, whereas participants scoring low on presence distribute their gaze more randomly over the scene. Another finding in line with this conclusion is that during strong presence experiences, the average saccade amplitude and velocity is diminished. However, we have to point out, that inter-individual differences in eye movements are substantial.

A possible answer to the inter-individual differences is the dissociation between the processing of the spatial localization and the object identification in the brain. These so-called ‘‘where” and ‘‘what” systems are anatomically associated with the dorsal respectively the ventral streams in the cortex [14]. Correspondingly, functional imaging studies found motion cues to increase the activation of the dorsal stream [15] and object features such as color to activate the ventral stream [16][17][18].

In dynamic environments such as the virtual roller coaster, spatial presence is associated with increased activity the parietal lobe regions, which in the first place to mediate spatial localization. In less dynamic environments spatial cues such as shadowing or object motion were found more important than object cues such as textures or geometric detail [19]. Accordingly, through an EEG study, it was found that high spatial presence activated the parietal lobe regions in the brain, which mainly mediate spatial localization [20].

The virtual roller coaster seems to elicit spatial presence in the first place through spatial cues such as optical flow. These seem to function quite well as the high overall ratings of self location and attention allocation suggest. As expected, some individuals did not experience spatial presence or scored low on spatial presence. These individuals seem to inspect more frequently objects surrounding the ride. They score low on spatial presence since they miss the spatial cues inducing motion and focus on objects such as two-dimensional trees. This viewing behavior could be both basis and result of failing to suspend the disbelief and accept the displayed environment as real.

Conclusions

We found differences in viewing behavior depending on the extent of spatial presence experienced. High spatial presence was associated with more fixations, shorter fixations, smaller saccades amplitudes, and decreased saccade velocity. Our findings underline the importance of attentional processes in spatial presence. In this dynamic environment, spatial cues (ie, optical flow) seem to form the basis for spatial presence experiences.

Acknowledgements

We want to thank our research assistants Eva Siegenthaler and Matthias Hartmann for their support in the conduct of this study.

References

[1] Lombard, M., & Ditton, T. B. (1997). At the heart of it all: The concept of presence. *Journal of Computer-mediated communication, 3*(2). Available at http://209.130.1.169/ jcmc/vol3/issue2/lombard.html.

[2] Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors, 40*(3), 354-375.

[3] Sadowsky, W., & Stanney, K. (2002). Measuring and managing presence in virtual environments, in K. M. Stanney (Ed.), *Handbook of Virtual Environments Technology*, Lawrence Erlbaum Associates: Hillsdale, New Jersey.

[4] Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections on memory and persuasion. *Journal of Computer Mediated Communication, 3*(2). Available at http://jcmc.indiana.edu/vol3/issue2/kim.html

[5] Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42*, 72-92.

[6] Biocca, F. (1997). The cyborg's dilemma: progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication*, *3*(2). Available: http://www.ascusc.org/jcmc/vol3/ issue2/biocca2.html.

[7] Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., et al. (2007). A process model of the formation of Spatial Presence experiences. *Media Psychology, 9*, 493-525.

[8] James, W. (1890). *The Principles of Psychology*. New York: Holt

[9] Henderson, J. M. (2003). Human gaze control during-real world scene perception. *Trends in Cognitive Sciences, 7*(11), 498-504.

[10] Bailenson, J. N. & Yee, N. (2005). Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16*, 814-819.

[11] URL: www.nolimitscoaster.com/

[12] URL: www.sr-research.com/

[13] Vorderer, P., Wirth, W., Saari, T., Gouveia, F. R., Biocca, F., et al. (2003). *Constructing Presence: Towards a two-level model of the formation of Spatial Presence.* Unpublished report to the European Community, Project Presence: MEC (IST-2001–37661). Hannover, Munich, Helsinki, Porto, Zurich.

[14] Mishkin, M., Unterleider, L. G., Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences, 6*, 414-417.

[15] DeAngelis, G., & Newsome, W. (1999). Organization of disparity-selective neurons in macaque area MT of the macaque. *Journal of Neuroscience 19*, 1398-1415.

[16] Dubner, R., & Zeki, S., (1971). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research, 35*, 528-532.

[17] Kayaert, G., Biederman, I., de Beeck, H. P., & Vongels, R., (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *European Journal of Neuroscience, 22*, 212-224.

[18] Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: position-specific tuning for boundary confirmation. *Neurophysiology, 86*, 2505-2519.

[19] Lee, S., & Kim, G. J. (2008). Effects of visual cues and sustained attention on spatial presence in virtual environments based on spatial and object distinction. *Interacting with Computers, 20*(4), 491-502.

[20] Baumgartner, T., Valko, L., Esslen, M., & Jäncke, L., (2006). Neural correlate of spatial presence in an arousing and noninteractive virtual reality: an EEG and psychophysiology study. *CyberPsychology & Behavior 9*, 30-45.