Head Turn Scaling Below the Threshold of Perception in Immersive Virtual Environments

Martin Westhoven[1], Dennis Paul[2] and Thomas Alexander[1]

[1]Human Factors Department, Fraunhofer FKIE
{*martin.westhoven, thomas.alexander*}*@fkie.fraunhofer.de*
[2]Institute for Psychology, Leuphana University Luneburg
Dennis.Paul@mailbox.tu-dresden.de

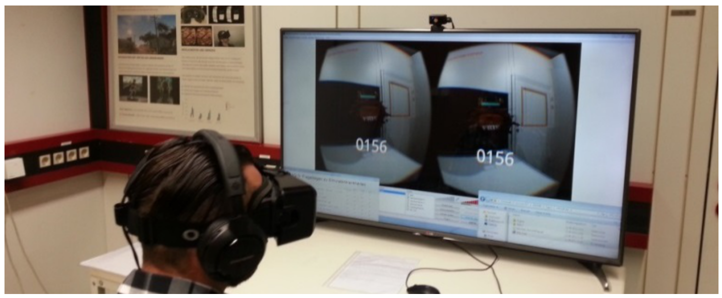

Figure 1: *Participant of the pretest for the empirical study.*

Abstract

Immersive virtual environments allow to experience presence, the feeling of being present in a virtual environment. When accessing virtual reality with virtual reality goggles, head tracking is used to update the virtual viewpoint according to the user's head movement. While typically used unmodified, the extent to which the virtual viewpoint follows the real head motion can be scaled. In this paper, the effect of scaling below the threshold of perception on presence during a target acquisition task was studied. It was assumed, that presence is reduced when head motion is scaled. No effect on presence, simulator sickness and performance was found. A significant effect on physical task load was found. The results yield information for further work and for the required verification of the used concept of presence. It can be assumed, that load can be modified by the scaling without significantly influencing the quality of presence.

Keywords: presence, perception, head tracking manipulation, empirical study, virtual reality, immersive virtual environments

Concepts: •Human-centered computing → Empirical studies in HCI; •Computing methodologies → Perception; Virtual reality;

VRST '16,, November 02 - 04, 2016, Garching bei München, Germany
ISBN: 978-1-4503-4491-3/16/11
DOI: http://dx.doi.org/10.1145/2993369.2993385

1 Introduction

Virtual Reality (VR) is not just a purely technical field, but rather a human-machine interface technology with significant effects on psychological experience. When modifying software, such factors have to be taken into account in the sense of a qualitative user experience for human computer interaction. A central concept is the so-called presence, the physical feeling of being in this unreal, but perceived as real, environment. The concept of presence is increasingly focused by scientists to better understand the psychological mechanisms behind virtual experiences. Most research aims at finding ergonomic designs for VR experiences and psychologic or explicitly psychiatric or psychotherapeutic applications of the technology. The questions why and to which extent presence occurs, what it is composed of and what effects it has are the motivation for this work.
Background is the idea of a specific modification in immersive virtual environments: A scaling of head rotation, which alters the translation of real head movement to its virtual counterpart. To be able to look around is central for users of virtual realities to feel present in the virtual environment instead of just looking upon it from outside. This work focuses on the effects of the modification of this basic interaction below the threshold of perception. For now, the modifications are chosen to be imperceptible to not alert users to the change in head-movement behaviour and thus add a possible break in immersion. From an application point of view, effects on simulator sickness symptoms, experienced task load and performance have to be considered. In terms of usability, modifying the user interface should not lead to reduced presence, which is after all a characteristic feature of immersive VR and could therefore significantly alter the quality of the experience. After a brief summary of related work focusing on modifications of virtual environments and effects on perception, an experiment about effects of scaling head rotations is presented, its results are discussed and finally a summary and an outlook on further work are given.

2 Background and Related Work

Head rotation is a necessary interaction when using (VR) head-mounted displays (HMDs), since commercially available VR-HMDs reach fields of view (FOVs) around 100°horizontally (e.g. Oculus Rift or HTC Vive), whereas the human FOV is about 200°horizontally and 125°vertically [Arthur 2000]. Head rotation incorporates a translation as well, since the camera position is offset from the rotational axis [Yao et al. 2014]. The head can normally be rotated around 90°to each side [Dvořák and Baumgartner 1997]. This rotation is normally translated one-to-one into the virtual world, but modifications can be used to vary the necessary head movement. Especially training or medical applications come to mind as application areas, but also teleoperation, where operator load could possibly be reduced by this.

2.1 Perception of Head-Movement

[Jerald et al. 2012] published several studies about the perception of the virtual scenes during head rotations. Scene movement can occur through intentional manipulation or through insufficient hard- and software. An example for the latter is latency, which remains a problem even with recent developments. Normally, the scene on the retina rotates in the opposite direction of the eye movement, which includes head movement. With VR HMDs, the scene usually only changes through head movement. If it changes with appropriate speed, the scene appears to be stable. By scaling the rotation, the scene optically appears to rotate with (boosting) or against (dampening) the head rotation; the scene moves slower or faster than the user moves his head. For their psychophysical experiments, [Jerald et al. 2012] constructed a latency free HMD to study synthetically added scene movement. [Jerald et al. 2008] found out, that during quasi-sinus-form head rotation and in a controlled visual scene the threshold of perception is dependent on the direction of scene movement relative to the head. Rotating in direction of the physical rotation the threshold was twice as high (11.2%) compared to rotating against the physical rotation (5.2%). Noteworthy is a report of [Jaekl et al. 2005], who state that a scene seems to be more stable when it rotates against the physical rotation. [Jerald et al. 2012] explain the difference to their results with latency in the HMD used by [Jaekl et al. 2005] and the fact, that the eye position is offset forwards from the rotational axis. Another experiment [Jerald et al. 2009] with simple head rotations and constant speed of the scene which moves against the physical rotational direction showed, that perception thresholds become larger with increasing head movement.
In so-called redirected walking experiments, the manipulation of rotation serves to manipulate users in the virtual environment. Redirected Walking is a method to bring natural walking in large virtual environments into smaller real locations. Optimally, it is possible to walk infinitely in the virtual world while never leaving the physical location. Manipulation of the visual input below the perception threshold influences the walking behavior of the user, who then unconsciously adapts his path. This manipulation takes the form of synthetic scene movement comparable with the head rotation experiments described above. Head and torso movement were highly correlated, since the participants moved their whole body to navigate the virtual environment [Razzaque et al. 2001; Steinicke et al. 2008].
[Steinicke et al. 2008; Steinicke et al. 2010] measured perception thresholds for Redirected Walking, which were much higher than with head rotation only: Participants could be made to rotate 68% / 49% more or 10% / 20% less than they perceived virtually. The visual content does not seem to influence the technique. A strong visual flow however heightened the sensitivity towards the manipulation. [Chen et al. 2014] investigated perception thresholds for modified head rotations. Background was kinesophobia treatment, a condition where patients reduce head rotation out of fear of their chronical neck pain. The reported thresholds were 10.7% for dampening and 13.7% for boosting rotations.

2.2 Possible Effects of Head-Movement Modifications

Modifying the head rotation for narrow FOVs (49.6°horizontal) can result in performance improvements for search tasks [Jay and Hubbold 2003]. With increasing FOV, performance does not drop [Kopper et al. 2011], but an interaction trend indicates, that modifying the rotation only increases performance for smaller FOVs. In a flight simulator study, [Le Ngoc and Kalawsky 2013] concluded that head rotation was found to be comparable with extending the display.
The experienced presence depends on the attributes of virtual environments, which can lead to different experiences. Several theories try to explain this relation, e.g. the theories of the so-called Igroup [Regenbrecht 1999; Schubert 2003; Schubert 2009; Schubert et al. 2001] and of [Slater et al. 1995]. Even though presence is used as a variable of user experience [Schubert et al. 2001], there are also influencing factors from outside the construct itself. The relation between technology and the influencing factors is explained by the Potential Action Coding Theory of presence (PACT) [Schubert 2003; Schubert et al. 2001]: Presence is the result of immersion, but not exclusively. Rather, there are cognitive processes involved which create a mental model of the virtual environment under the influence of the virtual stimuli. This also includes non-immersive stimuli, like normal 3D applications or even text-based applications. The interpretation of this mental representation of the virtual environment would then evoke the feeling of presence. Immersion is a technical component in the process of creating presence. [Slater et al. 1995] state, that immersion is the amount of congruence between sensory input with the subjective mental representation and with proprioception. Presence would then be a function of these congruences. While this work aims to investigate the influence of virtual stimuli on presence, [Hofmann 2013] studied the influence of presence on spatial perception. It was shown, that the state of presence can influence the estimation of virtual parameters and thus distort the spatial perception.
In their study about virtual movement, [Slater et al. 1995] compared in-place walking with walking by mouse interaction. Higher presence was found for the more natural in-place walking, which hints at a connection between proprioception and visual sense for the creation of presence. [Usoh et al. 1999] added real walking and verified the earlier results. The differences of scaled head tracking are far smaller which leads to the question, whether the relation between proprioception and visual perception is equally important in this case. Scaling head rotations should have an effect on the consistency of multimedia information [Witmer and Singer 1998], since there is a discrepancy in this regard. However, modifications can also counter a previously existent discrepancy, as observed by [Jaekl et al. 2005] regarding scene stability. The sensory aspect is part of the quality aspect of immersion according to [Schubert et al. 2001], which influences involvedness, the drawing of attention towards the virtual environment. Discrepancies could also influence the expectation factor [Witmer and Singer 1998], which finds its equivalent in predictability and interaction in the work of [Schubert et al. 2001]. Apart from involvedness, spatial presence is affected as well and thus, different scalings of head rotation could also influence the overall presence.
There are no conclusive results, but it is assumed, that presence and performance are positively correlated. This could possibly be caused by the many performance factors, which also raise presence [Schuemie et al. 2001; Witmer and Singer 1998]. E.g. [Slater et al. 1996] postulate, that higher presence leads to more realistic behav-

ior which may lead to higher performance. Complex tasks however could even suffer from high presence, since tasks with an information density, which surpasses the daily routine, profit from abstraction, which in turn could come into conflict with high presence [Regenbrecht 1999].

Studies of [Riley and Kaber 1999; Riley and Kaber 2001] report a negative correlation between telepresence and mental workload, since high cognitive load can lead to reduced involvedness. However, a base load seems to be required, as the constructs were also shown to correlate positively [Ma and Kaber 2006]. The relation between presence and Simulator Sickness is an open question and is in need of further research. Studies so far provide mixed results [Schuemie et al. 2001], although a negative relation seems plausible as sickness symptoms draw on the attention of the user [Witmer and Singer 1998]. Head rotations are reported to increase sickness symptoms [Lackner 1990]. It is noteworthy though, that this is often coupled with forward movement, since head rotations tend to increase vection, the illusion of self movement while observing a moving scene [Kim et al. 2015]. Also, forward motion in combination with head rotation along another axis can lead to coriolis and pseudo-coriolis effects, which can lead to intra-vestibular unbalance and therefore boost visually induced motion sickness [Keshavarz et al. 2014]. Visual flow and especially acceleration is a main factor of Simulator Sickness [LaViola 2000; Yao et al. 2014]. Head rotation also creates an acceleration, but since the virtual movement stems from the tracked real head movement, the discrepancy between visual and vestibular system should be low. Modifying the rotation to a greater extent could lead to increased Simulator Sickness symptoms, though. However, experiments on redirected walking did not result in increased symptoms [Razzaque et al. 2001].

While much work focused on the perception of head movement scaling, this work focuses on the effect of manipulating the scaling ratio on the experienced presence.

3 Method

3.1 Leading Question and Hypotheses

The leading question was if scaling of transmitted head rotation data has an influence on presence. The independent variable was therefore the rotation factor F_R. This factor was originally used in the context of measuring perception thresholds for redirected walking [Steinicke et al. 2010; Steinicke et al. 2008] and is defined as the ratio of the virtual to the real rotation of the head $F_R = R_{virtual}/R_{real}$. In other words, F_R defines to which extent the rotation is modified with $R_{real} = \alpha$ and $R_{virtual} = \alpha \cdot F_R$. Rotation around the upward pointing axis (horizontal rotation) with different scaling factors results in roughly three different behaviors of the changing viewpoint: For $F_R = 1$ the virtual rotation is stable relative to reality. For $F_R < 1$ the virtual scene optically moves in direction of the head rotation and for $F_R > 1$ in the opposite direction. Altering the speed of rotation of the FOV directly influences the perceived scene rotation relative to the head rotation. The rotation factors at the perception thresholds of comparable studies were at 0.59 / 0.68 and 1.1 / 1.24 for full body rotations [Steinicke et al. 2010; Steinicke et al. 2008] and 0.89 and 1.05 for scene rotation only [Jerald et al. 2008]. [Chen et al. 2014] report mean differential thresholds at 0.89 and 1.14. A threshold for scene stability is reported one-sided by [Jaekl et al. 2005] with 1.26. For search tasks, $F_R = 2$ is reported to yield better performance [Jay and Hubbold 2003].

It is assumed, that the perception threshold for full body rotations is higher due to the correlation of head and body rotation. Since no overall trend for the direction of asymmetric thresholds can be identified, the up and down scaling factors were chosen to be of equal proportions. Boosting or dampening of around 10% seems to be below the perception threshold of most people, corresponding to $F_R = 0.9$ and $F_R = 1.1$. Based upon those considerations, one main hypothesis (1) regarding presence and three side hypotheses (2), (3) and (4) regarding additional data are formulated:

Scaling of the virtual head rotation by the proposed factors...

(1) ... significantly decreases presence (significantly lower compared to $F_R = 1$).

(2) ... evokes stronger symptoms of simulator sickness (significantly higher compared to $F_R = 1$).

(3) ... increases load when dampening ($F_R < 1$; significantly higher compared to $F_R = 1$) and decreases load when boosting ($F_R > 1$; significantly higher compared to $F_R = 1$).

(4) ... decreases task performance when dampening ($F_R < 1$; significantly lower compared to $F_R = 1$) and increases task performance when boosting ($F_R > 1$; significantly higher compared to $F_R = 1$).

3.2 Apparatus

The virtual environment ran on an HP Z420 workstation computer with a Nvidia GeForce GTX 780 graphics card and an Oculus Rift DK2 as a VR-HMD. The full HD OLED display has a FOV of 100°and a resolution of 960 x 1080 pixels per eye. Head-tracking is implemented via infrared-LEDs on the HMD which are tracked by a camera and inertial sensors. Visual acuity as well as color and stereoscopic vision were checked with a vision screening instrument (Optovist, Vistec AG). The Igroup Presence Questionnaire (IPQ) [Schubert 2003; igroup 2008] was filled out manually, while the Simulator Sickness Questionnaire (SSQ) [Kennedy et al. 1993] and the NASA Task Load Index (NASA-TLX) [Hart and Staveland 1988; Hart 2006] were filled out in digital. All questionnaires were in German. The IPQ is of German origin, translations were used for the latter two. German translations of the SSQ and the NASA-TLX have been successfully applied before, e.g. the SSQ in [Mehlitz 2004] or the NASA-TLX in [Sepehr 1988].

The virtual environment was created using Unreal Engine 4.7.6 (Epic Games). The virtual environment was a virtual model of the real-world laboratory. The participants were seated roughly where they were sitting in the real laboratory. Rotating the head includes a slight translation, since the rotational axis is behind the center point of the eyes. To avoid inducing disparities, translation was therefore scaled as well. The scaling of both rotation and translation was implemented by replacing the distance between current and last frame with its respective scaled version. Ambient sound was played over headphones to eliminate background noise, e.g. originiating from people passing in the hallway in front of the lab. The logged data was the total time-on-target in ms, the number of contact breaks, the total head rotation in degree and the precise points in time for target locks and breaking contact again. Also, there were several questions regarding the head rotation. The first item was, whether the rotation felt natural or unnatural. For each treatment an additional item of the questionnaire was, whether a modification of the head tracking was felt.

3.3 Experiment Task and Design

The experimental task was to keep a hovering robot in focus while it moved horizontally in front of the participant on a circular path of 150°in a distance of about 1m. The virtual scene is depicted in figure 2.

Figure 2: *Sample scene as presented to the participants.*

For each experimental treatment, a different movement pattern consisting of 135 randomly generated pairs of movement distances and speeds along the path was used. The pattern was the same for each participant in a certain treatment, only the experimental condition was varied. The robot changed directions in a randomly determined pattern to make the moment of changing direction hard to predict. Also the speed was varied between each change of direction. A virtual laser ray was rendered to mark the viewfield's center. When the robot was in this center, a display in its upper region showed a blue-green color to indicate that it is being held in focus correctly, else it showed a red-orange. The participants had to 'charge' the robot by directing the laser to it. The longer the laser was held on top of the robot, the more points were earned. The overall score was displayed slightly below the center of the view to keep it visible at all times.
The independent variable, represented through the scaling factor F_R, was varied over three degrees. The first is a control category (1) with $F_R = 1$. The second category (2) is a dampening with $F_R = 0.9$ and the third (3) a boosting with $F_R = 1.1$. Dependent variables are primarily Presence, measured via IPQ, and secondarily simulator sickness via SSQ, load via NASA-TLX, and performance via overall time on target. A within-design was used. Each participant performed the task under the three conditions in a randomly generated order and on different days to avoid training and sequence effects.

3.4 Experiment Execution

The experiment was completed by $N = 13$ employees of a German research institute, one of them female and twelve male. The age ranged from 19 to 45 years with a mean age of $M = 30.31$ (Standard deviation $SD = 8.43$). No payment was offered, but a ticket for a fair was offered as an incentive for the best performing participant. Participants who were pregnant, had insufficient knowledge of the German language or reported medical problems with their necks were not included. The experiment had to be cancelled for a second female participant due to lacking language skills. Since presence does not seem to be influenced by gender, the remaining female participant was included for statistical analysis. The experiments took place from 9:30 to 15:30 local time. 30 minutes were planned for each treatment. In practice, the first treatment took about 25, the second about 15 and the third about 20 minutes. Preceeding the first treatment, a binocular vision screening was performed to measure color and stereoscopic vision and the visual acuity for corrected sight, meaning prescription glasses and the like were left on if they were to be used with the VR-goggles as well. Additionally the participants were asked for information on age, sex, vision impairments, attention related impairments (e.g. ADHD), weekly usage of first person 3D-Applications, experience with immersive virtual environments and to fill out the SSQ.
Throughout the experiments, the same examiner was present and stood about 1-2 meters diagonally behind the participant. The participants sat on a non-revolving chair to make sure they performed the rotation mainly with their heads. When they were set up, the virtual scene faded in and the robot moved into the front of the participant. After they confirmed that they were ready, the experiment started. After about five minutes, the experiment ended and the virtual scene faded out. The virtual view of the participants was shown on a regular screen as well, so the examiner could check if the participants were following their set task. In order to not influence the participants early, questions specifically regarding the head rotation were asked only after all experimental treatments were finished.

4 Results

All thirteen participants were able to complete the task in the virtual environment. The visual acuity measured during the vision screening ranged from 0.8 to 1.25 for binocular far sight. Four participants recognized all six color tables, the rest recognized five. All had stereoscopic vision. None reported ADHD. The data analysis was done using an analysis of variance (ANOVA) with repeated measurements. Using MANOVA with repeated measurements when analyzing multiple dependent variables is useful to minimize the first order error and to take interdependencies between them into account [Field 2009]. Since the variables correlated only minimally, thus violating an important prerequisite, multiple univariate ANOVAs were performed instead of a MANOVA.

4.1 Scaling of Head Rotation

Six participants reported that looking around was unusual, but did not directly relate to the scaling itself. Instead a too high physical effort, a missing reflection in a window to a neighbouring room in the virtual environment and a flickering of the target when it changed direction were reported. Single reports were also given on

- recognition of the target from a game (design inspiration)
- on treatment (2) being more pixellated
- on treatment (2) being more blurry after treatment (3)
- on the laser pointer being more inaccurate in treatment (1)

Four participants reported a recognition of a modification to the translation of real to virtual head movement in all three treatments. Two started with treatment (1), the others with treatment (2). Three of them reported unnatural movement, citing an unnatural movement necessary to follow the target. A further participant reported a modification in his last treatment, which was the control condition.

4.2 Presence

Following the Igroup, a so-called presence profile can be established, which allows the descriptive comparison of the three treatments considering their presence components [igroup 2008]. The three components are involvedness, spatial presence and reality judgement. All components together yield a general presence value. The values between the components are highest for general and spatial presence, followed by involvedness and degree of reality (see table 1). There seem to be no internal variations for the components with respect to different scaling factors. In general the up-scaling treatment resulted in a lower presence value with more deviations. Compared to the mean of 619 cases from the Igroup database, all components of presence scored higher than the mean value. Interestingly, the ratio of the component values for the three treatments

was lower in the involvement score compared to all others.
Reports of a feeling of unnaturalness correlate significantly negative with presence scores of the experimental treatments (1) and (3) (Bravais-Pearson-correlation) ((1): $r(11) = -.61$, $p = .025$; (3): $r(11) = -.71$, $p = .007$). With treatment (2), there is a small, non-significant dependency ($r(11) = -.50$, $p = .085$). Participants who noticed something unnatural thus also experienced lower presence. For further analysis, the overall presence score is used, which is the sum of all 14 IPQ items' scores ($0 - 84$). The box-whisker-plots show distribution and interval of the data for each treatment (Figure 3).

Table 1: *Overview of the presence components.*

Dependent Variable	Stage	Mean	Standard Deviation
General Presence	(1) One-to-one	4.31	0.95
	(2) Dampening	4.31	1.18
	(3) Boosting	3.69	1.70
Spatial Presence	(1) One-to-one	4.55	0.74
	(2) Dampening	4.52	1.00
	(3) Boosting	4.38	0.93
Involvedness	(1) One-to-one	3.35	0.72
	(2) Dampening	3.35	0.77
	(3) Boosting	3.21	1.14
Reality Judgement	(1) One-to-one	2.88	0.85
	(2) Dampening	2.96	0.98
	(3) Boosting	2.81	1.06

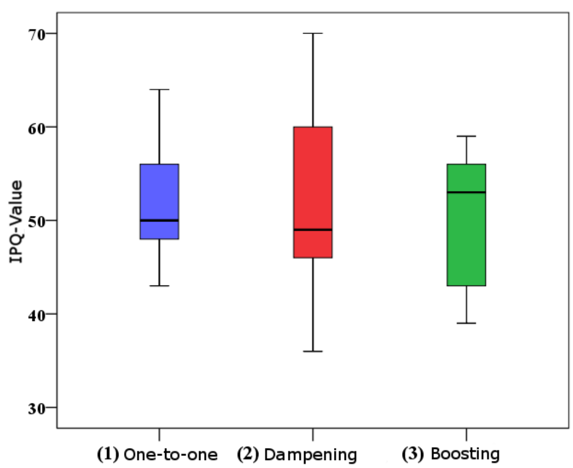

Figure 3: *Boxplots (Median, quartile, outliers) for presence (IPQ).*

The scores seem to be comparable, only level (2) shows a higher variance. The presence histograms of the treatments (1) (Figure 4) and (2) (Figure 5) depict the frequency distribution. They seem to be approximately normally distributed, in contrast to the histogram of treatment (3) (Figure 6).

Treatments (1) and (2) show a slight positive skewness S and a high negative kurtosis. Treatment (3) is slightly negatively skewed and rather flat-topped. The z-values of both characteristic values ($Z = (S0)/SE$) are each smaller than 1.96, which indicates non-significant skewness and kurtosis with $p < .05$. The mean values for the different treatments differ only minimally between treatment (1) and (2). The presence for an up-scaled rotation is lower than with both other conditions ($M = 49.69$), while the median is greater ($Mdn = 53$). Down-scaling of the rotation shows the greatest variance of all treatments ($SD = 10.03$) (Table 2). A precondition for an ANOVA is the normal distribution of the measured

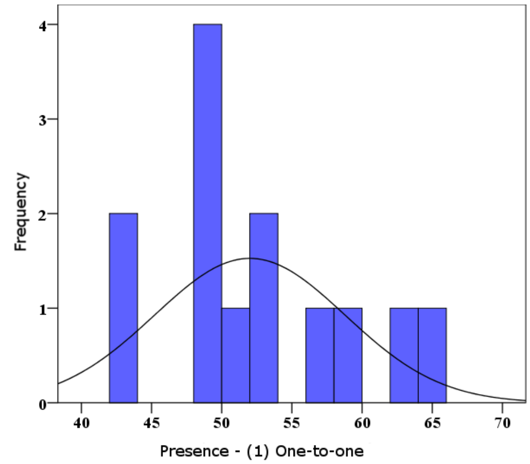

Figure 4: *Presence histogram for treatment (1).*

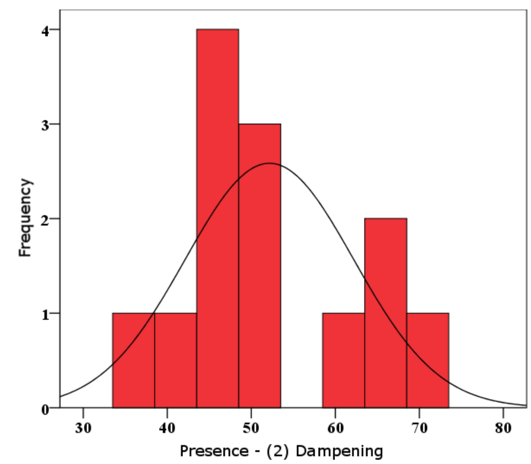

Figure 5: *Presence histogram for treatment (2).*

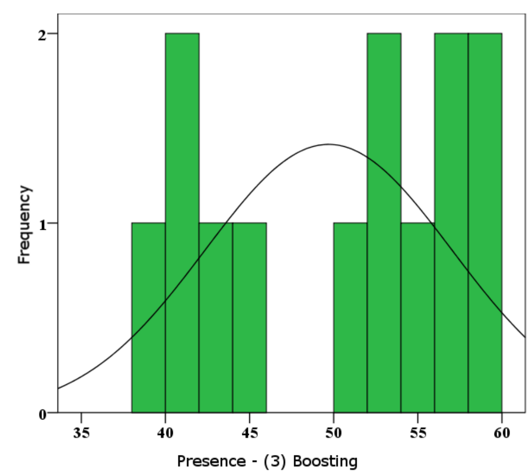

Figure 6: *Presence histogram for treatment (3).*

data in each treatment. Shapiro-Wilk-Tests for normal distribution were non-significant for each treatment. Thus, a normal distribution can be assumed. The p-value of treatment three is only minimally larger than the significance level of .05, which relates to the

very large kurtosis (Table 3). A two-factor analysis of variance on one factor was performed, with the scaling as a repeatedly measured inner subject factor and taking notice of the scaling as a non-repeated between-subjects factor. Studying the effect of the scaling tests hypothesis (1). The descriptive statistic shows that treatment (2) does not induce lower presence. No hypothesis was formed a priori for the interaction between scaling factors and noticing the scaling, since it was not clear, whether the participants would notice the modification at all. Nonetheless, it is interesting to see, whether the participants who noticed the modification had a different presence experience. Another precondition for performing an ANOVA with reaped measurements is sphericity, the homogeneity of the variances of all treatments and of the covariances of all pairwise combinations. The Mauchly-test checks the hypothesis $H0$ of sphericity being present (Table 4).
The probability of the test statistic is non-significant with .21, which implies $H0$ cannot be rejected and the F-value of the ANOVA is interpretable. The p-value of the scaling is non-significant ($F(2, 22) = 0.89$, $p = .214$), therefore a significant difference between the factors treatments cannot be assumed. The approximated effect (partial eta-squared) is $\eta^2 = .074$. The factor scaling thus explains 7.4% of the variance of presence. The test power of the observed power is $1 - \beta = .183$. The probability to find an empirical effect of $\eta^2 = .074$ under the given circumstances was 18.3%. The interaction of scaling and noticing the scaling is also non-significant ($F(2, 22) = 1.46$, $p = .127$) with $\eta^2 = .117$ and a power of 27.9% (Table 5).

Table 2: *Descriptive Statistics for Presence (IPQ).*

	(1) One-to-one	(2) Dampening	(3) Boosting
Mean	52.00	52.15	49.69
Mean SE	1.88	2.78	2.03
Median	50	49	53
SD	6.79	10.03	7.33
Variance	46.16	100.64	53.73
Skewness	0.56	0.45	−0.35
Skew. SE	0.62	0.62	0.62
Kurtosis	−0.56	−0.63	−1.58
Kurt. SE	1.19	1.19	1.19

Table 3: *Shapiro-Wilk-Test for normal distribution of presence.*

	Statistic	df	Significance
Presence - (1) One-to-one	.925	13	.290
Presence - (2) Dampening	.937	13	.418
Presence - (3) Boosting	.888	13	.092

Table 4: *Mauchly-Test for sphericity of presence.*

Inner-subject effect	Mauchly-W	Approximated χ^2	df	Sig.
Scaling	.732	3.12	2	.210

Accordingly, there is no evidence of interactions between the factors. Also, no between-subjects effect for noticing the scaling can be assumed due to non-significant results ($F(1, 11) = 0.12$, $p = .367$, $\eta^2 = .011$, $1 - \beta = .062$) (Table 6). A test for simple inner-subject contrasts compared treatments (2) and (3) with the control condition (Table 7). There was only one significant F-value found for the interaction of both factors ($F(1, 11) = 4.58$, $p = .028$) with $\eta^2 = .294$ and a test power of 49.7%. The profile diagram (Figure 7) depicts the interactions graphically. Both groups

Table 5: *Test for inner-subject effects of presence.*

Source	df	F	Sig.	Partial η^2	Observed Power
Scaling	2	0.89	.214	0.74	.183
Scaling * Notice	2	1.45	.127	0.117	.279

show the lowest presence for boosted rotation. For both dampening and boosting, presence is on a comparable level. Participants who noticed a modification showed a lower presence with one-to-one scaling. Note, that also in this configuration, subjects reported noticing a modification.

Overall there were no interdependencies between presence and other dependent variables. Only the negative correlation coefficients of the interdependency between presence and both measurements of simulator sickness were significant for treatment (1) ($r(11) = -.72$, $p = .002$ respectively $r(11) = .56$, $p = .046$). Variance analysis with repeated measurements for presence components yielded no significant results.

Table 6: *Test for between-subject effects of presence.*

Source	df	F	Sig.	Partial η^2	Observed Power
Constant term	1	484.60	.000	.978	1.000
Notice	1	0.12	.367	.011	.062

Table 7: *Test for inner-subject contrasts of presence.*

Source and Scale	df	F	Sig.	Partial η^2	Observed Power
Scaling lin.	1	0.95	.176	.079	.145
Scaling quad.	1	0.86	.187	.073	.136
Scaling * Notice linear	1	4.58	.028	.294	.497
Scaling * Notice quadratic	1	0.32	.292	.028	.081

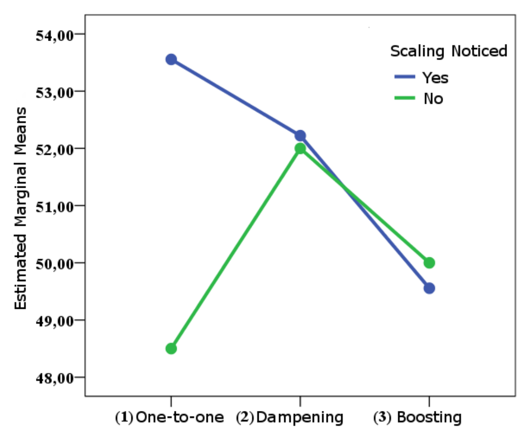

Figure 7: *Profile diagram for presence.*

4.3 Simulator Sickness

Figure 8 shows the SSQ scores for both measurements for all factor treatments. A difference for pre- and post-exposition measurement can be seen. Five single values of two participants are outliers. Since they stem from pre-exposition questioning, a poor shape on that day can be assumed. Both datasets are therefore excluded from further analysis, since otherwise the precondition of normal distribution for variance analysis cannot be assumed. The highest values were found for one-to-one scaling and the lowest were found for boosting (Table 8). Generally, symptoms of the oculomotorics and disorientation groups occurred the most. Due to mostly lower values, the distributions are slightly positively skewed. Nonetheless, the Shapiro-Wilk-Test shows normal distributions for all measurements except the first measurement of treatment (1), where the test is significant ($p = .013$).

A two-factorial analysis of variance with repeated measurements was performed on both factors with $n = 11$. It was planned for checking hypothesis (2), but already the mean values are descriptively lower than for treatment one. The scaling has three degrees and time of measurement further two. The Mauchly-Test for sphericity shows homogenous variances for the factor degrees and their correlations. No significant differences could be found between the factor degrees ($F(2, 20) = 0.81$, $p = .230$) with effect size $\eta^2 = .075$ and power of .168. A significant difference was found for the time of measurement ($F(1, 10) = 8.78$, $p = .007$) with $\eta^2 = .468$ and a power of .761.

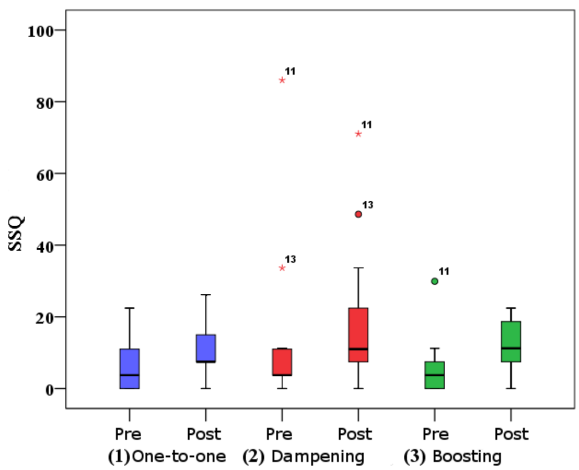

Figure 8: *Boxplots (Median, quartile, outliers) for Simulator Sickness (SSQ).*

Table 8: *Descriptive statistics for Simulator Sickness.*

	(1)		(2)		(3)	
	Pre	**Post**	**Pre**	**Post**	**Pre**	**Post**
Mean	7.53	11.59	4.74	11.52	3.74	10.54
SE Mean	2.60	2.74	1.23	2.92	1.13	1.94
Median	3.74	7.48	3.74	11.00	3.74	11.22
SD	8.61	9.10	4.09	9.68	3.74	6.44
Variance	74.10	82.86	16.76	93.63	13.99	41.45
Skewness	.86	.53	.42	1.21	.73	.34
SE Skew.	.66	.66	.66	.66	.66	.66
Kurtosis	-.92	-.67	-.95	1.91	-.13	.06
SE Kurt.	1.28	1.28	1.28	1.28	1.28	1.28

4.4 Load

Shapiro-Wilk-Tests for normal distribution were non-significant, hinting at normally distributed data, which is a requirement for the univariate ANOVA with repeated measurements. Mauchly tests on sphericity were also non-significant, allowing for the assumption of sphericity. A significant difference ($F(2, 24) = 3.87$, $p = .018$) between the treatments with an effect size of $\eta^2 = .244$ and corrected item-total correlation of 64.3% was found. Testing for simple inner-subject contrasts with the control treatment as the reference category, the quadratic trend becomes significant ($F(1, 12) = 6.17$, $p = .015$) with $\eta^2 = .340$ and $1 - \beta = .627$.

Pairwise post-hoc tests between the treatments with alpha-levels adapted by the Bonferroni-method serve to check hypothesis (3). In comparison to the control treatment, a significant difference was found ($p = .042$). The intervals of the RAW-TLX values are shown as boxplots in figure 9.

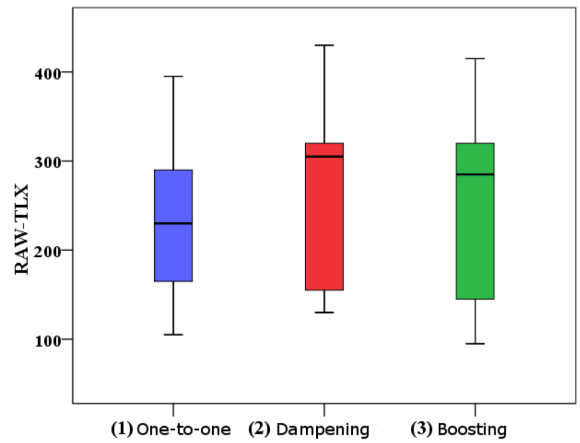

Figure 9: *Boxplots (Median, quartile, outliers) for load (NASA-TLX).*

Higher medians and means can be seen for the modifying treatments, compared to the control treatment (Table 9). Even though no weighting was performed by the participants, univariate analyses of variance with repeated measurements can be performed for the NASA-TLX sub-scales when considering this constraint. Under the assumption of sphericity, a significant result was found for the performance sub-scale ($F(2, 24) = 2.67$, $p = .045$, $\eta^2 = .182$, $1-\beta = .478$) and a trend for the effort sub-scale ($F(2, 24) = 2.17$, $p = .068$, $\eta^2 = .153$, $1 - \beta = .4$).

Individual performance (inverted scale) was rated lowest for the down-scaled condition ($M = 54.23$) and was also lower for up-scaling ($M = 48.46$) compared to one-to-one scaling ($M = 45.38$). Accordingly, a significant quadratic trend ($F(1, 12) = 7.71$, $p = .009$) shows for the inner-subject contrasts with $\eta^2 = .391$ and a power of 72.2%. Similarly, effort is highest for treatment (2) ($M = 44.23$) and higher for treatment (3) ($M = 40.00$) compared to the control treatment (1) ($M = 36.15$) and a significant quadratic trend of the inner-subject contrasts ($F(1, 12) = 3.46$, $p = .044$, $\eta^2 = .224$, $1 - \beta = .402$) was found as well.

4.5 Performance

Figure 10 shows an overview of the distribution of the overall time on target as a measure for performance during the five-minute lasting ($302.12s$) task. For the control category, mean and median

Table 9: *Descriptive Statistics for Load (NASA-TLX).*

	(1) One-to-one	(2) Dampening	(3) Boosting
Mean	230.38	257.69	245.38
Mean SE	27.67	27.37	28.98
Median	230	305	285
SD	99.76	98.69	104.49
Variance	9951.92	9740.06	10918.59

are slightly higher. On average for the three conditions, the target was in focus (1): 86%, (2): 85.1% and (3): 85.4% of the time. Compared to the other treatments, the scattering is relatively low for the boosted condition ($SD = 4.66$) (Table 10). The Shapiro-Wilk-Test ensured that all factor treatments of scaling are normally distributed. The Mauchly-Test preceding the univariate analysis of variance with repeated measurements indicates sphericity. No significant difference could be found between the experimental conditions ($F = 1.36$, $p = .138$, $\eta^2 = .102$, $1 - \beta = .264$). Pairwise post-hoc tests were planned to check hypothesis (4), but were not appropriate due to the non-significant ANOVA.

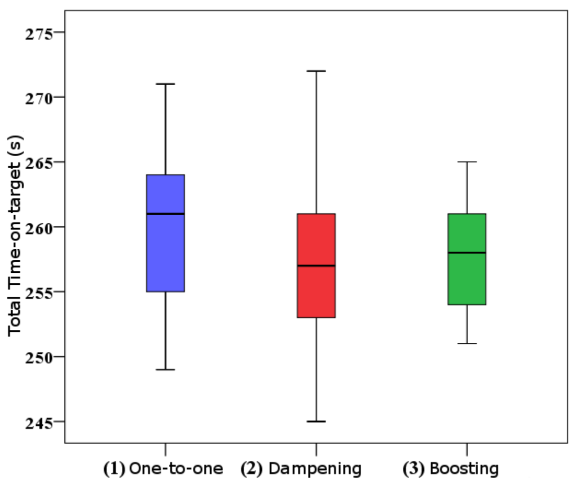

Figure 10: *Boxplots (Median, quartile, outliers) for performance (Total time on target).*

Table 10: *Descriptive Statistics for Performance (Time on target).*

	(1) One-to-one	(2) Dampening	(3) Boosting
Mean	259.85	257.23	257.92
Mean SE	1.84	2.11	1.29
Median	261	257	258
SD	6.626	7.61	4.66
Variance	43.81	57.86	21.74

5 Discussion

No influence of the head rotation scaling on presence in immersive virtual environments and also no significantly different effects between the treatments were found. The hypotheses have to be rejected for the tested range of scaling. In the following, significance and implications of these results are discussed. First, further experiments or replications are needed to fully falsify the hypotheses. According to the theory of optimal sample sizes, $n = 11$ participants are needed to safely find large main effects ($E = .8$) with significance-level $\alpha = .05$ for analysis of variance with repeated measurements. It ensures that an effect is backed up by a test power of $1 - \beta = .8$ and that smaller effects are disregarded. To find a large interaction effect of noticing the modification with the modification itself would require six participants per group, however [Bortz and Döring 2007]. In regards to large effects, this is an argument for an appropriate falsification. Based on literature, scaling factors of $-+10\%$ were chosen, which was expected to be below the perception threshold. Only 31% of the participants noticed the scaling after being asked. This result is consistent with the thresholds of 0.89 und 1.14 by [Chen et al. 2014], who also scaled head rotation in immersive virtual environments. Participants who noticed something unnatural also experienced lower presence. It is possible that the modification was only detectable in hindsight and because thinking about it was triggered by an item of the used questionnaire. This is supported by participants reporting modifications also in the control condition. Chosing scaling factors above the perception threshold could lead to further results on this topic.

Perception theories [Regenbrecht 1999; Schubert 2003; Schubert 2009; Schubert et al. 2001; Slater et al. 1995] explain how discrepancies between proprioception and visual perception can lead to reduced presence. This should also happen, if the conflict is not conscious. In this experiment, the discrepancy seemingly did not suffice to influence the experience of presence. Possibly the above-average presence niveau of the virtual environment influenced this. Further studies are needed to evaluate the influence of the general level of presence on perception of the scaling.

The spatial presence component was expected to be lowered by the scaling, because prediction and interaction would be hampered. That this was not found is interesting and should be investigated further. With reference to secondary dependent variables, simulator sickness was investigated. Scaling of the head rotation did not influence simulator sickness significantly. A non-significant reduction was found compared to the control condition, but this has to be interpreted as a random result. It can be assumed, that modifications of this magnitude can be used without inducing more or more severe sickness symptoms. It was expected, that dampening the head movement would increase load and boosting it would decrease load. This could not be verified. But significant differences were found nonetheless, possibly caused by the increased physical movement needed to get the same results as normally. This is backed up by reports of the participants, who said that they needed to turn their heads unnaturally far.

Boosting the rotation also increased load. Perhaps this can be explained by higher cognitive load to process the discrepant sensory input. No influence on performance was found. The target acquisition task was easy to understand and learn. The target was slow enough, that perhaps the modification was to no effect in this phase. Since a boosted rotation can also have adverse effects when the target abruptly changes direction, there may be no potential for better performance compared to the unmodified condition. Regarding the effect of boosting virtual movement on performance, [Teather and Stuerzlinger 2008] did not find any effects either. They studied boosting of translation in a head-tracked desktop setup where the head position controlled the perspective of a 3D scene. They reported a learning effect on task completion time over time however, which can possibly be attributed to the adaption of the vestibo-ocular reflex (VOR). Future work should therefore also focus on learning effects and possibly include eye-tracking to look for adaption of the VOR to the boosted movement.

The open outcome of the experiment prompts to investigate the method and execution for possible explanations. Repeated measurements allow for a small sample size and can reduce variance between participants. The treatments can however easily be influenced by those already given. The time between treatments varied between less than 24 hours and several days due to constraints in finding fitting appointments for all participants. Both could have a distorting effect on the results. Furthermore, the results cannot be generalized for each condition because the classic preconditions for independent measurements were not fulfilled, although efforts towards randomization and balancing were made [Bortz and Döring 2007]. In a between-subjects design, also the question regarding the perception of the scaling could have been asked directly after the treatment, because no further treatment would have followed. Asking the question several days after exposure is not optimal, since memories can be incomplete or distorted and also be influenced by memories of the other treatments. The results further show, that the question was not optimally formulated. It should have been clearer, that a difference in scaling between virtual and real world was sought after and not a difference between the conditions. Asked as it was, misunderstandings were possible, which can be seen in the reports of detected scaling in the one-to-one condition.

The scaling factors were chosen to be below the perception threshold to not break immersion. An increased range of the scaling factor could have yielded more contrasts and information about the perception threshold, which would have made comparison with those of [Chen et al. 2014] possible.

Usage of a virtual model of the real world laboratory was planned. It has to be noted though, that this could maybe affect the perception of presence. Tests with different virtual environments are needed to investigate influences of the environment and the associations the participants make with it. Fixation of a moving object is different from static objects. If the head rotation was prompted by fixating static points in the virtual room, the attention would not have been so strongly focussed on the target. There is also the question, whether different stimuli would lead to other results. Finally, a narrower movement interval could have guaranteed that participants would not have turned their upper bodies upon reaching their physical limits and therefore detect the unnaturally high movement necessary.

6 Summary and Outlook

Although modification of the head rotation in virtual environments was studied before, its effect on presence is usually not. Possible applications were discussed, where small imperceptible changes could be of benefit. Theories about presence, especially those of the Igroup [Regenbrecht 1999; Schubert 2003; Schubert 2009; Schubert et al. 2001] and of [Slater et al. 1995], indicate, that presence is essentially defined by the coordination of senses, but technical factors can, through unconscious heuristics, directly influence perception. Above a specific strength, discrepancies between the senses, e.g. as induced by modifications of the head rotation, could therefore lead to a reduced feeling of presence. The modifications in this work were accordingly chosen to be below the threshold of perception to not add a possible break in presence. Even though the specific hypothesis had to be rejected, it was shown that the scaling can induce different loads. The technique can thus be used to manipulate the real head rotation as [Chen et al. 2014] intended. A reduction of presence through rotational scaling could not be shown. Should presence indeed not be influenced, reducing the head rotation could be used to induce a higher load for e.g. training or medical purposes. Furthermore, the question why spatial presence is not reduced would need to be answered. A possible explanation could be the relatively small differences between the treatments. Further experiments are necessary to validate or falsify the model of presence proposed by [Schubert et al. 2001].

Also, the factors of immersion and interaction as well as their interactions with presence should be studied further. Another direction to better understand the interaction of presence and proprioception is the study of other scalings, like hand movements. Noteworthy is a theory especially advocated by [Slater 2004], which was not considered in this work: Post-experimental questionnaires could not adequately measure presence, since the question itself could lead to the formation of the phenomenon afterwards. There would be no neuronal or behavioral results to compare with and also in interdependencies with performance, presence could only be seen as an intermediate variable. Since the questionnaires' items are the only possibility to express the users' experiences, the data would not be verifiable. The theory does not question the construct itself, but the measurement method which is used the most. Future research should take this into consideration. Different behavioral responses to experiences in virtual environments as well as neuronal activity, mimics, eye movement and reaction times could give further clues to validate or falsify theories which postulate, that presence is invoked by science itself. Research on presence is still a very young field and the concept needs further development in theory and practical ways. When considering techniques to enhance interaction, like the studied modification of head rotation, from an ergonomic point of view, presence as the "...key to defining VR in terms of human experience ..." [Steuer 1992, p. 5] should not be diminished.

References

ARTHUR, K. W. 2000. *Effects of field of view on performance with head-mounted displays*. PhD thesis, University of North Carolina at Chapel Hill.

BORTZ, J., AND DÖRING, N. 2007. *Forschungsmethoden und Evaluation für Human-und Sozialwissenschaftler: Limitierte Sonderausgabe*. Springer-Verlag.

CHEN, K. B., PONTO, K., SESTO, M. E., AND RADWIN, R. G. 2014. Influence of altered visual feedback on neck movement for a virtual reality rehabilitative system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, 693–697.

DVOŘÁK, J., AND BAUMGARTNER, H. 1997. *Manuelle Medizin - Diagnostik*, 5., überarb. und erw. aufl. ed. Thieme, Stuttgart [u.a.].

FIELD, A. 2009. *Discovering statistics using SPSS*. Sage publications.

HART, S. G., AND STAVELAND, L. E. 1988. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology 52*, 139–183.

HART, S. G. 2006. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50*, 9, 904–908.

HOFMANN, J. 2013. *Raumwahrnehmung in virtuellen Umgebungen: Der Einfluss des Präsenzempfindens in Virtual Reality-Anwendungen für den industriellen Einsatz*. Springer-Verlag.

IGROUP, 2008. Igroup presence questionnaire (ipq).

JAEKL, P. M., JENKIN, M. R., AND HARRIS, L. R. 2005. Perceiving a stable world during active rotational and translational head movements. *Experimental brain research 163*, 3, 388–399.

JAY, C., AND HUBBOLD, R. 2003. Amplifying head movements with head-mounted displays. *Presence: Teleoperators and Virtual Environments 12*, 3, 268–276.

JERALD, J., PECK, T., STEINICKE, F., AND WHITTON, M. 2008. Sensitivity to scene motion for phases of head yaws. In *the 5th symposium*, S. Creem-Regehr and K. Myszkowski, Eds., 155.

JERALD, J., STEINICKE, F., AND WHITTON, M. 2009. Scene-motion thresholds correlate with angular head motions for immersive virtual environments. *ACHI ... the ... International Conference on Advances in Computer-Human Interactions 2009*, 69–74.

JERALD, J., WHITTON, M., AND BROOKS, F. P. 2012. Scene-motion thresholds during head yaw for immersive virtual environments. *ACM transactions on applied perception 9*, 1.

KENNEDY, R. S., LANE, N. E., BERBAUM, K. S., AND LILIENTHAL, M. G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology 3*, 3, 203–220.

KESHAVARZ, B., HETTINGER, L. J., KENNEDY, R. S., AND CAMPOS, J. L. 2014. Demonstrating the potential for dynamic auditory stimulation to contribute to motion sickness. *PloS one 9*, 7, e101016.

KIM, J., CHUNG, C. Y. L., NAKAMURA, S., PALMISANO, S., AND KHUU, S. K. 2015. The oculus rift: a cost-effective tool for studying visual-vestibular interactions in self-motion perception. *Frontiers in psychology 6*, 248.

KOPPER, R., STINSON, C., AND BOWMAN, D. 2011. Towards an understanding of the effects of amplified head rotations. In *The 3rd IEEE VR Workshop on Perceptual Illusions in Virtual Environments*.

LACKNER, J. 1990. Human orientation, adaptation and movement control. *Motion Sickness, Visual Displays and Armoured Vehicle Design*, 28–50.

LAVIOLA, J. J. 2000. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin 32*, 1, 47–56.

LE NGOC, L., AND KALAWSKY, R. S. 2013. Evaluating usability of amplified head rotations on base-to-final turn for flight simulation training devices. In *Virtual Reality (VR), 2013 IEEE*, 51–54.

MA, R., AND KABER, D. B. 2006. Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies 64*, 6, 541–552.

MAY, J. G., AND BADCOCK, D. R. 2002. Vision and virtual environments. *Handbook of Virtual Environments: Design, Implementation, and Applications. Mahwah, NJ: Lawrence Erlbaum Associates, Inc*, 29–64.

MEEHAN, M., RAZZAQUE, S., WHITTON, M. C., AND BROOKS, F. P. 2003. Effect of latency on presence in stressful virtual environments. In *IEEE Virtual Reality 2003*, 141–148.

MEHLITZ, M. 2004. *Aufbau eines medizinischen Virtual Reality-Labors und Entwicklung eines VR-gestützten neuropsychologischen Testsystems mit einer präklinischen und klinischen Evaluationsstudie*. Dissertation, Georg-August-Universität Göttingen, Göttingen, Deutschland.

RAZZAQUE, S., KOHN, Z., AND WHITTON, M. C. 2001. Redirected walking. 105–106.

REGENBRECHT, H. 1999. Faktoren für präsenz in virtueller architektur. *Bauhaus-Universität Weimar, Diss.*

RILEY, J. M., AND KABER, D. B. 1999. The effects of visual display type and navigational aid on performance, presence, and workload in virtual reality training of telerover navigation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 43*, 22, 1251–1255.

RILEY, J. M., AND KABER, D. B. 2001. Utility of situation awareness and attention for describing telepresence experiences in a virtual teleoperation task. In *Proceedings of the International Conference on Computer-Aided Ergonomics and Safety*, 88–93.

SCHUBERT, T., FRIEDMANN, F., AND REGENBRECHT, H. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments 10*, 3, 266–281.

SCHUBERT, T. 2003. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie 15*, 2, 69–71.

SCHUBERT, T. W. 2009. A new conception of spatial presence: Once again, with feeling. *Communication Theory 19*, 2, 161–187.

SCHUEMIE, M. J., VAN DER STRAATEN, P., KRIJN, M., AND VAN DER MAST, CHARLES A.P.G. 2001. Research on presence in virtual reality: A survey. *CyberPsychology & Behavior 4*, 2, 183–201.

SEPEHR, M. M. 1988. Assessment of subjective mental workload using nasa-task load index. In *Proceedings of the 7th European Annual Conference on Human Decision Making and Manual Control*, 69–75.

SLATER, M., USOH, M., AND STEED, A. 1995. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI) 2*, 3, 201–219.

SLATER, M., LINAKIS, V., USOH, M., KOOPER, R., AND STREET, G. 1996. Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In *ACM virtual reality software and technology (VRST)*, 163–172.

SLATER, M. 2004. How colorful was your day? why questionnaires cannot assess presence in virtual environments. *Presence 13*, 4, 484–493.

STEINICKE, F., BRUDER, G., ROPINSKI, T., AND HINRICHS, K. 2008. Moving towards generally applicable redirected walking.

STEINICKE, F., BRUDER, G., JERALD, J., FRENZ, H., AND LAPPE, M. 2010. Estimation of detection thresholds for redirected walking techniques. *IEEE transactions on visualization and computer graphics 16*, 1, 17–27.

STEUER, J. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of communication 42*, 4, 73–93.

TEATHER, R. J., AND STUERZLINGER, W. 2008. Exaggerated head motions for game viewpoint control. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, ACM, 240–243.

USOH, M., ARTHUR, K., WHITTON, M. C., BASTOS, R., STEED, A., SLATER, M., AND BROOKS JR, F. P. 1999. Walking > walking-in-place > flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 359–364.

WITMER, B. G., AND SINGER, M. J. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments 7*, 3, 225–240.

YAO, R., HEATH, T., DAVIES, A., FORSYTH, T., MITCHELL, N., AND HOBERMAN, P., Eds. 2014. *Oculus VR Best Practices Guide*.