Univ Access Inf Soc
DOI 10.1007/s10209-017-0581-5

LONG PAPER

Immersive 360° video user experience: impact of different variables in the sense of presence and cybersickness

David Narciso[1,2] · Maximino Bessa[1,2] · Miguel Melo[2] · António Coelho[2,3] · José Vasconcelos-Raposo[1,2]

Abstract Virtual Reality (VR) has been recently gaining interest from researchers and companies, contributing to the development of the associated technologies that aim to transport its users to a virtual environment by the stimulation of their senses. Technologies such as Head-Mounted Displays (HMD), capable of presenting 360° video in 3D, are becoming affordable and, consequently, more common among the average consumer, potentiating the creation of a market for VR experiences. The purpose of this study is to measure the influence of (a) video format (2D/monoscopic vs 3D/stereoscopic), (b) sound format (2D/stereo vs 3D/spatialized), and (c) gender on users' sense of presence and cybersickness, while experiencing a VR application using an HMD. Presence and cybersickness were measured using questionnaires as subjective measures. Portuguese versions of the Igroup Presence Questionnaire for presence and the Simulator Sickness Questionnaire for cybersickness were used. Results revealed no statistically significant differences between (a) VIDEO and (b) SOUND variables on both senses of presence and cybersickness. When paired with (a) VIDEO, the independent variable (c) Gender showed significant differences on almost all subscales of presence. Results suggest that the widely acknowledged differences in spatial ability between genders were a major factor contributing to this outcome.

✉ David Narciso
david.g.narciso@inesctec.pt

1 Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal

2 INESC TEC, Porto, Portugal

3 Faculdade de Engenharia da Universidade do Porto, Porto, Portugal

Keywords User Experience · Virtual Reality · Presence · Cybersickness · 360° 3D Video · 3D Sound

1 Introduction

Technologies associated with Virtual Reality (VR) have improved greatly over the last years. Immersive experiences capable of presenting 360 3D video and 3D sound are becoming more common to the average consumer. According to Bleumers et al. [4], 360° video can be described as moving images that have been captured so that the viewer can look around in any given angle as if they are turning the camera. 360° video can benefit from 3D technology since it leads to a higher sense of perceived depth and perceived sharpness [61]. 3D technology allows a faster perception of depth [45] and more precise cues regarding the spatial location, size, or shape of the 3D objects [23]. Although there is some research toward 3D displays (e.g. [16]), little attention on the literature has been given to the nature of the 360° content (captured video or computer-generated environment) on immersive 3D displays such as Computer Automatic Virtual Environments (CAVE) or Head-Mounted Displays (HMD). Immersive 3D displays are not to be confused with viewable 3D displays (e.g. [9]), where users can see 360° content around the display instead of being involved by it. To the best of our knowledge, a small amount of research has addressed the evaluation of 360° content visualization using an immersive 3D display (e.g. [26, 42]), and we were not able to identify studies comparing 2D versus 3D conditions using 360° content on the same immersive display. If we consider 2D versus 3D on non-immersive displays, there are more works available in the literature (e.g. [33]).

Spatialized 3D or surround sound is sound processed to give the listener the impression of a sound source within a three-dimensional environment. This is a more realistic experience

Published online: 24 October 2017

when listening to recorded sound compared to stereo, because stereo only varies across one axis, usually the x (horizontal) axis and within a limited angle of listening (typically 60°). Kramer [30] defines 3D sound as an attempt to add spatial characteristics to mediated sounds. 3D sound is created in stereo (2 channels), quadraphonic (4 channels), and especially in surround sound systems (5 or more channels) [38].

The aim of this paper is to study the influence of video and sound format on users' sense of presence and cybersickness, two widely recognized variables on the literature for evaluating VR environments. Common definitions of presence found on the literature are the ones proposed by Slater et al. [58] and Witmer et al. [65]. Slater et al. define presence as "a state of consciousness, the (psychological) sense of being in the VE," and corresponding modes of behavior [58]. Participants who are highly present should be more engaged in the VE than the outside world, and the VE should feel as a place visited and not just as images seen. Witmer et al. describe presence as "the subjective experience of being in one place or environment, even when one is physically situated in another" [65]. Due to the diverse backgrounds of scholars studying the concept of presence, there is a lack of a unified terminology. Some related terms found on the literature include telepresence, mediated presence, and virtual presence [35]. The term "telepresence" was originally associated with human operators and their sense of being transported to a remote work space via teleoperating systems. Mediated presence limits the concept of presence strictly to the real of mediated perception. Lastly, virtual presence, a term coined by [54], refers to presence caused by VR technologies, such as the work presented in this paper. For a detailed explanation and conceptualizations of presence in the literature, refer to [39]. In the year 2000, during an online discussion of Presence-L Listserv, a community of scholars interested in the concept defined presence has a psychological state or subjective perception even though part (or all) of an individual's current experiences generated by and/or filtered through human-made technology, part or all of the individual perceptions fail to accurately acknowledge the role of technology in the experience [47]. Because it is a perceptual illusion, presence is a property of a person, it can and does vary across individuals and across time for the same individual [39]. Presence can be divided into two types: physical/spatial and social presence [25]. Physical, or spatial, presence refers to the sense of physically being in another location, forgetting about the immersive technology involved, and accepting the virtual environment as a true environment. Social presence is generally referred to as the feeling of being together (and communicating) with someone.

Although systematic research into the causes and effects of presence began in the early 1990s, there is no commonly accepted paradigm for its assessment, resulting in different approaches to its measurement. Similar to the various definitions of presence, the various approaches to its measurement were introduced by scholars from diverse backgrounds [62]. Also worth mentioning is the concept of immersion which, although related, differs from the concept of presence. Slater [56, 57] distinguishes these concepts by referring to presence as the subjective experience of "being there" and immersion as an objective description of the technology. The distinction of concepts is relevant because even though a more immersive system tends to lead to a higher sense of presence, this is not always the case (e.g. [2]).

As aforementioned, the other variable used in our study to evaluate the VR environment is cybersickness. Cybersickness is an exhibition of symptoms (such as headaches, disorientation, stomach awareness, and nausea) that can occur during/after exposure to a VR environment [34]. There are several studies on cybersickness (e.g. [41, 60]) that show its influence over the usage of VR applications, and we believe that this is a key issue to address to make sure everyone can use VR technologies in an optimal way. Studying the sense of cybersickness is relevant to the field of accessibility and usability as it allows to identify which factors influence it so that we can posteriorly eliminate or reduce them, assuring that the highest number of people can enjoy access to this type of content. Similar terms found on the literature are simulator sickness and Visually Induced Motion Sickness (VIMS) [1]. LaViola Jr. [34] further describes two common types of sickness when users experience an immersive VE: cybersickness and motion sickness. Despite their similar symptomatology, they should not be confused with each other, their causes differentiate them. Cybersickness occurs strictly with visual stimulation. With motion sickness, although vision can be a contributing factor, vestibular stimulation alone can be sufficient to induce motion sickness. Motion sickness occurs if there is a conflict between visual, vestibular, and proprioceptive signals in response to a motion stimulus [50]. Depending on the VE characteristics, we consider any sickness caused as cybersickness, if there is no operator motion incorporated in the VE or motion sickness otherwise [29].

Presence and cybersickness are known to have a negative correlation [55, 65], experiencing symptoms from cybersickness leads to distraction from the VE and therefore lower sense of presence.

Presence and cybersickness measurement can be broadly divided into two general categories: subjective measures, in which a participant is asked for a conscious judgement of his/her psychological state/response related to the mediated environment, and objective corroborative measures, where an attempt to measure user responses that are automatic and without conscious deliberation is made [62].

2 Related work

Several works can be found in the literature regarding the comparison of 2D versus 3D content and its influence on the sense of presence and cybersickness and, although lower in numbers, there are also works comparing spatialized versus stereo sound on the same dependent variables. When evaluating 2D versus 3D content on the sense of presence, the majority of articles available in the literature report that a higher sense of presence is achieved when participants are presented with 3D content (e.g. [24, 46, 52]). There are also articles that report no statistically significant difference between 2D and 3D content (e.g. [3, 48, 49]) and others where the sense of presence decreased as the depth started to be exaggerated, reaching the limits of stereoscopic vision [14, 60]. Measuring the effect of 3D content on the sense of presence is a complex task as several factors can affect the resulting experience of the participant; according to Häkkinen et al., [18] these are the following: the stereoscopic content itself, the cognitive-emotional factors related to the viewer, the perceptual-physiological limitations of the viewer's visual system, its culture, and the viewing context.

Concerning 3D video content and its influence on the sense of cybersickness, several studies report a negative effect on the user, getting a higher sense of cybersickness in the 3D condition (e.g. [28, 40, 41, 52]). A study of visual fatigue performed by Choy et al. [10] found that participants watching 2D video sequences exhibited less visual fatigue compared to 3D video. Sarno [51] also performed an experiment on 3D visualization and found that participants who watched a movie segment in 3D experienced more ocular discomfort and feelings of disorientation when compared with the same movie segment in 2D. An interesting result of Sarno's experiment was that there were no significant effects of the 3D format on feelings of nausea.

Concerning 3D sound, there seems to be mixed evidence in the literature with authors arguing that presence increases with the addition of 3D sound (e.g. [20, 64]), and others with results indicating that increasing the number of audio channels does not significantly enhance the sense of presence [7, 13, 32, 36]. Lombard and Ditton [38] claim that "Despite the lack of hard evidence, it seems likely that dimensional audio, at least in many circumstances, evokes increased presence." Based on the literature, it seems that sound dimensionality has a positive influence on the sense of presence if the situation takes some sort of advantage from it, like the navigation task on Hendrix and Barfield's experiment [20]. Regarding sound and its influence on cybersickness, there are few works that can be identified in the literature. Keshavarz and Hecht [28] found that using sound or not does not have a significant difference in the sense of cybersickness. When comparing spatialized versus non-spatialized audio, Dicke et al. [12] found that predictable left-right audio movements resulted in a higher perception of cybersickness than random or no movements. A study performed by Gunther et al. [17] compared different audio conditions and achieved an interesting result: All participants who experienced motion sickness were from the spatialized sound condition (3 out of 10 participants in the audio 3D condition); however, the author says that although they cannot prove it conclusively, they have no reason to suspect a correlation between such variables.

Some works on both factors are present in the literature; however, our study differs in several aspects from all the works we found. We have not identified any work using 360° 3D captured video to study the use of stereoscopy, all the works use either non-immersive displays such as traditional monitors or TVs supporting some type of 3D technology or more immersive displays such as a CAVE environment that require the user to wear additional gear such as special glasses to see in 3D, whereas in our case the participant only needs to put on an HMD. Regarding sound, we also could not find any work in the literature that uses the same approach as ours, where instead of just using a regular Head-Related Transfer Function (HRTF), we also consider the rotation of the user's head to correctly deliver sound in the most realistic way depending on which direction the user is looking at.

Given the lack of work on this subject, we believe our work can provide valuable new insights on the impact of these new approaches on the sense of presence and cybersickness. The aim of this study is, therefore, to evaluate the sense of presence and cybersickness following a subjective methodology approach and compare it by video type, sound type, and gender with the goal of understanding the relation between such variables. In terms of sound, we compared 3D (with four channels—quadraphonic audio—with support to head tracking) with traditional 2D sound (one channel—mono audio—without support to head tracking), both played on stereo headphones. Lastly, regarding video, we compared video 360° in 2D against the exact same video with 3D technology.

3 Methods and material

3.1 Experimental setup

This is a quasi-experimental, cross-sectional study with a quantitative focus. Its main goal is to study the influence of different video formats, sound dimensionality, and gender on the feeling of presence and cybersickness on what concerns the 360° video visualization using HMD.

3.2 Sample

The sample consisted of 128 participants (63 male and 65 female) with ages ranging between 18 and 38 years ($M = 21.02$, SD = 4.604). Participants were selected using simple random sampling technique and were divided into four groups: the first group ($N = 31$) experienced 360 video with 2D sound; the second group ($N = 33$) experienced 360 video with 3D sound; the third group ($N = 32$) experienced 360 3D video with 2D sound; and lastly, the fourth group ($N = 32$) did the experience having as stimulus 360 3D video with 3D sound. All participants reported normal or corrected to normal vision.

3.3 Material

Two immersive videos were produced for this experiment. The videos were captured by making use of a 360° 3D mount for goPro cameras using 12 goPros. After the capture, the videos were processed to merge the multiple footage. A post-production was also performed to ensure that there were no artefacts such as stitching or color problems. Although both have the same content, their nature differs from one another as one of the videos is 2D and the other one 3D . The video was captured during the celebrations of St. John at Porto, an extremely popular street celebration of Portugal (Fig. 1) where they are surrounded by people and music, where participants could view the 360° scene through an HMD by looking toward any direction they wanted to explore the given scene. As the nature of the experiments depends directly on participants' feedback, great care was given to ensure that possible issues related to the synchronicity of the stimuli regarding the head movements and the video position were minimized to the extent technically possible.

3.4 Variables

The independent variables of this study are VIDEO (360° 2D video and 360° 3D video), SOUND (2D sound and 3D sound), and GENDER (Male and Female). The dependent variables are Overall presence, its subscales Immersion, Realness and Spatial Presence, and Overall cybersickness and its subscales Nausea, Oculomotor Discomfort, and Disorientation.

3.5 Metrics

For the present study, three metrics were used, namely a sociodemographic questionnaire and Portuguese versions of Igroup Presence Questionnaire (IPQ) [53] and Simulator Sickness Questionnaire (SSQ) [27]. The sociodemographic questionnaire was used in order to collect data such as age, gender, or previous experience with the technology used in

Fig. 1 Print screens of the video stimuli

the experiments. This enables us to characterize the sample groups and support theoretical interpretations.

To measure presence, we used IPQp [63], an adapted and validated version of the original IPQ to Portuguese that maintains the original validity of the questionnaire. It contains a total of 14 questions that must be answered using a five-point rating scale. To measure cybersickness, a Portuguese version of the SSQ was used. This 16-item questionnaire requires answers in a scale of 1–4 that describe the severity of the symptoms. This Portuguese version of the SSQ was obtained by following the back-translation method [8, 19] and performing the respective content validity assessment that involved the collaboration of four bilingual experts: one of them doctor in Psychology with expertise in psychometry and three of them doctors in Computer Science with expertise in VR.

3.6 Hypotheses

We hypothesized that as we delivered more realistic stimuli (i.e. with the 3D conditions of VIDEO and SOUND format), the participant's sense of presence would increase; we further hypothesized that the 3D VIDEO condition would increase cybersickness; SOUND format would not show a difference in cybersickness; and lastly, we hypothesized that Gender would not show a statistically significant difference in either presence or cybersickness nor would it have a significant interaction with other independent variables analyzed.

3.7 Experimental protocol

Each experimental scenario (360° 2D video with 2D sound, 360° 2D video with 3D sound, 360° 3D video with 2D sound, 360° 3D video with 3D sound) was randomly assigned between the participants. All the experiments were undertaken in a laboratory room where the experimenters had full control over the environmental variables. The overall luminance of the room was of approximately 40 cd/m^2, corresponding to a dim environment. There were no external sources of noise that could interfere with the experiments. For presenting the visual stimulus, a head-mounted display was used [44], and for the auditory stimulus, a pair of headphones with active noise cancelation was used [5].

The experimental procedure began by receiving participants at the experimental room and explaining to them what the experiment consisted of and how they would participate in it without revealing the purposes of the study in order to avoid bias. Then, the participants were asked to sign a consent form and to fill in a sociodemographic questionnaire. Having completed those, the participants were guided to the experimental apparatus where the experimenter helped to equip them with the head-mounted display and headphones (Fig. 2). After completing the stimuli delivery, the experimenter helped participants unequipping and guided them to a table where they could read the questionnaire instructions and fill in the IPQp questionnaire and the Portuguese version of the SSQ. The entire procedure lasted for approximately 15 min, including the preparation of the participant, the stimulus presentation, and completing the questionnaires.

Fig. 2 Participants performing the experiments (illustrative photos)

4 Results

A preliminary analysis was conducted to eliminate outliers. In total, 30 outliers were detected and removed from the sample to further analyse the collected data. The resulting sample consisted of 98 participants (53 male and 45 female). Of them, 21 tested the 360° 2D video with 2D sound, 26 experienced 360° video with 3D sound, 26 the 360° 3D video with 2D sound, and the remaining 25 the 360° 3D video with 3D sound. For the purposes of the present study, the normal distribution of the data was determined through *Skewness* and *Kurtosis*. *Skewness* values varied from − 0.132 to 1.303, and *Kurtosis* ranged from − 0.571 to 0.761; the values obtained showed a normal distribution for all variables [15], 2010). Thus, parametric statistics were used. To study the effect of the independent variables (2 (Gender) × 2 (Video format) × 2 (Sound format)), we performed multivariate analysis of variance (MANOVA).

4.1 Presence

A MANOVA of the independent variable Gender regarding presence showed no significant differences between conditions. $Wilks'\ \lambda = 0.979$, $Z(3, 88) = 0.618, p = 0.605, \eta_p^2 = 0.21$ and OP = 0.174. An ANOVA of independent variable Gender regarding overall presence and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variable Video regarding overall presence showed no statistically significant differences between conditions. $Wilks'\lambda = 0.981, Z(3, 88) = 0.534, p = 0.660, \eta_p^2 = 0.18$ and OP = 0.155.

An ANOVAof the independent variable Video regarding overall presence and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variable Sound regarding presence showed no statistically significant differences between conditions. $Wilks'\lambda = 0.998, Z(3, 88) = 0.049, p = 0.986, \eta_2^2 = 0.002$ and OP = 0.58.

An ANOVA of the independent variable Sound regarding overall presence and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variables Gender × Video regarding presence showed no statistically significant differences between conditions. $Wilks'\lambda, Z(3, 88) = 2.088, p = 0.108, \eta_2^2 = 0.066$ and OP = 0.518.

An ANOVA of the independent variables Gender × Video regarding overall presence and its subscales showed statistically significant differences. Results show that female participants scored higher spatial presence, realness, and overall presence in the video 3D condition while male participants achieved a higher spatial presence, realness, and overall presence in the video 2D condition.

A MANOVA of independent variables Gender × Sound regarding presence showed no statistically significant differences between conditions. $Wilks' \lambda = 0.921, Z(3, 88) = 2.520, p = 0.063, \eta_2^2 = 0.079$ and OP = 0.605.

An ANOVA of the independent variables Gender × Sound regarding overall presence and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variables Video × Sound regarding presence showed no statistically significant differences between conditions. $Wilks' \lambda = 0.966, Z(3, 88) = 1.018, p = 0.388, \eta_2^2 = 0.034$ and OP = 0.268.

Table 1 details the ANOVA of independent variables Gender × Video in regard to overall presence and its subscales.

An ANOVA of the independent variables Video × Sound regarding overall presence and its subscales showed no statistically significant differences between conditions. The most expressive values were obtained in the realness subscale of presence, with a $M(\mathrm{SD}) = 12,52(2,91)$ in the 2D video and 2D sound condition, $M(\mathrm{SD}) = 13,77(2,69)$ for the 2D video with 3D sound condition, $M(\mathrm{SD}) = 13,46(3,60)$ for the 3D video with 2D sound condition and $M(\mathrm{SD}) = 12,52(3,20)$ for the 3D video and 3D sound condition. The rest of the parameters were $Z = 3,084$, $p = 0,82$, $\eta_2^2 = 0,033$ and lastly an OP = 0,412.

A MANOVA of the independent variables Gender × Video × Sound regarding presence showed no statistically significant differences between conditions. $Wilks' \lambda = 0.984$, $Z(3, 88) = 0.476$, $p = 0.700$, $\eta_2^2 = 0.016$ and OP = 0.143. An ANOVA of the independent variables Gender × Video × Sound regarding overall presence and its subscales showed no statistically significant differences between conditions.

Regarding the marginal, it was estimated with a confidence interval of 95% and the scores were 95% CI [22.22, 23.62] for Spatial Presence, 95% CI m 2[12.48, 13.74] for Realness, 95% CI [11.91, 13.25] for Involvement, and 95% CI [47.15, 50.07] for Overall Presence.

4.2 Cybersickness

A MANOVA of the independent variable Gender regarding cybersickness showed no statistically significant differences between conditions. $Wilks' \lambda = 0.989$, $Z(3, 88) = 0.320$, $p = 0.811$, $\eta_2^2 = 0.11$ and OP = 0.110.

An ANOVA of the independent variable Gender regarding overall cybersickness and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variable Video regarding cybersickness showed no statistically significant differences between conditions. $Wilks' \lambda = 0.982$, $Z(3, 88) = 0.532$, $p = 0.662$, $\eta_2^2 = 0.18$ and OP = 0.155.

An ANOVA of the independent variable Video regarding overall cybersickness and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variable Sound regarding cybersickness showed no statistically significant differences between conditions. $Wilks' \lambda = 0.985$, $Z(3, 88) = 0.438$, $p = 0.726$, $\eta_2^2 = 0.015$ and OP = 0.135.

An ANOVA of the independent variable Sound regarding overall cybersickness and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variables Gender × Video regarding cybersickness showed no statistically significant differences between conditions. $Wilks' \lambda = 0.942$, $Z(3, 88) = 1.807$, $p = 0.152$, $\eta_2^2 = 0.058$ and OP = 0.455.

Table 2 details the ANOVA of independent variables Gender × Video in regard to overall cybersickness and its subscales.

An ANOVA of the independent variables Gender × Video regarding overall cybersickness and its subscales showed statistically significant differences. Results showed that female participants scored higher nausea in the 2D video condition while male participants achieved higher nausea in the 3D video condition.

A MANOVA of the independent variables Gender × Sound regarding cybersickness showed no statistically significant differences between conditions. $Wilks' \lambda = 0.990$, $Z(3, 88) = 0.305$, $p = 0.822$, $\eta_2^2 = 0.010$ and OP = 0.107.

Table 1 ANOVA of independent variables Gender × Video in regard to overall presence and its subscales

	Male		Female					
	Video 3D	Video 3D	Video 2D	Video 3D				
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*Z*	*p*	eta	OP
Spatial Presence	24.09 (3.12)	22.50 (3.06)	21.67 (3.80)	23.43 (3.76)	5.440	0.022	0.057	0.636
Realness	13.83 (3.00)	12.53 (3.41)	13.63 (2.58)	13.67 (3.38)	3.941	0.050	0.042	0.502
Involvement	12.78 (3.86)	12.93 (3.10)	11.54 (2.48)	13.00 (3.65)	0.794	0.375	0.009	0.143
Overall Presence	50.70 (7.38)	47.97 (6.01)	45.83 (6.34)	50.10 (9.05)	5.731	0.019	0.060	0.658

Table 2 ANOVA of independent variables Gender × Video in regard to overall cybersickness and its subscales

	Male		Female					
	Video 3D	Video 3D	Video 2D	Video 3D				
	M (SD)	*M* (SD)	*M* (SD)	*M* (SD)	*Z*	*p*	eta	*OP*
Nausea	1.66 (3.70)	4.45 (6.00)	3.98 (5.57)	2.27 (4.16)	4.680	0.033	0.049	0.572
Oculomotor discomfort	8.57 (9.51)	10.61 (10.83)	12.00 (13.02)	6.86 (8.27)	2.884	0.093	0.031	0.390
Disorientation	13.31 (16.51)	13.92 (17.61)	18.56 (19.97)	13.92 (17.61)	0.626	0.431	0.007	0.123
Overall cybersickness	8.46 (9.18)	10.72 (10.88)	12.47 (11.71)	8.01 (8.47)	2.758	0.100	0.030	0.376

An ANOVA of the independent variables Gender × Sound regarding overall cybersickness and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variables Video × Sound regarding cybersickness showed no statistically significant differences between conditions. $Wilks'\lambda = 0.977$, $Z(3, 88) = 0.706$, $p = 0.551$, $\eta_2^2 = 0.023$ and $\mathrm{OP} = 0.194$.

An ANOVA of the independent variables Video × Sound regarding overall cybersickness and its subscales showed no statistically significant differences between conditions.

A MANOVA of the independent variables Gender × Video × Sound regarding cybersickness showed no statistically significant differences between conditions. $Wilks'\lambda = 0.990$, $Z(3, 88) = 0.307$, $p = 0.820$, $\eta_2^2 = 0.010$ and $\mathrm{OP} = 0.107$.

An ANOVA of independent variables Gender × Video × Sound regarding overall cybersickness and its subscales showed no statistically significant differences between conditions.

Regarding the marginal, it was estimated with a confidence interval of 95% and the scores were 95% CI [2.05, 4.14] for Nausea, 95% CI [7.39, 11,79] for Oculomotor discomfort, 95% CI [11.32, 18.78] for Disorientation, and 95% CI [7.87, 12.11] for Overall Cybersickness.

5 Discussion

The main research question was whether Video format, Sound format, and/or Gender influence the senses of presence or cybersickness of a participant in a VE. The hypotheses for the outcomes of this study were described in Sect. 3.6. The following paragraphs present a comparison of the hypotheses, in the same order in which they were described, with the results of the statistical analyses. In short, the results do not comply with any of our hypothesis apart from the assumptions regarding the sound format and cybersickness, as results showed no statistically significant differences between Video and Sound conditions in the senses of presence or cybersickness.

Regarding Video format and the sense of presence, we expected that combining 3D technology with 360° video and an immersive media (HMD) for visualization would benefit the VR experiment, resulting in higher levels of presence felt by the participants. Interestingly, the obtained results showed no statistically significant differences between video format and presence. This result is particularly surprising in the realness subscale of the IPQp questionnaire where we expected better results given the increase in consistency between virtual and real worlds provided by the 3D technology. This outcome can be broadly attributed to the fact that visual discomfort is expected to occur with such technology [52], even when disparity values do not surpass one degree limit of visual angle [31]. Given the negative correlation between presence and cybersickness, we postulate that visual discomfort has, to some extent, influenced this result. Further, we speculate that the sense of presence could increase in the 3D condition if some task requiring close range interaction was used, similar to Narayan's experiment [42]. His experiment showed a positive influence of 3D visualization on task performance (the task consisted of two users handling a virtual object with accuracy at close range). However, this was not feasible in our study due to a limitation of the type of content used (see Study limitations and recommendations Section for further details). Snow and Williges [59] also suggest that beneficial effects of stereopsis might increase with the difficulty of a participant's task, meaning that a task with a higher level of difficulty could also contribute to an increased sense of presence. Another possible explanation concerning the content of the video itself can be derived from Bowman & McMahan, where they state that, when visualizing a simple, regular, and easy to understand environment, less immersive systems can provide the same results as more immersive ones [6]. Finally, one aspect that should be noted is the content of the VE: Participants were presented with a video of a physical place in the real world instead of a computer-generated place. The fact that we used a video could have influenced the results as an environment with high level of realness may have put the user in an already high level of presence making the difference between 2D Video and 3D Video conditions less noticeable.

Concerning Sound format and the fact that it did not show significant differences in the sense of presence, we advance with a possible explanation: the environment is generally noisy (in this case, a local celebration where the participant is surrounded by people and music) and the participant is overrun with sounds that make it harder to take advantage of the discrete audio channels. The same explanation was used by [13] in an experiment where changing the sound dimensionality did not show significant effects in presence. We speculate that the sense of presence could be sharpened if the participant had to rely on directional audio cues to complete a task, like it occurred in Hendrix and Barfield's experiment [20]. However, in our study, we could not employ tasks due to our decision regarding the type of content used.

In terms of video format and cybersickness, we expected higher values for the 3D Video condition as other works in the literature have shown; however, on our study this did not happen. We believe the reason for this to happen is strongly related to the fact that a fixed point of view was used instead of a moving environment (e.g. [28]), which led to the low level of dropouts and low cybersickness experienced.

Regarding sound and its influence on the sense of cybersickness, there are few works in the literature available on the subject. One of the existent works [12] claims that spatialized sound may increase the sense of cybersickness if the visual sense is not synchronized with what one hears. However, in our case, the stimulus was synchronized so we had no reason to think that a higher sense of cybersickness would be felt in the 3D sound condition. We believe that the generally noisy environment contributed to this outcome since if there were more sound sources that had more emphasis/intensity on a certain direction, and if those were not synchronized, it could increase cybersickness.

When analysing Gender together with the other independent variables, we found a statistically significant effect on Video and Gender on both presence and cybersickness, showing that, with 3D Video condition, female participants reported greater scores on overall presence, spatial presence, and realness and a lower score of nausea. Male participants, in turn, had the reverse pattern, showing a higher overall presence, spatial presence, and realness and lower nausea with 2D Video and lower presence scores and higher nausea with 3D Video. A fairly similar result is found on Narayan's [42] study where the absence of stereoscopic vision had a greater negative effect on female participants that is in line with previous studies in psychology that show a difference in cognitive abilities across genders [21]. A participant's sense of presence is influenced by the attributes of the VR platform, the features of the VR environment, the task itself, and user characteristics [43]. A possible explanation would be that user characteristics influenced the resulting experience, namely the differences between genders, such as the widely acknowledged differences in spatial ability and rotation [11, 37]: Females perceived 3D Video as more natural, while males perceived 2D Video as more natural, leading to lower levels of nausea, and higher levels on the sense of presence.

5.1 Threats to validity

Some precautions were taken to avoid threatening the validity of the study. First, for each participant, the HMD was adjusted to minimize the risk of discomfort when using the HMD. Second, the 360 video was obtained from an array of 12 GoPro's and respective videos were stitched using Kolor Autopano Video Pro, thus ensuring a correct stitching and synchronization of the final video. The final video has a resolution of 4000×2000 in the 2D condition and a resolution of 4000×4000 in the 3D condition (two videos of 4000×2000, rendering one in each eye) and has a frame rate of 47 frames per second (FPS), giving the user a perception of fluid movement. Lastly, the head movements were tracked at a frame rate of 75 FPS, the recommended frame rate for the HMD used (Oculus Rift DK2).

A limitation of this study concerns the type of content used for the experiment. We presented users with a 360° video of a physical place in the real world, which limits any possible interactions during the experiment as we cannot dynamically change the course of actions/events in a captured video in the same way as on a virtual simulation. Results of an experiment by [28] regarding the effect of content type (video *vs*. simulation) on cybersickness showed that participants reported higher feelings of cybersickness in the video condition. We speculate that varying the content type would also inherently influence the sense of presence.

Another possible limitation of the study concerns the content of the video: The video is of a local celebration where people gather around and listen to music. The point of view is located at a relatively far distance (approx. 3–4 m) and, although there was movement, few people passed near the camera rig. We speculate that lack of near movement can contribute to 3D not being able to provide a noticeable difference on the sense of presence, since it is more effective at close range [6].

Based on these limitations, future work will contemplate two additional sets of experiences: first to study the effects of varying the content type (Video *vs*. simulation) on the sense of presence and cybersickness, and second to study the addition of interaction (No task *vs*. task requiring close range handling of a virtual object) on the same dependent variables.

6 Conclusions

The work presented in this paper compared different 360° video formats (2D/monoscopic and 3D/ stereoscopic) combined with different sound formats (2D/non-spatialized and 3D/spatialized) across genders in order to understand their impact on the sense of presence and cybersickness felt in a VR experience. This fact is important for areas of application such as the immersive journalism since it shows that simpler and more affordable solutions can be used for accessing such contents such as the google cardboard.

Although our results do not comply with the majority of the papers found in the literature, we need to take into consideration several factors in which our study differs. Over the years, technology properties such as FOV, frame rate, resolution, image density (PPI), and so on change, which will inherently influence the user's sense of presence and cybersickness. Even if we take into account just the 3D technology, it presents several pitfalls and traps that were more likely to occur in older studies: varying viewing conditions, different image intensities, ghosting, flicker, the speed/accuracy tradeoff, subjects' stereo acuity, and the degree of difficulty in the discrimination task [22]. Finally, few studies had positional and/or rotational head tracking, which has been seen as an enhancement on the sense of presence (e.g. [20, 59]).

Both video and audio formats did not produce a significant effect on the sense of presence or cybersickness. We believe the major factor as to why this occurred is related to the nature of the content used in the experiments. It did not take advantage of the added depth of stereoscopy because people were the only non-static element and they did not make much movement nor did they pass very often in front of the camera rig. As for sound, the environment was generally noisy making the participant being overrun with sounds from every direction, thus making it harder to benefit from discrete sound channels. The comparison of Gender and Video revealed significant differences between male and female, which implies that there is a need for additional care when dealing with a specific gender.

Future work is necessary to verify if using the appropriate content will cause significant differences on the sense of presence and/or cybersickness. For this, we propose to perform a set of experiments having as reference a taxonomy of scenes and tasks. Also, future work toward understanding the differences in visual systems between genders is needed to come closer to an explanation to why women felt a higher sense of presence and lower cybersickness with the 3D Video and why men had the same effect though on 2D Video. We believe that if the participant is in a considerably high level of presence and if there is no content or task that takes advantage of the added depth or localized sounds, 3D Video and 3D Sound will not make a significant difference.

Acknowledgements This work was supported by the project "TEC4Growth Pervasive Intelligence, Enhancers and Proofs of Concept with Industrial Impact/NORTE-01- 0145-FEDER-000020" financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF). The research works were conducted at MASSIVE VR Laboratory, an output of the project REC I/EEI-SII/0360/2012 entitled "MASSIVE - Multimodal Acknowledgeable multiSenSory Immersive Virtual Environments" financed by the European Union (COMPETE, QREN and FEDER).

References

1. Bando, T., Iijima, A., Yano, S.: Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: a review. Displays **33**(2), 76–83 (2012)
2. Baños, R.M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., Rey, B.: Immersion and emotion: their impact on the sense of presence. Cyberpsychol. Behav. **7**(6), 734–741 (2004)
3. Baños, R.M., Botella, C., Rubió, I., Quero, S., García-Palacios, A., Alcañiz, M.: Presence and emotions in virtual environments: the influence of stereoscopy. Cyberpsychol. Behav. **11**(1), 1–8 (2008)
4. Bleumers, L., Van den Broeck, W., Lievens, B., Pierson, J.: Seeing the bigger picture: a user perspective on 360 tv. In: Proceedings of the 10th European Conference on Interactive Tv and Video, ACM, New York, NY, USA, EuroiTV '12, pp. 115–124, doi:10.1145/2325616.2325640 (2012)
5. Bose (2017) QuietComfort 15 - acoustic noise cancelling headphones. https://www.bose.com/en_us/support/products/headphones_support/over_ear_headphones_support/qc15.html
6. Bowman, D.A., McMahan, R.P.: Virtual reality: how much immersion is enough? Computer **40**(7), 36–43 (2007)
7. Brinkman, W.P., Hoekstra, A.R., van EGMOND, R.: The effect of 3D audio and other audio techniques on virtual reality experience. IOS Press, Amsterdam (2015)
8. Brislin, R.W.: Back-translation for cross-cultural research. J. Cross Cult. Psychol. **1**(3), 185–216 (1970)
9. Butler, A., Hilliges, O., Izadi, S., Hodges, S., Molyneaux, D., Kim, D., Kong, D.: Vermeer: direct interaction with a 360 viewable 3d display. In: Proceedings of the 24th annual ACM symposium on User interface software and technology, ACM, pp 569–576 (2011)
10. Choy, S.M., Chiu, K.H., Cheng, E., Burnett, I.: 3d fatigue from stereoscopic 3d video displays: comparing objective and subjective tests using electroencephalography. In: TENCON 2015–2015 IEEE Region 10 Conference, IEEE, pp 1–4 (2015)
11. Coluccia, E., Louse, G.: Gender differences in spatial orientation: a review. J. Environ. Psychol. **24**(3), 329–340 (2004)
12. Dicke, C., Aaltonen, V., Billinghurst, M.: Occurrence of simulator sickness in spatial sound spaces and 3d auditory displays. International Conference on Auditory Display (2009)
13. Freeman, J., Lessiter, J.: Here, there and everywhere: the effects of multichannel audio on presence. In: Fourth Annual International Workshop on Presence, Georgia Institute of Technology (2001)
14. Freeman, J., Avons, S.E., Meddis, R., Pearson, D.E., IJsselsteijn, W.: Using behavioral realism to estimate presence: a study of the utility of postural responses to motion stimuli. Presence **9**(2), 149–164 (2000)
15. George, D., Mallery, P.: SPSS for windows step by step: a simple guide and reference, 11.0 Update. Allyn and Bacon (2010)
16. Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., et al blue-c: a spatially immersive display and 3d video portal for telepresence. In: ACM Transactions on Graphics (TOG), ACM, vol 22, pp 819–827 (2003)

17. Gunther, R., Kazman, R., MacGregor, C.: Using 3d sound as a navigational aid in virtual environments. Behav. Inform. Technol. **23**(6), 435–446 (2004)
18. Häkkinen, J., Kawai, T., Takatalo, J., Leisti, T., Radun, J., Hirsaho, A., Nyman, G.: Measuring stereoscopic image quality experience with interpretation based quality methodology. In: Electronic Imaging 2008, International Society for Optics and Photonics, pp. 68,081B–68,081B (2008)
19. Hambleton, R., Zenisky, A.L.: Translating and adapting tests for cross-cultural assessments. Cross Cult. Res. Methods Psychol. 46–70 (2011)
20. Hendrix, C., Barfield, W.: The sense of presence within auditory virtual environments. Presence Teleoperators Virtual Environ. **5**(3), 290–301 (1996)
21. Hoffman, M., Gneezy, U., List, J.A.: Nurture affects gender differences in spatial abilities. Proc. Natl. Acad. Sci. **108**(36), 14,786–14,788, doi:10.1073/pnas.1015182108, http://www.pnas.org/content/108/36/14786.abstract, http://www.pnas.org/content/108/36/14786.full.pdf (2011)
22. Hsu, J., Pizlo, Z., Babbs, C., Chelberg, D.M., Delp III, E.J.: Design of studies to test the effectiveness of stereo imaging truth or dare: is stereo viewing really better? In: IS&T/SPIE 1994 International Symposium on Electronic Imaging: Science and Technology, International Society for Optics and Photonics, pp 211–222 (1994)
23. Hubona, G.S., Wheeler, P.N., Shirah, G.W., Brandt, M.: The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization. ACM Trans. Comput. Hum. Interact. (TOCHI) **6**(3), 214–242 (1999)
24. IJsselsteijn, W., de Ridder, H., Freeman, J., Avons, S.E., Bouwhuis, D.: Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. Presence **10**(3), 298–311 (2001)
25. IJsselsteijn, W.A., de Ridder, H., Freeman, J., Avons, S.E.: Presence: concept, determinants, and measurement. In: Electronic Imaging, International Society for Optics and Photonics, pp. 520–529 (2000)
26. Juan, M.C., Pérez, D.: Comparison of the levels of presence and anxiety in an acrophobic environment viewed via hmd or cave. Presence **18**(3), 232–248 (2009)
27. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**(3), 203–220 (1993)
28. Keshavarz, B., Hecht, H.: Visually induced motion sickness and presence in videogames: the role of sound. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56. SAGE Publications, pp. 1763–1767 (2012)
29. Kolasinski, E.M.: Simulator sickness in virtual environments. Tech. rep, DTIC Document (1995)
30. Kramer, G.: Sound and communication in virtual reality, communication in the age of virtual reality. pp. 259–276 (1995)
31. Lambooij, M.T., IJsselsteijn, W.A., Heynderickx, I.: Visual discomfort in stereoscopic displays: a review. In: International Society for Optics and Photonics Electronic Imaging 2007, vol. 6490 (2007)
32. Larsson, P., Västfjäll D., Olsson, P., Kleiner, M.: When what you hear is what you see: presence and auditory-visual integration in virtual environments. In: 10th Annual International Workshop on Presence (2007)
33. LaViola, J.J. Jr, Litwiller, T.: Evaluating the benefits of 3d stereo in modern video games. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, CHI '11, pp. 2345–2354, doi:10.1145/1978942.1979286, (2011)
34. LaViola Jr., J.J.: A discussion of cybersickness in virtual environments. ACM SIGCHI Bull. **32**(1), 47–56 (2000)
35. Lee, K.M.: Presence, explicated. Commun. Theory **14**(1), 27–50 (2004)
36. Lessiter, J., Freeman, J.: Really hear? the effects of audio quality on presence. In: Fourth Annual International Workshop on Presence, Citeseer (2001)
37. Linn, M.C., Petersen, A.C.: Emergence and characterization of sex differences in spatial ability: a meta-analysis. Child Dev. 1479–1498 (1985)
38. Lombard, M., Ditton, T.: At the heart of it all: the concept of presence. J. Comput. Mediat. Commun. **3**(2), (1997a)
39. Lombard, M., Ditton, T.: At the heart of it all: the concept of presence. J. Comput. Mediat. Commun. **3**(2), (1997b)
40. McCauley, M.E., Sharkey, T.J.: Cybersickness: perception of self-motion in virtual environments. Presence Teleoperators Virtual Environ. **1**(3), 311–318 (1992)
41. Naqvi, S.A.A., Badruddin, N., Malik, A.S., Hazabbah, W., Abdullah, B.: Does 3d produce more symptoms of visually induced motion sickness? In: 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp. 6405–6408 (2013)
42. Narayan, M., Waugh, L., Zhang, X., Bafna, P., Bowman, D.: Quantifying the benefits of immersion for collaboration in virtual environments. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, ACM, pp. 78–81 (2005)
43. Nash, E.B., Edwards, G.W., Thompson, J.A., Barfield, W.: A review of presence and performance in virtual environments. Int. J. Hum. Comput. Interact. **12**(1), 1–41 (2000)
44. Oculus Oculus Rift - development kit 2. https://www3.oculus.com/en-us/dk2/ (2017)
45. Palmisano, S.: Consistent stereoscopic information increases the perceived speed of vection in depth. Perception **31**(4), 463–480 (2002)
46. Patel, H., Stefani, O., Sharples, S., Hoffmann, H., Karaseitanidis, I., Amditis, A.: Human centred design of 3-d interaction devices to control virtual environments. Int. J. Hum. Comput. Stud. **64**(3), 207–220 (2006)
47. Presence Research IS (2000) The concept of presence: explication statement. Retrieved on May 5, 2016
48. Raja, D.: The effects of immersion on 3d information visualization. PhD thesis, Virginia Polytechnic Institute and State University (2006)
49. Rand, D., Kizony, R., Feintuch, U., Katz, N., Josman, N., Weiss, P.L.T., et al.: Comparison of two vr platforms for rehabilitation: video capture versus hmd. Presence Teleoperators Virtual Environ. **14**(2), 147–160 (2005)
50. Reason, J.T., Brand, J.J.: Motion Sickness. Academic press, Cambridge (1975)
51. Sarno, D.M.: The effect of stereoscopic (3d) movies on psychological and physiological experiences. PhD thesis, Bridgewater State University (2015)
52. Schild, J., LaViola, J., Masuch, M.: Understanding user experience in stereoscopic 3d games. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 89–98 (2012)
53. Schubert, T., Friedmann, F., Regenbrecht, H.: The experience of presence: factor analytic insights. Presence **10**(3), 266–281 (2001)
54. Sheridan, T.B.: Musings on telepresence and virtual presence. Presence Teleoperators Virtual Environ. **1**(1), 120–126 (1992)
55. Singer, M.J., Witmer, S. US Army Research Institute. Interim Report (1996)
56. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. R. Soc. Lond. B Biol. Sci. **364**(1535), 3549–3557 (2009)
57. Slater, M., Usoh, M., Steed, A.: Depth of presence in virtual environments. Presence Teleoperators Virtual Environ. **3**(2), 130–144 (1994)

58. Slater, M., Linakis, V., Usoh, M., Kooper, R., Street, G.: Immersion, presence, and performance in virtual environments: an experiment with tri-dimensional chess. In: ACM Virtual Reality Software and Technology (VRST), vol. 163. ACM Press New York, NY, p. 72 (1996)
59. Snow, M.P., Williges, R.C.: Empirical models based on free-modulus magnitude estimation of perceived presence in virtual environments. Hum. Factors J. Hum. Factors Ergon. Soc. **40**(3), 386–402 (1998)
60. Takatalo, J., Kawai, T., Kaistinen, J., Nyman, G., Häkkinen, J.: User experience in 3d stereoscopic games. Media Psychol. **14**(4), 387–414 (2011)
61. Tam, W.J., Stelmach, L.B., Corriveau, P.J.: Psychovisual aspects of viewing stereoscopic video sequences. In: Photonics West'98 Electronic Imaging, vol 3295. pp 226–235, doi:10.1117/12.307169 (1998)
62. Van Baren, J., IJsselsteijn, W.: Measuring presence: A guide to current measurement approaches. Deliverable of the OmniPres project IST-2001-39237 (2004)
63. Vasconcelos-Raposo, J., Bessa, M., Melo, M., Barbosa, L., Rodrigues, R., Teixeira, C.M., Cabral, L., Augusto Sousa, A.: *ipq* adaptation and validation of the igroup presence questionnaire in a portuguese sample. Presence Teleoperators Virtual Environ. **25**(3), 191–203 (2016)
64. Västfjäll, D.: The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. Cyberpsychol. Behav. **6**(2), 181–188 (2003)
65. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence Teleoperators Virtual Environ. **7**(3), 225–240 (1998)