ORIGINAL ARTICLE

Quasi-3D: reducing convergence effort improves visual comfort of head-mounted stereoscopic displays

Vittorio Dalmasso[1] · Michela Moretti[1] · Claudio de'Sperati[1]

Abstract

The diffusion of virtual reality urges to solve the problem of vergence-accommodation conflict arising when viewing stereoscopic displays, which causes visual stress. We addressed this issue with an approach based on reducing ocular convergence effort. In virtual environments, vergence can be controlled by manipulating the binocular separation of the virtual cameras. Using this technique, we implemented two quasi-3D conditions characterized by binocular image separations intermediate between 3D (stereoscopic) and 2D (monoscopic). In a first experiment, focused on perceptual aspects, ten participants performed a visuo-manual pursuit task while wearing a head-mounted display (HMD) in head-constrained (non-immersive) condition for an overall exposure time of ~ 7 min. Passing from 3D to quasi-3D and 2D conditions, progressively resulted in a decrease of vergence eye movements—both mean convergence angle (static vergence) and vergence excursion (dynamic vergence)—and an increase of hand pursuit spatial error, with the target perceived further from the observer and larger. Decreased static and dynamic vergence predicted decreases in asthenopia trial-wise. In a second experiment, focused on tolerance aspects, fourteen participants performed a detection task in near-vision while wearing an HMD in head-free (immersive) condition for an overall exposure time of ~ 20 min. Passing from 3D to quasi-3D and 2D conditions, there was a general decrease of both subjective and objective visual stress indicators (ocular convergence discomfort ratings, cyber-sickness symptoms and skin conductance level). Decreased static and dynamic vergence predicted the decrease in these indicators. Remarkably, skin conductance level predicted all subjective symptoms, both trial-wise and session-wise, suggesting that it could become an objective replacement of visual stress self-reports. We conclude that relieving convergence effort by reducing binocular image separation in virtual environments can be a simple and effective way to decrease visual stress caused by stereoscopic HMDs. The negative side-effect—worsening of spatial vision—arguably would become unnoticed or compensated over time. This initial proof-of-concept study should be extended by future large-scale studies testing additional environments, tasks, displays, users, and exposure times.

Keywords Asthenopia · Cyber-sickness · Depth perception · Head-mounted display · Near vision · Sensory-motor adaptation · Skin conductance · Physiological signals · Stereoscopic display · Spatial vision · Vergence eye movements · Virtual reality · Visual stress

1 Introduction

Finding the best conditions that ensure a comfortable user experience is an important goal of 3D visual technologies. Immersive solutions with head-mounted displays (HMD), for example, provide reasonably realistic visual experiences but fall short as to the achieved comfort. Cyber-sickness symptoms such as nausea, disorientation and ocular difficulties are often reported by users when navigating virtual environments with HMDs (Rebenitsch and Owen 2016; Gallagher and Ferre 2018).

Vittorio Dalmasso and Michela Moretti have contributed equally to this work.

✉ Claudio de'Sperati
desperati.claudio@unisr.it

Vittorio Dalmasso
v.dalmasso1@studenti.unisr.it

[1] Laboratory of Action, Perception and Cognition, School of Psychology, Vita-Salute San Raffaele University, Milan, Italy

In addition to the visuo-vestibular conflict that can give rise to motion sickness (Akiduki et al. 2003), another conflict can create adverse conditions when viewing stereoscopic displays, namely, the so-called Vergence-Accommodation Conflict (VAC), which has been reported to generate ocular discomfort or asthenopia (Rushton et al. 1994; Hoffman et al. 2008; Shibata et al. 2011; Zhou et al. 2021). VAC is produced by a mismatch between the focal distance and the vergence distance: the fixed focal length imposed by the display optics is in conflict with the varying vergence eye movements required to visually inspect the represented 3D scene. In short, vergence eye movements are made in the absence of concurrent accommodation changes. It is therefore not surprising that several methods aimed at reducing VAC have been proposed, often aimed at driving accommodation to meet the vergence requirements (Kramida 2016; Koulieris et al. 2017; Lyu et al. 2021).

Here we explored a method to reduce VAC based on an opposite strategy, namely, modifying vergence rather than modifying accommodation. The rationale is to reduce the convergence effort by reducing horizontal binocular image separation, in turn obtained by reducing the horizontal virtual camera separation that generates the stereoscopic 3D environment. By reducing image separation, it is expected that the vergence angle between the two eyes is accordingly reduced, and, consequently, also VAC is reduced. When image separation is null, a monoscopic visual stimulus is obtained (Rebenitsch and Owen 2016), a condition corresponding to a null vergence angle (as when viewing an object at infinity). This condition is achieved by so-called bi-ocular displays (Rushton et al. 1994). Intermediate image separations can be thought of as quasi-3D conditions. An opposite manipulation, i.e., hyperstereopsis, was used for example to study visual discomfort and adaptive mechanisms (Kooi and Toet 2004; Priot et al. 2010).

In addition to the benefit of reducing the VAC, decreasing the convergence effort could have positive effects on its own. Indeed, excessive near vision has been indicated as a cause of eyestrain, especially when using digital devices (Collier and Rosenfield 2011; Cooper et al. 2011; Coles-Brennan et al. 2019). For example, ocular comfort during reading has been reported to be correlated with the degree of exophoria (Collier and Rosenfield 2011), thus predicting a positive effect of decreasing the convergence effort over and beyond VAC reduction. The so-called 20-20-20 rule (take a 20-s break every 20 min and look at something 20-feet away) recommended by the American Optometric Association goes in this direction. More generally, decreasing the oculomotor load by reducing the amount of eye movements may count as a small yet positive benefit (Zangemeister et al. 2020).

However, manipulating the binocular image separation is known to alter depth perception (Jones et al. 2001; Gao et al. 2018), and in fact this method has been exploited in stereoscopic digital cinema ("Playing with camera's interocular distance will create size effects on objects, landscapes and actors, and make them feel giant or small", (Mendiburu 2009/2012), Chapter 2). Thus, it is important to consider both the pros of possible improvements in visual comfort and the cons of associated spatial distortion.

This study was aimed at assessing spatial vision and visual stress with different amounts of image separation in a virtual environment, a manipulation that is expected to affect convergence effort. In a first experiment, focused on spatial vision, a group of participants performed a visuomotor pursuit task under non-immersive (head-restrained) condition, while in a second experiment, focused on visual stress, another group of participants performed a near-vision detection task under immersive (head-free) condition.

2 Experiment 1

In this experiment we focused on spatial vision. We measured target localization in depth and apparent target size in four viewing conditions characterized by a different amount of binocular image separation. Subjective sensation of ocular fatigue (asthenopia) was also assessed. Convergence effort was computed in terms of both static vergence (mean vergence angle over time) and dynamic vergence (mean amplitude of vergence eye movements over time, respectively), as it is has been suggested that it is the variable, rather than the fixed, decoupling of accommodation and vergence that is more problematic (Wann et al. 1995). For this reason, we used a task that is expected to evoke continuous vergence eye movements (visuo-manual pursuit task in depth). In this experiment, head-constrained condition allowed us to experimentally control vergence eye movements, which would be impossible in a head-free condition unless using poorly realistic head-centered target motion.

2.1 Methods

2.1.1 Participants

Ten healthy participants aged between 20 and 24 (3 males) were recruited on a voluntary basis. They were right-handed and had normal or corrected-to normal vision. The study was approved by the Ethical Committee of the Milan University ("Parere 50/19"). Signed informed consent was obtained from all participants.

2.1.2 Stimuli and task

The visual stimuli were created under Unity3D, which was used to build an executable application driven in run-time by a Matlab script to control the experiment. The crucial

experimental manipulation consisted in reducing the horizontal separation between the two virtual cameras in 33.33% steps (Fig. 1), such that in one viewing condition the camera separation matched the individual inter-pupillary distance (ordinary stereoscopic viewing, we call it "3D" condition), while in another viewing condition the camera separation was null (monoscopic viewing, we call it "2D" condition). Other two viewing conditions were characterized by intermediate amounts of camera separation (66.66% and 33.33% of inter-pupillary distance, we call these two quasi-3D conditions "3D-" and "3D--", respectively). The experiment was conducted using both a reduced environment consisting of a dark grey uniform background and a structured environment—a simplified office. The underlying rationale was to compare a very reduced environment, typical of stimuli used
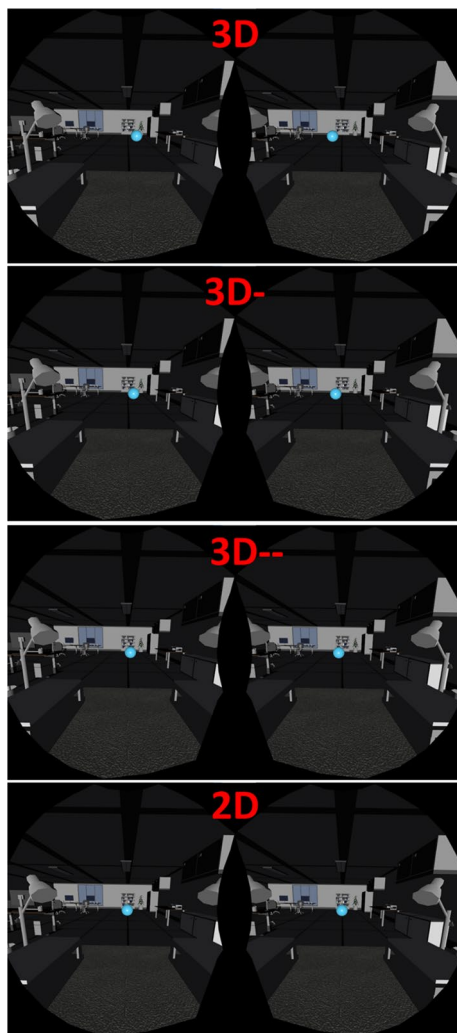

Fig. 1 The visual stimulus used in the first experiment, emphasizing the four viewing conditions with a different binocular image separation (3D, 3D-, 3D--, 2D). The cyan sphere is the target. Only the structured visual environment is shown. The title was not present during the experiment

in visual psychophysics, and a (slightly) more ecologically-relevant environment.

In each of the two virtual environments, participants performed an open loop manual pursuit task, which consisted in tracking with the dominant hand a visual target (rendered inside the virtual environment as a fluctuating solid cyan sphere with a 4 cm diameter) moving along a circular trajectory (10 cm diameter) in front of the observer on the horizontal gaze plane. The center of the trajectory was aligned with the body mid-sagittal plane. The target moved at constant velocity at 0.333 cycles / s for 18 s (6 cycles). For each of the 4×2 experimental conditions (4 viewing conditions $\times$ 2 environments) there were 3 trials with trajectories centered at 16, 20 or 24 cm away from the participant's ocular plane, covering overall a depth range between 11 and 29 cm, thus well into peripersonal space. The presentations of the 3 trials were separated by a 2-s blank period, and their sequence was randomized. Overall, the total exposure time was 7.2 min. Importantly, participants did not see their hand or the hand controller (see below), such that only proprioceptive afferent information was available to track the moving target (open-loop condition). The presentation sequence for the 4 viewing conditions was counterbalanced across participants (with 2 sequences: 3D→3D-→3D--→2D; and 2D→3D--→3D-→3D), while the presentation sequence for the 2 environments was fixed (first the reduced then the structured environment). The fixed sequence served to have one condition in which the target was the only spatial element in the scene, before the exposure to the spatially structured environment could be used to construe a memory-based visuo-motor spatial representation also in the reduced environment condition.

2.1.3 Hardware

The experiment was run on a Dell tower PC Aurora (with Intel Core i9-9900K at 3.6 GHz) equipped with a GeForce GTX 2070 Nvidia graphic board. The virtual environment was displayed through an HTC Vive Pro-Eye HMD (field of view = 110 deg, resolution = 2560 × 1200, frame rate = 90 fps), which also featured embedded binocular eye tracking through a Tobii device (sampling frequency = 90 Hz, nominal accuracy < 1 deg). Factory calibration, which included vertical offset of the HMD, inter-pupillary distance regulation and eye movements calibration, was performed prior to the beginning of the recording session. Head and hand movements were concurrently recorded through the Vive tracking devices with an accuracy of < 9 mm (Verdelet et al. 2019), and downsampled to 90 Hz (outside-in tracker, two-lighthouse system). For head movements recording, the HMD with integrated position sensors was used, while for hand movements recording the Vive hand-held controller was used.

2.1.4 Procedure

Participants were invited to sit as comfortably as possible inside the recording frame, and then helped to wear the visor, resting the upper part of the front on a forehead support, just above the border of the visor (Fig. 2). They were asked to keep their posture as stable as possible during the experimental sessions, compatibly with the hand pursuit movements. This solution was somewhat less comfortable than a free-moving, standing posture, but had the advantage of keeping a fixed distance between the eyes and the target: without this caution, participants might have idiosyncratically preferred to shift the head toward the target to improve pursuit, or to accompany the pursuit with head movements, possibly differently for the three pursuit trajectories, and this would have produced inconsistent eye movements data, which would be especially damaging for vergence data (the target motion was specified in world coordinates, not in head-centered coordinates).

Overall, the experiment lasted about 30 min, including the preparation of the experimental setup and the familiarization period with the virtual environment and the task. We took all the required measures against Covid-19. Besides the general laboratory measures (face mask, gloves, etc.), we adopted individual disposable masks tailored for HMDs, and disinfected the visor face pad, the forehead support and the hand controller after each use.

2.1.5 Data analyses

Signal synchronization was attained by offline aligning the time stamps of target, head, hand and eye movements recordings. Eye position, vergence and hand position signals were low-pass filtered (cut-off frequency: 30 Hz, 0.5 Hz

Fig. 2 The experimental setup. Thanks to a top-down forehead support, the wooden structure ensured a good head restraint while at the same time leaving free space for arm/hand movements in front of the participant

and 0.5 Hz, respectively). The first pursuit cycle was always discarded.

We stipulated that negative vergence means binocular convergence, such that zero-vergence indicates parallel ocular axes. We also defined static vergence as the mean value of vergence angle during a trial, and dynamic vergence as the mean excursion of oscillatory vergence movements associated to target motion in depth. To measure the excursion, a sinusoidal model was fitted to the vergence trace averaged over the 5 pursuit cycles, taking the peak-to-peak fitted amplitude as a measure of dynamic vergence.

To assess the subjective judgments of asthenopia, an 11-point numeric rating scale (NRS-11) was used, while for the apparent target size a direct diameter estimate was asked. At the end of each trial, participants answered verbally the following questions: (1) "Rate in a 0–10 scale how strong was your feeling of ocular fatigue"; (2) "Tell how many cm the target diameter was". The experimenter noted the participant's answers.

For the statistical analyses, we used an approach based on Generalized Linear Mixed Models (GLMM), which can accommodate different distributions for the dependent variables, including ordinal variables, and which can handle the within-subjects experimental design (subjects as random effect). The restricted maximum likelihood method was used, setting covariance structure to compound symmetry (similarly to ANOVA) and Satterthwaite approximation for the degrees of freedom. Convergence was always achieved. In Table 1 are listed the distributions and links used for each dependent variable (DV), as well as the typology (continuous or categorical) of the independent variables (IV). Distributions and link functions have been chosen based on the actual DV distributions and AIC values. Robust estimation was used with rating responses to mitigate possible errors in self-reports. The models had the general form: DV- $1 + IV1$ $[+ IV2 + IV1{:}IV2 + \ldots] + (1 \mid SUBJ)$.

In these models based on regression, the viewing condition was a continuous variable and not a categorical variable (we recall that each viewing condition is associated to a given camera binocular distance, in turn yielding a given binocular image separation). Indeed, the term quasi-3D does not only identify the two intermediate conditions tested in the present study, i.e., 3D- and 3D--, but the entire range between stereoscopic (3D) and monoscopic (2D) viewing.

Signal processing and plots have been performed with Matlab, while statistical analyses have been performed with SPSS (v.28). Only statistically significant results are reported in the main text.

2.2 Results

In the 3D condition, observers performed the manual pursuit task rather precisely and accurately, with quite regular

Table 1 List of distributions, links and variable typology used for the GLMM analyses

Experiment 1

  DV

    Pursuit error, depth (PD): Gaussian + Identity

    Pursuit error, lateral (PL): Gaussian + Identity

    Static vergence (SV): Gaussian + Identity

    Dynamic vergence (DV): Gaussian + Identity

    Target distance (TD): Gaussian + Identity

    Target size (TS): Gaussian + Identity

    Asthenopia (TS): Multinomial + Cumulative logit

  IV

    Viewing condition (VC): Continuous

    Environment (EN): Categorical

    Sequence (SEQ): Categorical

    Static vergence (SV): Continuous

    Dynamic vergence (SV): Continuous

Experiment 2

  DV

    Stimulus distance (SD): Gaussian + Identity

    Static vergence (SV): Gaussian + Identity

    Dynamic vergence (DV): Gamma + Log

    Convergence discomfort (CD): Multinomial + Cumulative logit

    Simulation sickness questionnaire, total score (SSQ_T): Multinomial + Cumulative logit

    Simulation sickness questionnaire, ocular score (SSQ_O): Multinomial + Cumulative logit

    Skin conductance level (SCL): Gamma + Log

    Heart rate (HR): Gamma + Log

    Heart rate variability (HRV): Gaussian + Identity

  IV

    Viewing condition (VC): Continuous

    Repetition (REP): Categorical

    Sequence (SEQ): Categorical

    Static vergence (SV): Continuous

    Dynamic vergence (SV): Continuous

    Skin conductance level (SCL): Continuous

    Ocular dominance (OD): Categorical

    Lens correction (LC): Categorical

    Interpupillary distance (IPD): Continuous

trajectories (Fig. 3, left panels). By reducing binocular image separation (3D-, 3D-- and 2D), manual pursuit tended to become less accurate, however, increasingly stepping away from the target, as can be better appreciated in the average spatial plots (Fig. 4).

To quantify pursuit accuracy, we measured the pursuit spatial error as the average signed distance between hand position and target position in the depth plane and in the horizontal plane (the vertical plane was not relevant). On average, spatial error in the depth plane passed from -0.2 cm (S.D. = 2.0) in the 3D condition to 10.2 cm (S.D. = 6.2) in the 2D condition, with 3D- and 3D-- conditions yielding intermediate errors (Fig. 6a), indicating that the target appeared beyond its actual position.

The results of the GLMM analyses showed that the viewing condition significantly affected manual pursuit in the depth plane [$F(1,67) = 156.154$, $p < 0.001$] and in the horizontal plane [$F(1,67) = 5.021$, $p = 0.028$], the latter likely being due to a small asymmetry of pursuit movement performed with one hand. Smooth pursuit errors showed an increasing trend when passing from 3 to 2D.

Even though observers did not receive instructions to track the target with the eyes, smooth pursuit eye movements were always present. Indeed, as shown in the individual examples (Fig. 3, right panels), smooth pursuit eye movements tracked continuously the moving target, which included a smooth vergence change dictated by the continuous target depth change. This can be better seen in Fig. 5, where the vergence traces were averaged over the 5 cycles of each trial and over subjects. Reducing binocular image separation resulted in two clear effects: a decrease of mean convergence (static convergence) and a decrease of vergence excursion (dynamic vergence). By averaging over the three trajectories, the former passed from a convergence angle of -15.19 deg (S.D. = 1.17) in the 3D condition to an almost null convergence angle of -0.41 deg (S.D. = 0.61) in the 2D condition, while the latter passed from a peak-to-peak amplitude of 5.44 deg (S.D. = 0.85) in the 3D condition to a peak-to-peak amplitude of 0.72 deg (S.D. = 0.44) in the 2D condition (Fig. 6b). The effect of the viewing condition was statistically significant in both cases [static vergence: $F(1,67) = 5971.043$, $p < 0.001$; dynamic vergence: $F(1,67) = 479.211$, $p < 0.001$].

A consequence of reducing binocular image separation was a significant change of target's apparent size [$F(1,60) = 104.944$, $p < 0.001$], which on average passed from an estimated diameter of 2.4 cm in the 3D viewing condition (S.D. = 0.6) to 8.3 cm in the 2D viewing condition (S.D. = 2.5), with intermediate figures for intermediate image separations (Fig. 6a). The interaction viewing condition × environment was statistically significant [$F(1,60) = 7.019$, $p = 0.010$].

Probing the subjective feeling of ocular fatigue (asthenopia) showed a significant effect of the viewing condition [$F(1,68) = 13.152$, $p < 0.001$]. Also, participants judged the structured environment to be significantly more effortful for the eyes [$F(1,68) = 5.098$, $p = 0.027$], possibly because the structured environment was always presented in the second session.

To further investigate the factors at the origin of asthenopia, we assessed the trial-wise relationship between asthenopia and vergence regardless of the viewing condition, controlling for the viewing condition presentation order.

Fig. 3 Example of hand movements recordings (left panels, spatial plots from top view) and binocular horizontal eye movements recordings (right panels, temporal plots) in a representative participant in the four viewing conditions (from top to bottom: 3D, 3D-, 3D--, 2D). Target trajectories in the spatial plots and vergence eye movements in the temporal plots are also shown. In the temporal plots, the three trajectories (near, intermediate and far) are plotted separately over time, and reproduce the actual (randomized) presentation order for this participant. Only recordings from the structured visual environment are shown
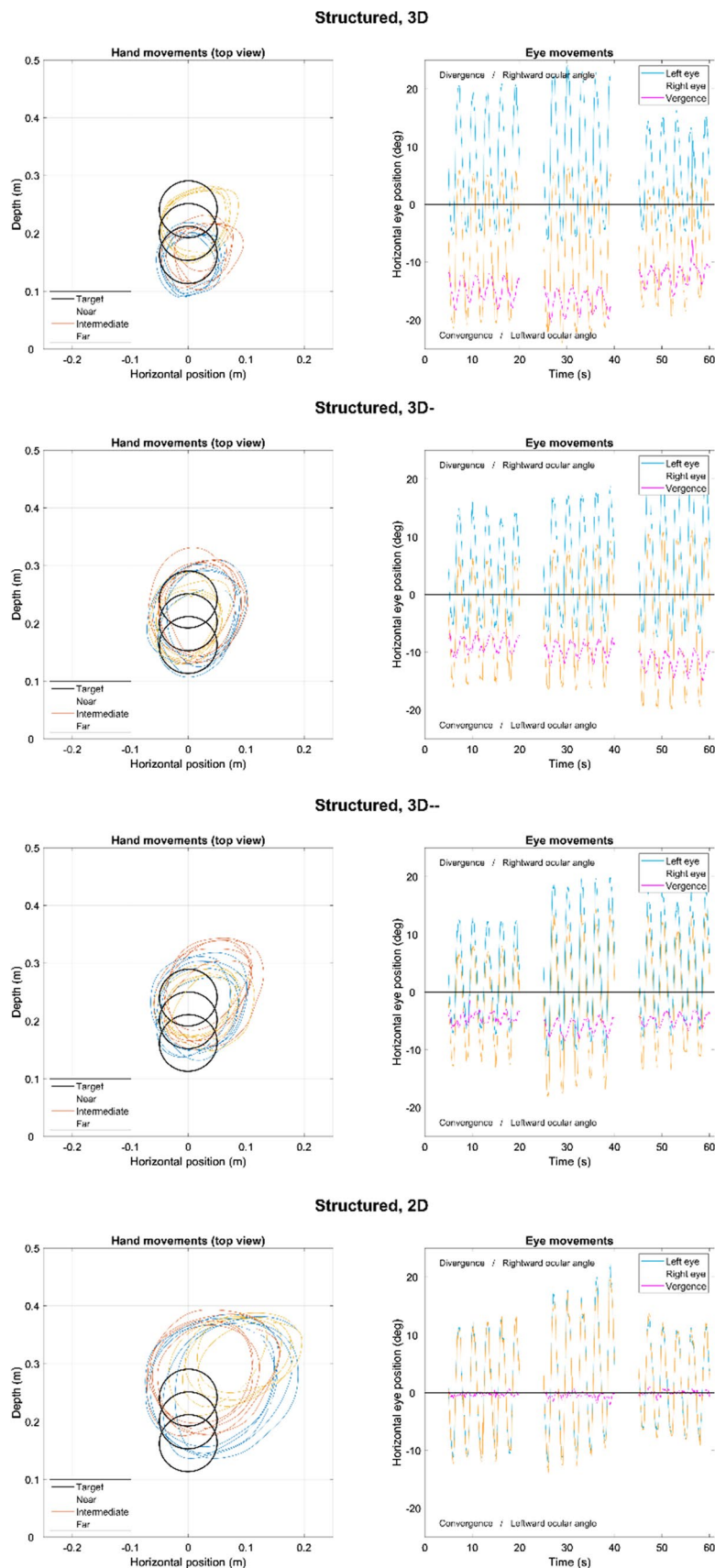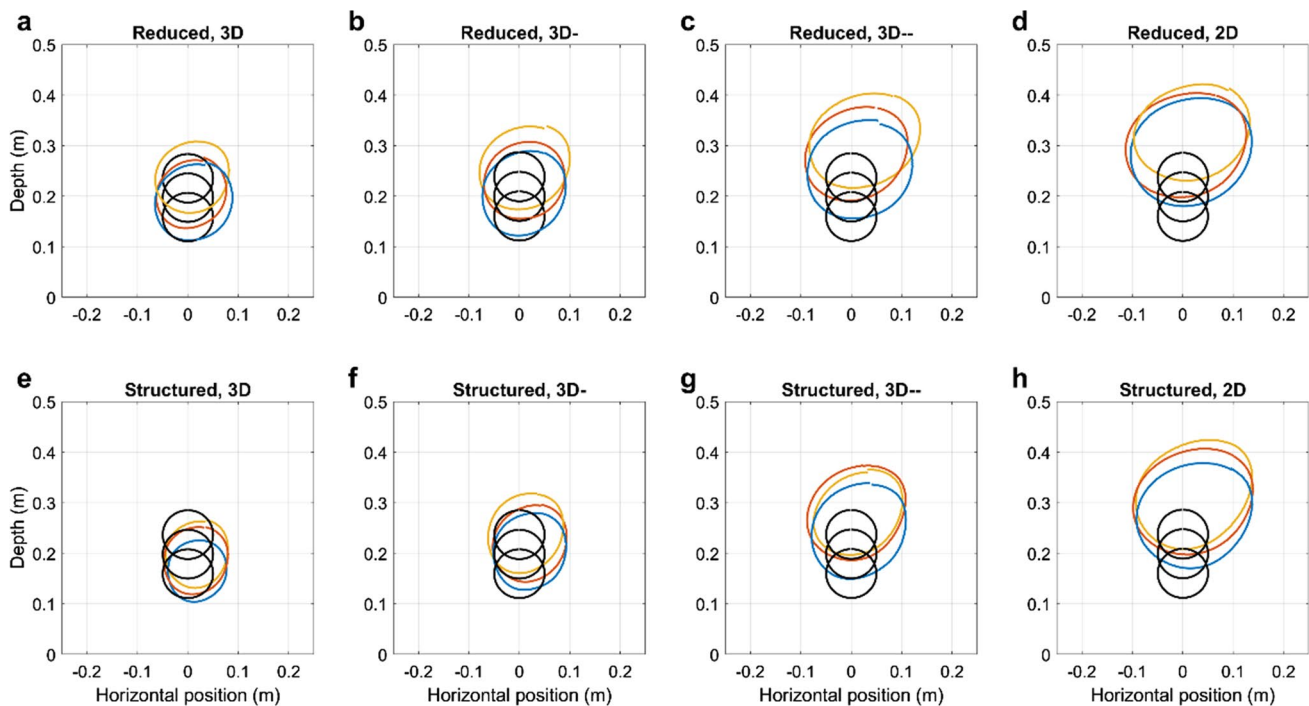
Fig. 4 Mean hand trajectories (top view) during the pursuit task in the four viewing conditions and the two visual environments. The three hand trajectories (near, intermediate and far) are depicted in different colors, while the three target trajectories are in black. Coordinates [0,0] indicate the inter-ocular point. Same conventions as in Fig. 3

We found a significant tendency towards a decrease of asthenopia symptoms with decreasing ocular convergence angle (static vergence: $F(1,63) = 5.389$, $p = 0.024$) and with decreasing amount of vergence eye movements (dynamic vergence: $F(1,63) = 4.074$, $p = 0.048$).

The results of the reported analyses are summarized in Table 2.

2.3 Discussion

A first finding of this experiment was a clear reduction of vergence from the 3D to the 2D viewing condition, passing through the two conditions with intermediate image separation (3D- and 3D--). This held true for both the mean vergence angle (static vergence) and the vergence excursion (dynamic vergence) during the manual pursuit task. Thus, the convergence effort that is normally required to visually explore 3D scenes in stereoscopic displays was systematically reduced by reducing binocular image separation.

Reducing image separation introduced a spatial localization error, as assessed through a visuo-manual pursuit task. The mean error increased from the 3D to the 2D viewing condition, passing through the two intermediate conditions. This held true for the near-to-far direction (translation in the depth plane), indicating that the target appeared beyond its actual position, but also for the right-left direction, where

the error showed a small and almost constant bias. The mean localization error in depth could become rather large—in the 2D condition it was 10.2 cm, i.e., 51% of the average target distance (20 cm).

However, measuring spatial localization using the visuo-manual pursuit task could have overestimated the perceptual localization error, because involuntary wrist rotations might have accompanied manual pursuit movements, with the effect of partly dissociating the controller's sensed position from the effective hand position (the center of the Vive controller's handle is ~ 10 cm below the controller's sensed position). We measured the mean controller's translational component resulting from wrist rotation in the horizontal and depth planes, which in the 2D viewing condition reached 2.5 and 1.9 cm, respectively. Considering that wrist rotation should be synergic with the global hand movement, and not opposing to it, this means that the tendency to rotate the wrist could have contributed up to about one-fourth of the translational component of manual pursuit error in depth (2.5 cm out of 10.2 cm). Therefore, the effective spatial localization error in depth in the 2D viewing condition, i.e., when the error was maximal, could be less than 10.2 cm: by taking the 2.5 cm figure above to be the maximal motor artefact introduced by wrist rotation, it would reduce to 7.7 cm (38% instead of 51%). However, it should be borne in mind that we cannot know whether participants performed
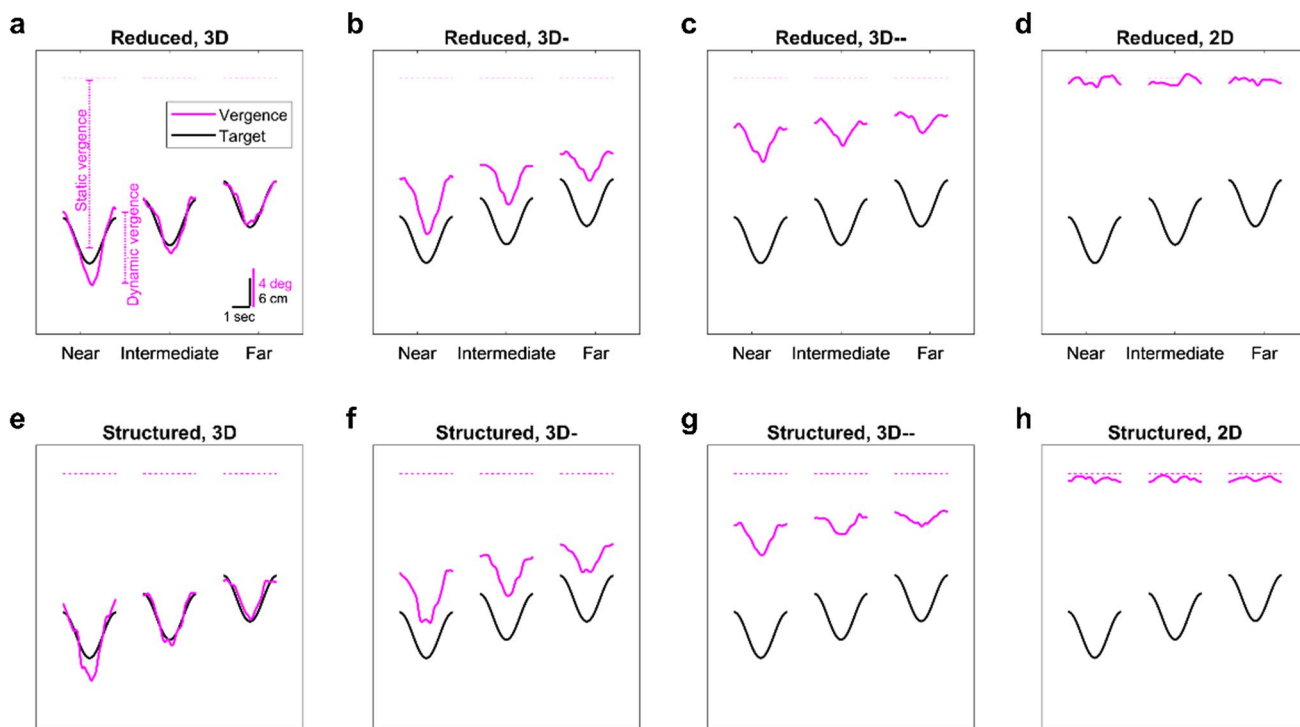
Fig. 5 Mean vergence eye movements during the pursuit task in the four viewing conditions and the two visual environments. The three trajectories (near, intermediate and far) are plotted separately over time, ordered for increasing distance from the participant. The horizontal dotted lines represent the zero-vergence condition (parallel ocular axes). Target motion is also shown, with arbitrary offset for graphical purpose. Downward deflection represents far-to-near eye movements (convergence) and far-to-near target motion. Panel a) also reports an example of computation of static and dynamic vergence (vertical magenta bars), as well as calibration bars for horizontal and vertical axes (all panels have the same axes)

Fig. 6 Effects of viewing condition on vergence and spatial vision. Data from the two visual environments are collapsed together. Data-points and error bars represent means and standard deviations across participants
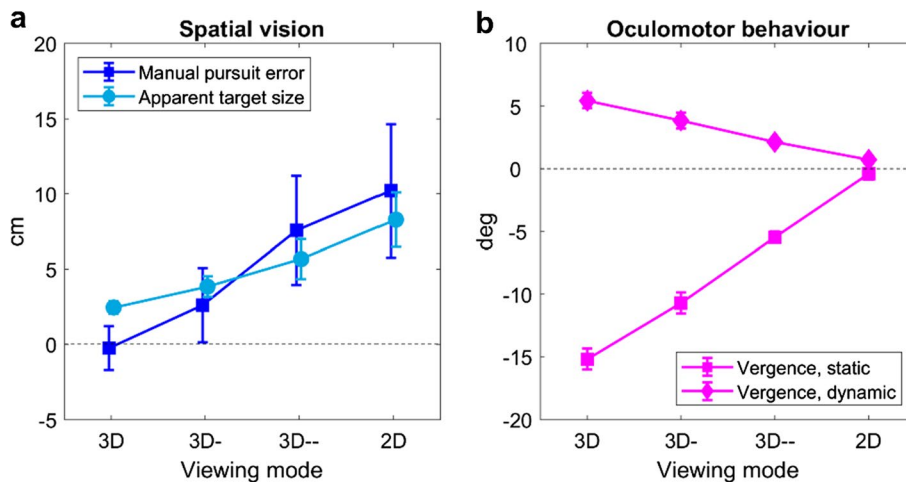
the pursuit movement by placing the center of the hand on the target—as implied by the task instruction—or whether instead the controller was functionally integrated in the action plan and used as a reaching tool. In the latter case, wrist rotation should be considered as an integral part of the pursuit movement. For this reason, it is safer to consider an approximate range of spatial error in the 2D viewing condition to be between 7 and 10 cm in the peripersonal space towards the extrapersonal space, at least in our experimental conditions—and correspondingly smaller errors in the 3D-- and 3D- viewing conditions.

We attribute this perceptual distortion—the target perceived further from the observer with decreasing image separation—to differences in vergence: the degree of binocular

Table 2 GLMM results for the first experiment

DV	IV	F	df1	df2	p value	Coeff	Lower	Upper
Pursuit error, depth	ViewingCondition	156.154	1	67	**< 0.001**	0.034	0.025	0.042
	Environment	3.069	1	67	0.084			
	ViewingCondition × Environment	0.813	1	67	0.370			
Pursuit error, lateral	ViewingCondition	5.021	1	67	**0.028**	0.003	0.000	0.007
	Environment	1.072	1	67	0.304			
	ViewingCondition × Environment	0.123	1	67	0.727			
Static vergence	ViewingCondition	5971.043	1	67	**< 0.001**	5.054	4.873	5.235
	Environment	1.528	1	67	0.221			
	ViewingCondition × Environment	2.192	1	67	0.143			
Dynamic vergence	ViewingCondition	479.211	1	67	**< 0.001**	− 1.565	− 1.770	− 1.360
	Environment	0.097	1	67	0.756			
	ViewingCondition × Environment	0.117	1	67	0.733			
Target size	ViewingCondition	104.944	1	60	**< 0.001**	0.014	0.009	0.020
	Environment	3.744	1	60	0.058			
	ViewingCondition × Environment	7.019	1	60	**0.010**			
Asthenopia	ViewingCondition	13.152	1	68	**< 0.001**	− 0.481	− 0.893	− 0.069
	Environment	5.098	1	68	**0.027**			
	ViewingCondition × Environment	0.271	1	13	0.611			
	Static vergence	25.905	1	63	**< 0.001**	− 0.094	− 0.131	− 0.057
	Dynamic vergence	12.070	1	63	**< 0.001**	0.242	0.103	0.382

For each dependent variable (DV), the fixed-effects tables report the effects of the independent variables (IV), i.e. the main effect of viewing condition (ViewingCondition), the main effect of environment (Environment) and the interaction between viewing condition and environment (ViewingCondition × Environment). For each predictor, the values for the F-statistic (F), degrees of freedom (df) and significance (p value) are reported, as well as the estimated coefficient and 95% confidence interval (lower and upper) when appropriate. For asthenopia, the effects of static and dynamic vergence were assessed through additional, distinct analyses because of the correlation among predictors. Statistically significant results are reported in bold

alignment is an information that the visual system exploits to recover the distance of an object, especially inside the peripersonal space, where reaching and grasping movements are fundamental to interact with the (virtual) environment (Mon-Williams and Dijkerman 1999; Melmoth et al. 2007; Naceri et al. 2011). However, it should be borne in mind that vergence signals (carrying motor information) and retinal disparity signals (carrying sensory information) are difficult to disentangle under binocular vision (Linton 2020).

This distortion in spatial localization may in turn explain another result of this study, namely, the increase of target's apparent size when the binocular image separation was reduced. The latter can be interpreted in the light of the Emmert's law: when two objects that have the same retinal size are perceived to be at a different distance from the observer, the further one is perceived to be larger (Gregory 2008). In the present study, reducing image separation increased the target's apparent distance from the observer; hence, it looked larger.

Another finding emerging from this experiment is that asthenopia depended on the viewing condition, and more specifically on vergence, both static and dynamic: the more parallel the ocular axes and the less pronounced the vergence eye movements, the less intense the feeling of asthenopia. This finding is in keeping with the observations that maintaining the eyes excessively converged for near vision can be associated to adverse visual symptoms (Collier and Rosenfield 2011; Cooper et al. 2011; Coles-Brennan et al. 2019), although difficulties in performing vergence eye movements may not be necessarily accompanied by a subjective sensation of visual fatigue (Zheng et al. 2021). The additional burden carried by the vergence accommodation conflict (VAC) may exacerbate the cost of vergence eye movements when using stereoscopic displays, where controlling vergence would thus be particularly important.

However, a note of caution should be made in regard of the asthenopia measure, because it is not totally obvious what exactly our participants judged when providing their ratings. Indeed, we posed a very simple and generic question ("Rate in a 0–10 scale how strong was your feeling of ocular fatigue"), without differentiating among the constellation of symptoms that can accompany cyber-sickness (Urvoy et al. 2013; Rebenitsch and Owen 2016; Gallagher and Ferre 2018; Kim et al. 2018). What we might have probed here could be a somewhat indistinct and generic sensation of visual discomfort.

3 Experiment 2

In the first experiment, participants wore an HMD but the head-restrained condition that was used to keep experimental control on vergence prevented immersivity, which is a fundamental feature of virtual reality. To test the effects of image separation under immersive condition, we ran a second experiment in which participants could freely move their head. To force as much as possible near vision, we used a detection task that required close visual inspection. In addition to introduce immersivity, this experiment also aimed at achieving a more thorough assessment of visual stress. For this, we specifically probed subjective convergence discomfort, measured cyber-sickness symptoms with a standard questionnaire, and recorded autonomic activity.

3.1 Methods

3.1.1 Participants

Sample size was chosen based on a power analysis. Indeed, while in the first experiment we used convenience sampling without concern for statistical power (an a-posteriori power analysis, made with MorePower 6.0.4, indicated a Cohen's $f^2 = 0.47$ for the main effect of Viewing condition, a rather large effect size), in this experiment we computed a priori the sample size required to ensure that, with $\alpha = 0.05$ and power $= 0.8$, the main effect of Viewing condition would be detected with Cohen's $f^2 \geq 0.30$ (medium effect), the main effect of Repetition would be detected with Cohen's $f^2 \geq 0.25$ (medium effect), and the interaction Viewing condition $\times$ Repetition would be detected with Cohen's $f^2 \geq 0.12$ (small effect). The resulting sample size was N = 14. Thus, fourteen healthy participants aged between 21 and 29 (mean age = 22.5 years, 3 males) were recruited on a voluntary basis. They had normal or corrected-to-normal vision, with no history of strabismus. The study was approved by the Ethical Committee of the Milan University ("Parere 50/19"). Signed informed consent was obtained from all participants.

3.1.2 Stimuli and task

In this experiment, we used the same manipulations of horizontal camera separation already tested in the first experiment (i.e., 3D, 3D-, 3D-- and 2D viewing conditions, counterbalanced), and also the structured environment that was designed for the first experiment (a simplified office). The moving target was replaced by a stationary sphere (4 cm diameter) presented 15 cm in front of the observer (world coordinates). The sphere had a heterogeneous surface texture and rotated on its vertical axis at constant speed (0.2 rotations per second). The participants' task was to detect a small, unpredictable change in the surface texture (detection task). For that, a cover story was prepared as follows: *"Dear participant, we are conducting a study on visual perception in virtual reality, using as a scenario an environment where familiar elements and science fiction elements coexist. The familiar element is an office, while the sci-fi element is a rotating semi-organic sphere. The sphere has self-evolutionary capabilities: every now and then, rarely and unpredictably, it produces a tiny change on its surface. Your task is to detect the change. For this you will need to inspect the surface of the sphere very closely. We tell you right away that it is extremely difficult to perceive the change, and you will have to look at the sphere very closely. During the session, you will need to tell whether you detected the change. Immediately thereafter, you will also need to rate, from 0 to 10, how uncomfortable you found eye convergence (the tendency to "cross-eye"), where 0 = no discomfort and 10 = intolerable discomfort. You will give these answers verbally every minute after an audible signal. The test consists of 4 experimental sessions. Each session lasts about 5 min and is preceded by 30 s of relaxation. At the end of each session, you will answer verbally a short questionnaire on your comfort, after which you can take a short break. Thanks for your participation.".* Because the detection task was not relevant to the study's goal (in fact no change on the rotating sphere occurred, it was just functional to focus participants' attention on a near-vision task), it will not be considered in the analyses.

3.1.3 Hardware and devices

Hardware was the same used in Experiment 1, except that the controller was not used because this experiment did not involve a manual task. In addition, a device for psychophysiological recordings was used to measure heart beats (photoplethysmography) and skin conductance (Shimmer 3, Shimmer Research Ltd.; sampling frequency = 51 Hz).

3.1.4 Procedure

Participants were invited to sit on a chair placed inside the recording room. As a first step, two electrodes were placed on the proximal phalanges of the forefinger and ring finger of the non-dominant hand for skin conductance recording, while the photoplethysmography sensor plate was placed on the distal phalange of the middle finger. Participants were then helped to wear the HMD. After checking that the participant felt comfortably and could move the head easily, and after the calibration phase was performed, the experiment started. The session scheme is illustrated in Fig. 7. Overall, the experiment lasted about 40 min, including the
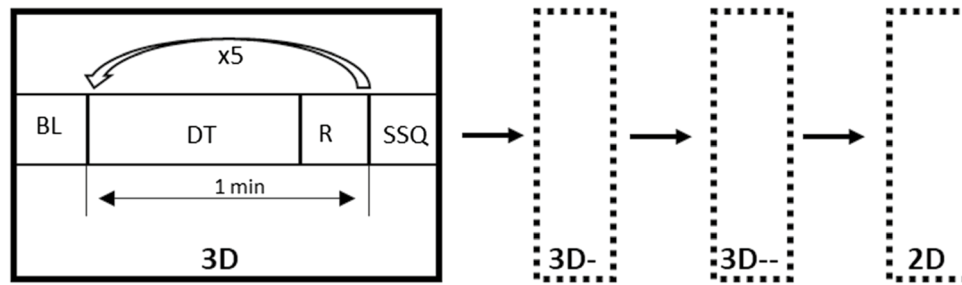
Fig. 7 Schematic illustration of the procedure in the second experiment. Time flows from left to right. The left box represents a session, which started with a 30-s baseline period (BL), followed by the detection task (DT) and the rating of convergence discomfort (R). The latter two phases (DT and R) were repeated 5 times, without interruptions, and lasted 5 min. At the end, the SSQ was administered. Except for the BL and SSQ periods (during which the system returned to the default Unity landscape, a ground surface with mountains' profile on the horizon), the visual stimulus was always present throughout the session. The dotted boxes represent the 3 subsequent sessions, identical in structure to the first one, administered in sequence with a brief pause between them. Each session is characterized by a given visual condition (3D, 3D-, 3D--, 2D). In this figure the 3D→2D sequence is represented, but in the experiment it was counterbalanced across participants with the 2D→3D sequence

preparation of the experimental setup and the familiarization period with the virtual environment and the task.

3.1.5 Data analyses

Signal processing related to head and eye movements was the same as in experiment 1. However, whereas static vergence (SV) was again computed as the mean vergence angle along a trial, dynamic vergence (DV) was measured as its standard deviation. The reason is that in this experiment the vergence changes were not driven by target motion (predictable) but by head movements (unpredictable).

Additional signal processing was performed to measure skin conductance and heart rate. Both skin and heart signals were smoothed offline (2-steps moving average, 17-points). Instantaneous heart rate was obtained from the smoothed inter-beat interval series. For each trial, we computed the mean skin conductance level (SCL) as well as the mean heart rate (HR) and heart rate variability (HRV), the latter computed as the standard deviation of instantaneous heart rate. SCL was obtained by extracting the slow component of skin conductance signal (low-pass cut-off: 0.1 Hz), which filtered out the non-specific skin conductance responses (NS-SCR) as well as possible movement artefacts.

To assess cyber-sickness symptoms, we used the Simulator Sickness Questionnaire (SSQ), a psychometric tool created to appraise intolerance when using visual interfaces (Kennedy et al. 1993), also widely used in virtual reality contexts to measure cyber-sickness (Bimberg et al. 2020). In this study we used both the SSQ total score (SSQ_T), which covers a rather wide symptoms range (nausea, oculomotor and disorientation), and a score (SSQ_O) obtained by summing the individual scores of the three SSQ items specifically indexing ocular problems (item 4: asthenopia; item 5: difficulty focusing; item 11: blurred vision). We refer to the SSQ_T score as indexing cyber-sickness symptoms and the SSQ_O score as indexing ocular symptoms.

To assess discomfort due to convergence effort (convergence discomfort, CD), an 11-point numeric rating scale (NRS-11) was used ("Rate in a 0–10 scale how strong was the discomfort due to ocular convergence"). CD was intended to be more specifically related to convergence effort than the generic asthenopia ratings used in the first experiment ("Rate in a 0–10 scale how strong was your feeling of ocular fatigue"). CD was complimentary to SSQ in that convergence discomfort is not specifically targeted by SSQ's items.

For the statistical analyses, we used the same GLMM approach described for the first experiment.

In Table 1 are listed the distributions and links used for each dependent variable (DV), as well as the typology (continuous or categorical) of the independent variables (IV). Signal processing and plots have been performed with Matlab, while statistical analyses have been performed with SPSS. Only statistically significant results are reported in the main text.

3.2 Results

3.2.1 Viewing distance and vergence eye movements

In this experiment, where participants were free to move the head, vergence eye movements depended largely on participants' head movements and not on target motion as in the first experiment. Figure 8a shows an example of viewing distance and vergence traces recorded from a single participant during the 5 repetitions (trials) in the 4 viewing conditions, while Fig. 1Sa (Supplementary Information) shows the average traces across participants. Viewing distance was computed as the instantaneous difference between head
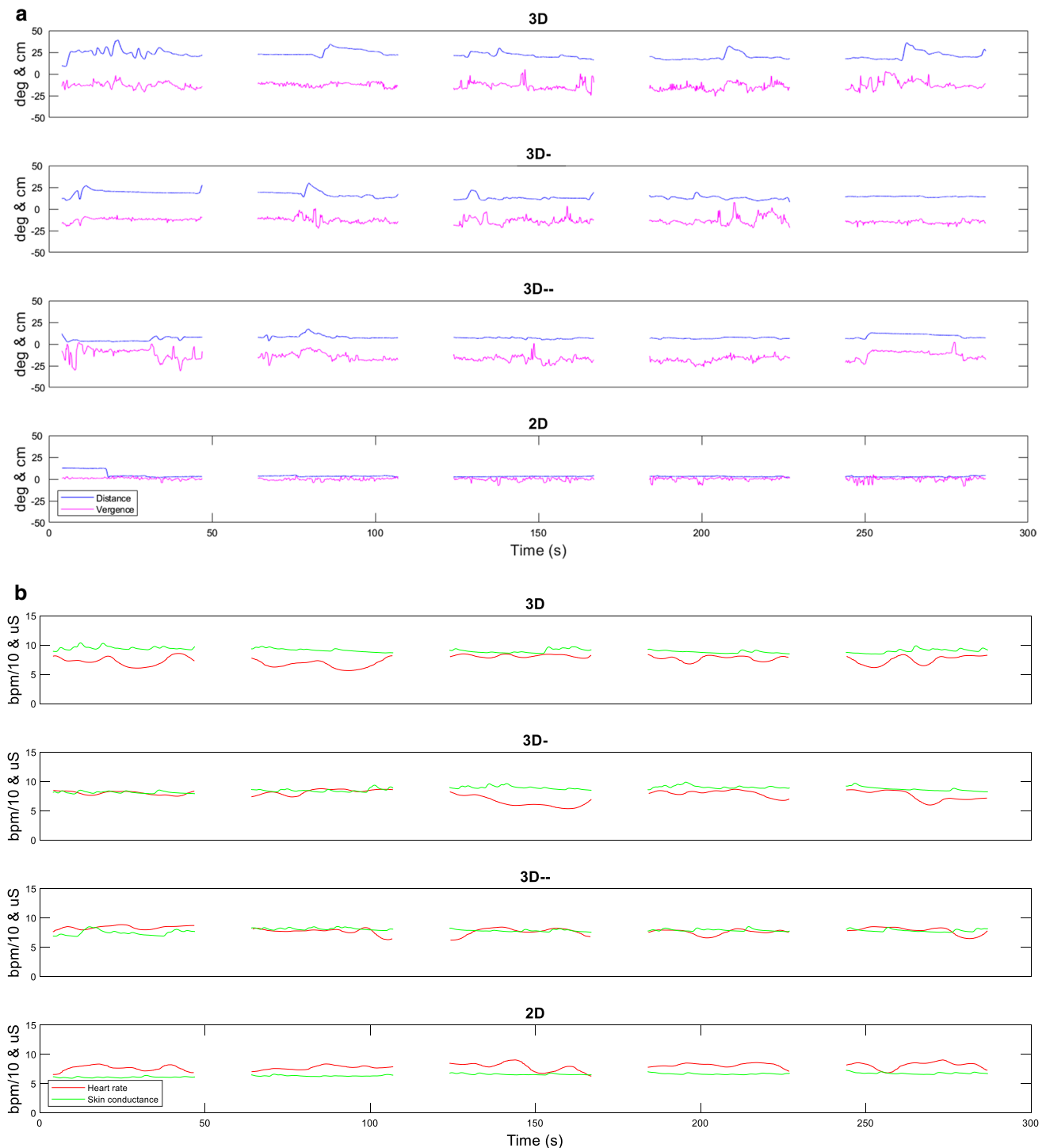
Fig. 8 Example of vergence and distance (**a**, top panels) and skin conductance and heart rate (**b**, bottom panels) recordings in the four viewing conditions (from top to bottom: 3D, 3D-, 3D--, 2D) and five repetitions (from left to right). Participants rated convergence dis-comfort verbally in correspondence of the blank periods. The visual stimulus was present throughout the session. These traces are taken from a single participant, while the average traces across participants are shown in the Supplementary Information (Fig. 1S)

pose and target pose (head-target distance). Note that distance and vergence may sometimes be decoupled, because participants may shift the gaze between objects located at different depth in the scene other than the target, or between two points on the target (e.g., the proximal equator and a pole), without moving the head. Convergence break can be

another cause of decoupling at very near viewing distance (Scheiman et al. 2003), which would be present mostly in the 3D condition, when convergence requirement is maximal. A further possible reason for decoupling will be addressed in the Discussion.

On average, distance and vergence depended on the viewing condition (Fig. 9a): participants tended to stay closer to the target when passing from the 3D to the 2D condition [main effect of viewing condition on distance, $F(1,257) = 86.226$, $p < 0.001$], and this reduction was accompanied by a decrease in vergence [main effect of viewing condition for static vergence, $F(1,257) = 225.845$, $p < 0.001$; main effect of viewing condition for dynamic vergence, $F(1,257) = 103.786$, $p < 0.001$].

3.2.2 Symptoms

Subjective discomfort due to convergence effort (CD) depended on the viewing condition (Fig. 9b), with ratings significantly decreasing when passing from the 3D to the 2D condition [main effect of viewing condition, $F(1,262) = 36.072$, $p < 0.001$]. Vergence—both static and dynamic—was a significant predictor of subjective convergence discomfort (CD), regardless of the viewing condition [static vergence, $F(1,269) = 12.745$, $p < 0.001$; dynamic vergence, $F(1,269) = 7.725$, $p < 0.006$].

The cyber-sickness symptomatology, as assessed through the SSQ_T, also changed significantly with the viewing condition [$F(1,36) = 9.738$, $p = 0.004$], indicating symptom relief when passing from 3 to 2D. This held true also for SSQ_O, i.e., the more specific ocular symptoms [$F(1,49) = 12.219$, $p < 0.001$]. Both symptoms were

significantly predicted by vergence eye movements [SSQ_T: static vergence, $F(1,35) = 6.006$, $p = 0.019$; dynamic vergence $F(1,35) = 4.307$, $p = 0.045$; SSQ_O: static vergence, $F(1,48) = 6.039$, $p = 0.018$; dynamic vergence $F(1,48) = 8.143$, $p = 0.006$].

3.2.3 Autonomic activity

Figure 8b shows an example of skin conductance and heart rate recordings from a single participant, while the corresponding average traces are shown in Fig. 1Sb (Supplementary Information). Figure 9c reports the aggregated SCL data. The effect of viewing condition on SCL was statistically significant [$F(1,257) = 36.713$, $p < 0.001$], with a progressive decrease when passing from 3 to 2D.

Regardless of visual condition and repetition, static vergence (SV), but not dynamic vergence (DV), turned out to be a significant predictor of SCL [$F(1,268) = 20.238$, $p < 0.001$]. The lack of relation between SCL, which is the expression of a tonic mechanism, and dynamic vergence is in keeping with the transient nature of dynamic vergence, whereas static vergence shares with SCL the fact of being a sustained, tonic process.

Remarkably, regardless of the viewing condition and vergence effort, SCL predicted subjective convergence discomfort trial-wise (CD: $F(1,269) = 7.390$, $p = 0.007$), and also cyber-sickness and ocular symptoms session-wise (SSQ_T: $F(1,35) = 11.864$, $p = 0.002$; SSQ_O: $F(1,10) = 5.092$, $p = 0.047$). These results have been obtained with three additional, distinct analyses because of the potential correlation among predictors.
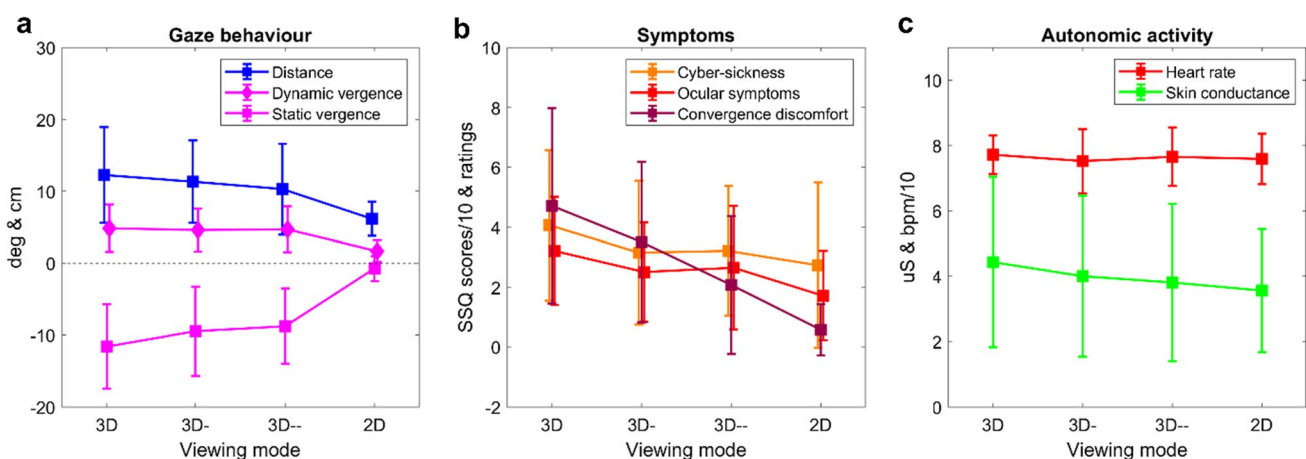

Fig. 9 Effects of viewing condition on gaze behaviour (**a**, head-target distance, static and dynamic vergence), symptoms (**b**, SSQ_T, SSQ_O, and convergence discomfort ratings) and autonomic activity (**c**, skin conductance and heart rate). SSQ total scores and bpm values have been divided by a factor 10 for graphical clarity. Data-points and error bars represent means and standard deviations across participants. In panel **b**, means and standard deviations have been computed from the median values of the raw ordinal scores across repetitions in each participant

We did not find evidence that heart rate (HR) or heart rate variability (HRV) depended on the viewing condition, repetition, repetition × viewing condition; also, neither static nor dynamic vergence predicted HR or HRV (data not shown).

3.2.4 Vergence and visual characteristics

We further explored possible influences of individual visual characteristics on vergence eye movements, which in turn may affect other variables that are dependent on vergence. To this end, we used a model where vergence, either static or dynamic, was the dependent variable, and presence of eye dominance, presence of visual correction and interpupillary distance were fixed-effects predictors. Participants were included as random intercept. However, none of these variables significantly predicted vergence, either static or dynamic (data not shown).

The results of the reported analyses are summarized in Table 3.

3.2.5 Dropping 2D

To rule out the possibility that the weight of the 2D viewing condition was predominant and thus that the results could not be replicated by considering only the viewing conditions from 3 to 3D-- (a scenario which could suggest that quasi-3D is in fact not capable to exert a substantial effect or at least that it can exert at most a minor effect), we repeated the analyses by dropping the 2D viewing condition. Any

Table 3 GLMM results for the second experiment

DV	IV	F	$df1$	$df2$	p value	Coeff	Lower	Upper
Distance	ViewingCondition	86.226	1	257	**< 0.001**	− 0.020	− 0.029	− 0.011
	Environment	0.302	4	257	0.876			
	ViewingCondition × Repetition	0.064	4	257	0.992			
Static vergence	ViewingCondition	225.845	1	257	**< 0.001**	3.354	2.379	4.330
	Repetition	0.071	4	257	0.991			
	ViewingCondition × Repetition	0.030	4	257	0.998			
Dynamic vergence	ViewingCondition	103.786	1	257	**< 0.001**	− 0.326	− 0.450	− 0.202
	Repetition	0.743	4	257	0.564			
	ViewingCondition × Repetition	0.615	4	257	0.652			
Convergence discomfort	ViewingCondition	36.072	1	262	**< 0.001**	− 1.606	− 2.163	− 1.049
	Repetition	1.033	4	262	0.391			
	ViewingCondition × Repetition	1.506	4	262	0.201			
	Static vergence	12.745	1	269	**< 0.001**	− 0.236	− 0.367	− 0.106
	Dynamic vergence	7.725	1	269	**0.006**	0.266	0.078	0.455
	Skin conductance level	7.390	1	269	**0.007**	1.043	0.287	1.798
SSQ_T	ViewingCondition	9.738	1	36	**0.004**	− 0.708	− 1.169	− 0.248
	Static vergence	6.006	1	35	**0.019**	− 0.114	− 0.209	− 0.020
	Dynamic vergence	4.307	1	35	**0.045**	0.264	0.006	0.523
	Skin conductance level	11.864	1	35	**0.002**	1.141	0.469	1.814
SSQ_O	ViewingCondition	12.219	1	49	**< 0.001**	− 0.734	− 1.155	− 0.312
	Static vergence	6.039	1	48	**0.018**	− 0.107	− 0.195	− 0.019
	Dynamic vergence	8.143	1	48	**0.006**	0.262	0.077	0.477
	Skin conductance level	5.092	1	10	**0.047**	0.458	0.008	0.907
Skin conductance level	ViewingCondition	36.713	1	257	**< 0.001**	− 0.058	− 0.111	− 0.004
	Repetition	0.146	4	257	0.965			
	ViewingCondition × Repetition	0.185	4	257	0.946			
	Static vergence	20.238	1	268	**< 0.001**	− 0.011	− 0.016	− 0.006
	Dynamic vergence	1.040	1	272	0.309	0.007	− 0.006	0.020

For each dependent variable (DV), the fixed-effects tables report the effects of the independent variables (IV), i.e. the main effect of viewing condition (ViewingCondition), the main effect of repetition (Repetition) and the interaction between viewing condition and repetition (ViewingCondition × Repetition). For each predictor, the values for the F-statistic (F), degrees of freedom (df) and significance (p value) are reported, as well as the estimated coefficient and 95% confidence interval (lower and upper) for the continuous predictors. For CD, SSQ_T and SSQ_O, the effects of static and dynamic vergence and skin conductance level were assessed through additional, distinct analyses because of the correlation among predictors. Statistically significant results are reported in bold

effect surviving the 2D drop would certificate that Quasi-3D is effective.

The results showed that the main effect of the viewing condition survived the drop of the 2D condition in several cases, notably, head-target distance [$F(1,187) = 9.771$, $p = 0.002$], static vergence [$F(1,187) = 29.994$, $p = 0.002$], convergence discomfort [$F(1,192) = 12.678$, $p < 0.001$], cyber-sickness symptoms [$F(1,24) = 8.906$, $p = 0.006$] and skin conductance [$F(1,187) = 121.446$, $p < 0.001$], whereas in two cases the effect was not statistically significant any longer, namely, dynamic vergence and SSQ_O. The main effect of repetition and the interaction viewing condition × repetition were never statistically significant. We submit that, in this scenario where the sample size was reduced by 25% (1 out of 4 experimental conditions dropped), lack of statistical significance for SSQ_O depends on a type II error. Dynamic vergence looks instead flat over the 3D-to-3D--binocularity range, almost dropping to zero at 2D (Fig. 9a), which suggests that it is sufficient a moderate degree of stereopsis to trigger a relatively constant amount of vergence eye movements. In sum, from this sort of litmus test without 2D we conclude that Quasi-3D is indeed effective in reducing vergence and tolerance, even though obviously in the 2D viewing condition the effects are maximal.

3.3 Discussion

In this experiment, participants performed a near-vision detection task under immersive condition. We found that decreasing the vergence demand by manipulating the binocular image separation improved subjective self-rated symptoms and decreased skin conductance. The results strengthened and extended considerably a finding of the first experiment, namely, that passing from 3 to 2D improved asthenopia, a finding that prompted us to visual stress.

The term visual stress is used here somewhat liberally to denote an undesired condition characterized by discomfort in the visual domain. In this study we did not attempt to split visual stress into visual, ocular and extra-ocular symptoms (Lema and Anbesu 2022), and, rather, in the second experiment we indexed operationally the subjective component of visual stress into a general cyber-sickness level (SSQ_T), a more specific ocular level (SSQ_O) and a very specific level (convergence discomfort rating). These three measures showed a consistent pattern of results, as they decreased passing from 3 to 2D through quasi-3D. Also, their changes were predicted by changes in vergence, both static and dynamic.

Importantly, visual stress was documented not only through subjective, self-rated symptoms but also through objective skin conductance recordings, which is a rather direct measure of sympathetic activity. Besides being associated to the primary cooling function, increased skin

conductance indicates increased arousal, and is considered to be a reliable real-time physiological correlate of stress (Cacioppo et al. 2000; Alberdi et al. 2016). Notice that we used SCL and not NS-SCR as a measure of tonic arousal because of the possible presence of movement artifacts in the latter, which in our experiment could be associated to active exploration of the visual stimulus, and also because the precise functional meaning of these two electrodermal components is a debated issue in psychophysiology (Boucsein et al. 2012). Thus, on the basis of the observations that SCL (i) decreased when passing from 3 to 2D, (ii) increased with increasing vergence, and (iii) predicted all three subjective symptoms, we conclude that skin conductance is a reliable index of visual stress. To our knowledge, this is the first time an influence of near vision on skin conductance is reported in the absence of a threatening stimulus, and also the first time skin conductance is associated to visual stress. These findings add to the attempts of objectively characterizing visual stress through physiological and behavioural indexes (Nahar et al. 2007; Combe and Fujii 2011; Mork et al. 2018, 2020; Lema and Anbesu 2022; Liu et al. 2022).

At variance with skin conductance, neither heart rate nor heart rate variability changed with the viewing condition, and were not predicted by convergence effort. Although heart rate was previously shown to be a poor indicator of visual stress (Naqvi et al. 2013; Mork et al. 2018), yet this decoupling between the two physiological measures may look somewhat paradoxical. However, it should be borne in mind that heart rate and skin conductance are not controlled in the same way. This is true at both the anatomical level (the former receiving both sympathetic and parasympathetic innervation and the latter receiving only sympathetic innervation) and the functional level (an increase in skin conductance can even coexist with bradycardia (Löw et al. 2015)). Such autonomic decoupling is reminiscent of the decoupling between skin conductance and pupil size during voluntary pupil size control (Eberhardt et al. 2021). We do not know what neural mechanisms caused the decoupling observed in our experiment, but it might be the result of either a lower sympathetic drive on heart rate, as compared to that on skin sweating, or an interplay between the two autonomic branches on heart activity without either one actually prevailing. Whatever the reason of this decoupling, skin conductance measures clearly showed that sympathetic arousal was higher when ocular convergence was more pronounced and when participants complained more.

4 General discussion

In this study we have manipulated the virtual camera separation in a virtual environment to produce 3D, quasi-3D (3D- and 3D--) and 2D viewing conditions, which were associated

to increasingly reduced vergence eye movements. In two experiments, spatial vision and visual stress were assessed under head-constrained and head-free conditions, with participants performing a visuo-motor and a visual detection task while wearing an HMD. Note that, although blocking head movements and blocking the visual shift are not equivalent manipulations, constraining head movements yields in fact a kind of non-immersive viewing condition, i.e., one that mimics watching a visual scene without moving the head. We preferred to label this condition as non-immersive rather than, say, poorly-immersive or the like, as it lacks a crucial ingredient of immersivity, namely, head-contingent visual shifts.

The main finding of this study is about visual stress, which was addressed especially in the second experiment. Converging evidence obtained through both subjective (asthenopia ratings, convergence discomfort ratings and SSQ scores) and objective (skin conductance) measures showed that reducing vergence eye movements results in reduced visual stress. It is difficult to tell how relevant this reduction effectively is. On the one hand, in the second experiment passing from 3D to quasi-3D (considering 3D- and 3D-- together) was not sufficient to improve cybersickness symptoms from"bad" to "concerning", at least according to a qualitative interpretation (Bimberg et al. 2020). On the other hand, when expressed as percentage, the reduction in cybersickness was ~ 12%, and the corresponding reduction of convergence discomfort was ~ 30%, which cannot be considered marginal. Thus, we think it is reasonable suggesting that, overall, the reduction of visual stress achieved by quasi-3D was remarkable. We add that, especially in the long term, even a small reduction could be appreciated by users.

Another relevant finding is the relation between visual stress and skin conductance. Physiological measures are often considered to be more reliable than self-reports, although it is not obvious that they necessarily grant a proper operationalization of psychological constructs or even simple sensations (in the present case visual stress). Remarkably, SCL predicted all visual stress symptoms, including the very specific symptoms associated to convergence discomfort (CD) and ocular symptoms (SSQ_O). Thus, this study suggests that SCL may be taken as a useful index capable to objectively signal visual stress in virtual reality contexts. Besides being an objective index, another advantage of SCL over self-reports is that it can be measured while participants perform a task, thus potentially saving the time needed to administer questionnaires. However, further research is needed to fully understand whether and under what circumstances SCL could be a valid replacement of subjective indexes of visual stress. For example, it is difficult to completely rule out that the apparently specific association of SCL with CD and SSQ_O actually depended on participants projecting a generic discomfort sensation onto reported feelings of ocular stress. Yet, even if SCL would ultimately turn out to be a mere indicator of generic discomfort, measuring skin conductance could nonetheless be a valuable tool to quantify user experience in virtual reality contexts even when threatening stimuli are not involved (e.g., Combe and Fujii 2011).

The reduction of visual stress may have depended on VAC reduction (Hoffman et al. 2008; Shibata et al. 2011) and/or vergence reduction per se (Collier and Rosenfield 2011; Cooper et al. 2011; Coles-Brennan et al. 2019). Because this study was conducted using a commercial HMD with fixed focal length, we cannot establish the relative weight of these two factors. Comparing devices with different focal lengths could in principle disentangle the two factors, but such comparison would be technically difficult to implement. However, the important consideration here is that reducing the binocular image separation produced a beneficial effect. In this sense, quasi-3D might be intended as complementary to the above-mentioned 20–20-20 rule in virtual reality: both strategies aim at relieving the convergence effort, but quasi-3D is an automatic setting and does not require users to interrupt the current task.

Reducing visual stress was accompanied by potentially negative side-effects. Indeed, in keeping with previous studies on stereoscopic displays (Jones et al. 2001; Mendiburu 2009/2012; Gao et al. 2018), and as shown in the first experiment, reducing the binocular image separation altered depth perception, in particular making objects to look farer and larger. Moreover, on intuitive grounds, it is likely that the general impression of global tridimensional appearance of the virtual environment progressively degraded passing from 3 to 2D through quasi-3D. These changes would add to other factors impacting on spatial perception when wearing an HMD (Banks et al. 2016).

Although clearly spatial distortion is an undesired effect, we submit that ultimately it may not have important negative consequences. Indeed, a main problem of objects looking farther than their objective location would manifest during multimodal interactions with objects, such as for example when making a reaching/grasping hand movement towards a target, or in spatial localization of acoustic objects: the systematic error introduced in visual spatial coordinates could result in objects being missed or mis-localized (Melmoth et al. 2007). By contrast, the spatial relationships within the visual modality would remain consistent, as the error applies to all visual objects, including body parts. Fortunately, tolerance to both audio-visual and visual-haptic discrepancies is quite large, as studies on ventriloquism and rubber hand illusion have shown (Recanzone 2009; Limanowski 2022), and sensory-motor plasticity could easily compensate visuo-motor mismatches through recalibration and realignment mechanisms (Prablanc et al. 2020), also in VR (Anglin et al.

2017). Similar arguments apply to size distortion. Thus, it is likely that these discrepancies in multimodal perception, including the sense of embodiment, simply go unnoticed or are compensated. After all, this is what happens in everyday life whenever an optical change is introduced: getting adapted to new spectacles or monocular patching is just a matter of time. The vergence system itself can undergo adaptation (Neveu et al. 2010; Kim et al. 2011). In the case of prism-altered perceived distance, adaptation involves a both a (fast) motor component and a (slower) visual component (Priot et al. 2010, 2011). Note that a given viewing mode may not need to be re-learned (Welch et al. 1993; Martin et al. 1996), thus even the transition back to real-life 3D following a long exposure to virtual Quasi-3D, which in principle might generate a negative side-effect, may turn out to be reasonably well tolerated. Thus, spatial distortion may not necessarily be a relevant side-effect, although clearly this aspect should be empirically evaluated each time according to the actual task. For example, in the fields of stereographic photography and digital cinema, the idea of training spatial vision has been explored along with the possibility of using binocular image separation to produce special effects (Mendiburu 2009/2012). Further research is needed to fully clarify the extent to which changes of 3D/2D viewing mode are well tolerated in applications outside the laboratory, for example in simulation or training contexts.

This study did not investigate long-term effects but only exposures in the order of minutes. With longer exposure times, adverse symptomatology, especially oculomotor symptoms, tends to grow with exposure times of up to 7.5 h, although the precise temporal dynamics of this process is still unclear (Dużmańska et al. 2018; Chen and Weng 2022). Splitting exposure into multiple sessions may mitigate the problem (Kennedy et al. 2000). By contrast, adaptation to spatial distortion can just improve over time (Prablanc et al. 2020). Thus, on intuitive ground, it is possible that, in the long-term, quasi-3D could both improve pros (benefits on visual stress could be even more appreciated) and reduce cons (favour acquaintance with altered multimodal mappings), as compared to the effects of short-term exposures addressed in this study.

Among the potential areas where quasi-3D could be useful, we envision virtual surgery and virtual examination of complex molecules, where both active manipulation and attentive near-vision are required. Especially in the first case, however, it would be important to ensure full adaptation to the new visuo-motor condition. Reading in a virtual world (e.g., Metaverse) is another situation where relieving convergence effort could be important. Comparing our approach to existing solutions and testing how it works in relevant and operative environments will

be crucial to promote the passage of quasi-3D from an experimental proof of concept to higher technology readiness levels (TRL).

A limitation of this study is that we did not measure participants' phoria, i.e., a latent misalignment of the two eyes, thus precluding the possibility to determine whether and how much individuals with different vergence characteristics (e.g., with more or less heterophoria) are differently susceptible to the 3D-2D manipulations that we have applied—and, in turn, whether, how much and for how long these manipulations can modify the degree of phoria (Shibata et al. 2011; Rebenitsch and Owen 2016), even possibly being exploitable as vergence training (Alvarez et al. 2010; Alvarez 2015). Similarly, we did not measure participants' accommodation, precluding the possibility of finding a relationship between accommodative capabilities (e.g., age-related changes) and the manipulations used in the present study.

5 Conclusion

In going beyond the traditional 2D / 3D dichotomy, this study has shed light on an intermediate viewing mode—quasi-3D—through which decreasing convergence effort. The main message is that quasi-3D viewing could represent a novel and easy solution to reduce visual stress when wearing an HMD, with costs in terms of spatial distortion that in many cases would likely be negligible. Although the exact balance between costs and benefits should be tested in relevant and operational environments, and with further tasks, displays, users and exposure times, we think that having established a proof of concept is a valuable starting point, especially considering that the issue of discomfort in virtual reality has not received a definitive answer yet.

Data availability Data will be made available upon reasonable request.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any conflict of interests.

References

Akiduki H, Nishiike S, Watanabe H, Matsuoka K, Kubo T, Takeda N (2003) Visual-vestibular conflict induced by virtual reality in humans. Neurosci Lett 340:197–200

Alberdi A, Aztiria A, Basarab A (2016) Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review. J Biomed Inform 59:49–75

Alvarez TL (2015) A pilot study of disparity vergence and near dissociated phoria in convergence insufficiency patients before vs. after vergence therapy. Front Human Neurosci 9:419

Alvarez TL, Vicci VR, Alkan Y, Kim EH, Gohel S, Barrett AM, Chiaravalloti N, Biswal BB (2010) Vision therapy in adults with convergence insufficiency: clinical and functional magnetic resonance imaging measures. Optom vis Sci 87:E985-1002

Anglin JM, Sugiyama T, Liew SL (2017) Visuomotor adaptation in head-mounted virtual reality versus conventional training. Sci Rep 7:45469

Banks MS, Hoffman DM, Kim J, Wetzstein G (2016) 3D Displays. Annu Rev vis Sci 2:397–435

Bimberg P, Weissker T, Kulik A (2020) On the usage of the simulator sickness questionnaire for virtual reality research. In: IEEE conference on virtual reality and 3D user interfaces (VRW). IEEE, City, pp 464–467

Boucsein W, Fowles DC, Grimnes S, Ben-Shakhar G, roth WT, Dawson ME, Filion DL (2012) Publication recommendations for electrodermal measurements. Psychophysiology 49:1017–1034

Cacioppo JT, Tassinary LG, Berntson GG (2000) Handbook of psychophysiology. Cambridge University Press, Cambridge

Chen S, Weng D (2022) The temporal pattern of VR sickness during 7.5-h virtual immersion. Virtual Reality 26:817–822

Coles-Brennan C, Sulley A, Young G (2019) Management of digital eye strain. Clin Exp Optom 102:18–29

Collier JD, Rosenfield M (2011) Accommodation and convergence during sustained computer work. Optometry 82:434–440

Combe E, Fujii N (2011) Depth perception and defensive system activation in a 3-d environment. Front Psychol 2:205

Cooper JS, Burns CR, Cotter SA, Daum KM, Griffin JR, Scheiman MM (2011) Optometric clinical practice guideline care of the patient with accommodative and vergence dysfunction. American Optometric Association, Washington

Dużmańska N, Strojny P, Strojny A (2018) Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. Front Psychol 9:2132

Eberhardt LV, Gron G, Ulrich M, Huckauf A, Strauch C (2021) Direct voluntary control of pupil constriction and dilation: Exploratory evidence from pupillometry, optometry, skin conductance, perception, and functional MRI. Int J Psychophysiol 168:33–42

Gallagher M, Ferre ER (2018) Cybersickness: a multisensory integration perspective. Multisens Res 31:645–674

Gao Z, Hwang A, Zhai G, Peli E (2018) Correcting geometric distortions in stereoscopic 3D imaging. PLoS ONE 13:e0205032

Gregory RL (2008) Emmert's Law and the moon illusion. Spat vis 21:407–420

Hoffman DM, Girshick AR, Akeley K, Banks MS (2008) Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. J vis 8(33):31–30

Jones G, Lee D, Holliman N, Ezra D (2001) Controlling perceived depth in stereoscopic images. In: Woods AJ, Bolas MT, Merritt JO, Benton SA (eds) Proceedings of the stereoscopic displays and virtual reality systems VIII, SPIE, vol 4297, City

Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG (1993) Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 3:203–220

Kennedy RS, Stanney KM, Dunlap WP (2000) Duration and exposure to virtual environments: sickness curves during and across sessions. Presence Teleoper Virtual Environ 9:10

Kim EH, Vicci VR, Granger-Donetti B, Alvarez TL (2011) Short-term adaptations of the dynamic disparity vergence and phoria systems. Exp Brain Res 212:267–278

Kim HK, Park J, Choi Y, Choe M (2018) Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl Ergon 69:66–73

Kooi FL, Toet A (2004) Visual comfort of binocular and 3D displays. Displays 25:99–108

Koulieris GA, Bui B, Banks MS, Drettakis G (2017) Accommodation and comfort in head-mounted displays. ACM Trans Graph (TOG) 36:1–11

Kramida G (2016) Resolving the vergence-accommodation conflict in head-mounted displays. IEEE Trans Visual Comput Graph 22:1912–1931

Lema AK, Anbesu EW (2022) Computer vision syndrome and its determinants: a systematic review and meta-analysis. SAGE Open Med 10:205031212211424400

Limanowski J (2022) Precision control for a flexible body representation. Neurosci Biobehav Rev 134:104401

Linton P (2020) Does vision extract absolute distance from vergence? Atten Percept Psychophys 82:3176–3195

Liu Z, Zhang K, Gao S, Yang J, Qiu W (2022) Correlation between eye movements and asthenopia: a prospective observational study. J Clin Med 11:7043

Löw A, Weymar M, Hamm AO (2015) When threat is near, get out of here: dynamics of defensive behavior during freezing and active avoidance. Psychol Sci 26:1706–1716

Lyu J, Ng CJ, Bang SP, Yoon G (2021) Binocular accommodative response with extended depth of focus under controlled convergences. J vis 21:21

Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT (1996) Throwing while looking through prisms: II. Specificity and storage of multiple gaze—throw calibrations. Brain 119:1199–1211

Melmoth DR, Storoni M, Todd G, Finlay AL, Grant S (2007) Dissociation between vergence and binocular disparity cues in the control of prehension. Exp Brain Res 183:283–298

Mendiburu B (2009/2012) 3D movie making: stereoscopic digital cinema from script to screen. Routledge, London.

Mon-Williams M, Dijkerman HC (1999) The use of vergence information in the programming of prehension. Exp Brain Res 128:578–582

Mork R, Falkenberg HK, Fostervold KI, Thorud HMS (2018) Visual and psychological stress during computer work in healthy, young females-physiological responses. Int Arch Occup Environ Health 91:811–830

Mork R, Falkenberg HK, Fostervold KI, Thorud HS (2020) Discomfort glare and psychological stress during computer work: subjective responses and associations between neck pain and trapezius muscle blood flow. Int Arch Occup Environ Health 93:29–42

Naceri A, Chellali R, Hoinville T (2011) Depth perception within peripersonal space using head-mounted display. Presence Teleoper Virtual Environ 20:254–272

Nahar NK, Sheedy JE, Hayes J, Tai YC (2007) Objective measurements of lower-level visual stress. Optom vis Sci 84:620–629

Naqvi SAA, Badruddin N, Malik AS, Hazabbah W, Abdullah B (2013) Does 3D produce more symptoms of visually induced motion sickness? In: 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), City, pp 6405–6408

Neveu P, Priot AE, Plantier J, Roumes C (2010) Short exposure to telestereoscope affects the oculomotor system. Ophthalmic Physiol Opt 30:806–815

Prablanc C, Panico F, Fleury L, Pisella L, Nijboer T, Kitazawa S, Rossetti Y (2020) Adapting terminology: clarifying prism adaptation vocabulary, concepts, and methods. Neurosci Res 153:8–21

Priot A-E, Laboissière R, Sillan O, Roumes C, Prablanc C (2010) Adaptation of egocentric distance perception under telestereoscopic viewing within reaching space. Exp Brain Res 202:825–836

Priot AE, Laboissiere R, Plantier J, Prablanc C, Roumes C (2011) Partitioning the components of visuomotor adaptation to prism-altered distance. Neuropsychologia 49:498–506

Rebenitsch L, Owen C (2016) Review on cybersickness in applications and visual displays. Virtual Reality 20:101–125

Recanzone GH (2009) Interactions of auditory and visual stimuli in space and time. Hear Res 258:89–99

Rushton S, Mon-Williams M, Wann J (1994) Binocular vision in a bi-ocular world: New-generation head-mounted displays avoid causing visual deficit. Displays 15:255–260

Scheiman M, Gallaway M, Frantz KA, Peters RJ, Hatch S, Cuff M, Mitchell GL (2003) Nearpoint of convergence: test procedure, target selection, and normative data. Optom vis Sci 80:214–225

Shibata T, Kim J, Hoffman DM, Banks MS (2011) Visual discomfort with stereo displays: effects of viewing distance and direction of vergence-accommodation conflict. Proc SPIE Int Soc Opt Eng 7863:78630P78631-78630P78639

Urvoy M, Barkowsky M, Le Callet P (2013) How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors. Ann Telecommun Annales Des Télécommunications 68:641–655

Verdelet G, Desoche C, Volland F, Farnè A, Coudert A, Hermann R, Salemme R (2019) Assessing spatial and temporal reliability of the vive System as a tool for naturalistic behavioural research. In: International conference on 3D immersion (IC3D), City, pp 1–8

Wann JP, Rushton S, Mon-Williams M (1995) Natural problems for stereoscopic depth perception in virtual environments. Vis Res 35:2731–2736

Welch RB, Bridgeman B, Anand S, Browman KE (1993) Alternating prism exposure causes dual adaptation and generalization to a novel displacement. Percept Psychophys 54:195–204

Zangemeister WH, Heesen C, Rohr D, Gold SM (2020) Oculomotor fatigue and neuropsychological assessments mirror multiple sclerosis fatigue. J Eye Move Res 13

Zheng F, Hou F, Chen R, Mei J, Huang P, Chen B, Wang Y (2021) Investigation of the relationship between subjective symptoms of visual fatigue and visual functions. Front Neurosci 15:686740

Zhou Y, Zhang J, Fang F (2021) Vergence-accommodation conflict in optical see-through display: review and prospect. Results Opt 5:100160