

Eye posture and screen alignment with simulated see-through head-mounted displays

Agostino Gibaldi

Magic Leap Switzerland GmbH,
Zürich, Switzerland



Yinghua Liu

Magic Leap, Inc., Plantation, FL, USA



Christos Kaspiris-Rousellis

Magic Leap Switzerland GmbH,
Zürich, Switzerland



Madhumitha S. Mahadevan

Magic Leap, Inc., Plantation, FL, USA



Jenny C. A. Read

Biosciences Institute,
Newcastle University Newcastle upon Tyne, UK



Björn N. S. Vlaskamp

Magic Leap, Inc., Plantation, FL, USA
Google, Seattle, WA, USA



Gerrit W. Maus

Magic Leap Switzerland GmbH,
Zürich, Switzerland



When rendering the visual scene for near-eye head-mounted displays, accurate knowledge of the geometry of the displays, scene objects, and eyes is required for the correct generation of the binocular images. Despite possible design and calibration efforts, these quantities are subject to positional and measurement errors, resulting in some misalignment of the images projected to each eye. Previous research investigated the effects in virtual reality (VR) setups that triggered such symptoms as eye strain and nausea. This work aimed at investigating the effects of binocular vertical misalignment (BVM) in see-through augmented reality (AR). In such devices, two conflicting environments coexist. One environment corresponds to the real world, which lies in the background and forms geometrically aligned images on the retinas. The other environment corresponds to the augmented content, which stands out as foreground and might be subject to misalignment. We simulated a see-through AR environment using a standard three-dimensional (3D) stereoscopic display to have full control and high accuracy of the real and augmented contents. Participants were involved in a visual search task that forced them to alternatively interact with the real and the augmented contents while being exposed to different amounts of BVM. The measured eye posture

indicated that the compensation for vertical misalignment is equally shared by the sensory (binocular fusion) and the motor (vertical vergence) components of binocular vision. The sensitivity of each participant varied, both in terms of perceived discomfort and misalignment tolerance, suggesting that a per-user calibration might be useful for a comfortable visual experience.

Introduction

“Zero optical image differences and zero alignment errors are not possible with binocular devices” (Self, 1986). Based on this assumption, the issue of vertical binocular image alignment has been long studied in the context of various types of equipment requiring binocular optics such as hand-held binoculars and binocular microscopes (Jacobs, 1943) or heads-up displays in airplane cockpits (Gold & Hyman, 1970; Gold, 1971). In a seminal literature review focusing on helmet-mounted displays, Self (1986) found that binocular vertical misalignment (BVM) can be tolerated by the visual system when it is within the range of

Citation: Gibaldi, A., Liu, Y., Kaspiris-Rousellis, C., Mahadevan, M. S., Read, J. C. A., Vlaskamp, B. N. S., & Maus, G. W. (2025). Eye posture and screen alignment with simulated see-through head-mounted displays. *Journal of Vision*, 25(1):9, 1–15, <https://doi.org/10.1167/jov.25.1.9>.

<https://doi.org/10.1167/jov.25.1.9>

Received May 14, 2024; published January 9, 2025

ISSN 1534-7362 Copyright 2025 The Authors



3.4 to 34.5 arcmin, indicating a large variability across studies. The author attributed the result to two factors. First, the visual scene content (complexity, static or dynamic, background) had a direct impact on the result. For example, when a complex background was present, the tolerable differences between the left and right images were smaller by a factor of 10 relative to a uniform background. Second, the difference in the criteria of what is tolerable for visual comfort led to tighter values compared with limits based on binocular image fusion and diplopia. This likely happens because, even if vertical misalignment becomes uncomfortable at some value, the visual system is generally still able to fuse the views from the two eyes before diplopia occurs. A more recent review by [Gavrilescu, Battista, Ibbotson, and Gibbs \(2015\)](#) evidenced similar issues, focusing attention on the great variability of subjects' tolerance to visual fatigue. These considerations showed the significance of standardizing the assessment of individual tolerance to BVM, as it may be used to predict the individual susceptibility to visual discomfort ([Zhang, Nourrit, & de Bougrenet de la Tocnaye, 2017](#)).

More recent studies have investigated the issue of BVM with respect to visual comfort on stereoscopic three-dimensional (3D) displays. [Speranza and Wilcox \(2002\)](#) exposed participants to a 3D movie viewing of roughly 35-minute duration, ensuring a relatively long exposure to BVM. In this configuration, the authors found that 15 to 20 arcmin of BVM can be tolerated before discomfort arises. [Kooi and Toet \(2004\)](#) used a different approach, focusing mainly on the stimulus onset rather than the exposure duration. Participants were asked to compare a reference unmanipulated stereoscopic image with a version of the same image with added BVM, each presented for 5 seconds, and then rate the amount of discomfort. In each trial, a different amount of BVM was added. The authors tested relatively large values of BVM (specifically, 1 and 2 prism diopters, or 34.2 and 68.4 arcmin) and demonstrated that 34.2 arcmin is well above the value that can be tolerated. [Tyler, Likova, Atanassov, Ramachandra, and Goma \(2012\)](#) performed an experiment similar to that of [Kooi and Toet \(2004\)](#) but assessed discomfort at finer levels of BVM. They found that values below 5 arcmin of BVM did not induce any symptoms of discomfort and that discomfort was perceived as moderate up to a value of 12 arcmin of BVM.

The issue of BVM has re-emerged with the widespread adoption of virtual reality (VR) and augmented reality (AR) displays; however, the results from the previous studies cannot be easily translated to AR/VR setups. Such devices exhibit two opposing issues related to BVM and discomfort. On one hand, adaptation of the eye posture can occur relatively quickly ([Kim, Vicci, Granger-Donetti, & Alvarez, 2011](#)) and is expected to mitigate symptoms of visual

discomfort for longer exposures. On the other hand, visual discomfort is known to increase cumulatively over time during continuous exposure ([Collins, Brown, Bowman, & Caird, 1991](#)). In VR setups, the issue of BVM is only seldom considered (e.g., see review in [Souchet, Lourdeaux, Pagani, & Rebenitsch, 2023](#)), and it is primarily related to individual interpupillary distances and the distance between the optical elements ([Hibbard, van Dam, & Scarfe, 2020](#)). Likewise, subsequent works in AR are mainly based on guidelines from [Self \(1986\)](#) and lack further investigation (e.g., see review in [Cakmakci & Rolland, 2006](#)). To gain a better understanding of the effects of BVM in AR setups, it is important to investigate how the visual system responds to misalignments. Eye movements are well adapted to the natural environment statistics ([Gibaldi & Banks, 2019](#); [Aizenman et al., 2023](#)) and are able to quickly adapt to sudden changes in the environment ([Kim et al., 2011](#)), particularly to vertical misalignments ([Schor & McCandless, 1995](#)). The level of vertical disparity that can be tolerated goes well above the limits suggested by the various studies, approximately 2.5° for small stimuli ([Bharadwaj et al., 2007](#)) and up to 7° for large stimuli ([Kertesz, 1981](#)). In these studies, the stimulus was shown at the central portion of the field of view, and generally the experiments were performed in a dark room. As a result, these configurations resembled the visual stimulation of a VR setup, where any vertical misalignment is coherent over the entire field of view.

In this study, we focused on the specific case of see-through AR. In see-through AR devices, the real world is always visible through the optics, and the images formed on the retinas are (by definition) binocularly aligned, whereas the augmented content is displayed using optical means that are commonly subject to some level of misalignment. In such configurations, contrary to VR, the visual system would have to implement different strategies, depending on whether the gazed object belongs to the real world or the augmented content. We devised a simulation mimicking the real environment in the background (optically aligned by definition) and an augmented environment in the foreground (which may be subject to misalignment) using a standard 3D display with shutter glasses that allowed for complete and accurate control of the BVM. Within this environment, we implemented an engaging visual search task that can be performed only if binocular vision is properly functioning. Tolerance to BVM was quantified based on task performance and perceived visual discomfort by also modulating the size of the augmented content. To better understand how the visual system responded to different levels of vertical misalignment, eye alignment was monitored during the course of the experimental session using a Nonius ([McKee and Levi 1987](#)) task.

Methods

Experimental setup

The main goal of this study was to evaluate the effects of BVM on eye posture and visual comfort in an AR environment. Although we could use an AR headset for this purpose, that would not allow for full control over the level of BVM due to various concurrent factors such as calibration errors, incorrect device donning, device deformations, and inaccuracies in the eye position. Consequently, we simulated an AR environment in a standard 3D stereoscopic setup (see [Figure 1A](#)). The setup consisted of an active 3D monitor (VG248QE;

Asus, Taipei, Taiwan) at a resolution of 1920×1080 pixels and 120-Hz frame rate (60 Hz to each eye), and a pair of shutter glasses (NVIDIA 3D Vision 2; NVIDIA, Santa Clara, CA). A chin rest was used to stabilize the participant's position at a fixed distance of 740 mm from the display over the course of the experiment. The stimuli were generated using MATLAB (MathWorks, Natick, MA) and Psychtoolbox ([Brainard, 1997](#); [Pelli, 1997](#); [Kleiner, Brainard, & Pelli, 2007](#)).

The room where the experiments were conducted was lit by incandescent lamps, resulting in ambient illuminance of 76 lux. Such a setup was able to provide a field of view of $39.5^\circ \times 22.8^\circ$ (45.6° diagonally) and an effective angular resolution of ~ 48 pixels per degree (ppd). This field of view was generally much

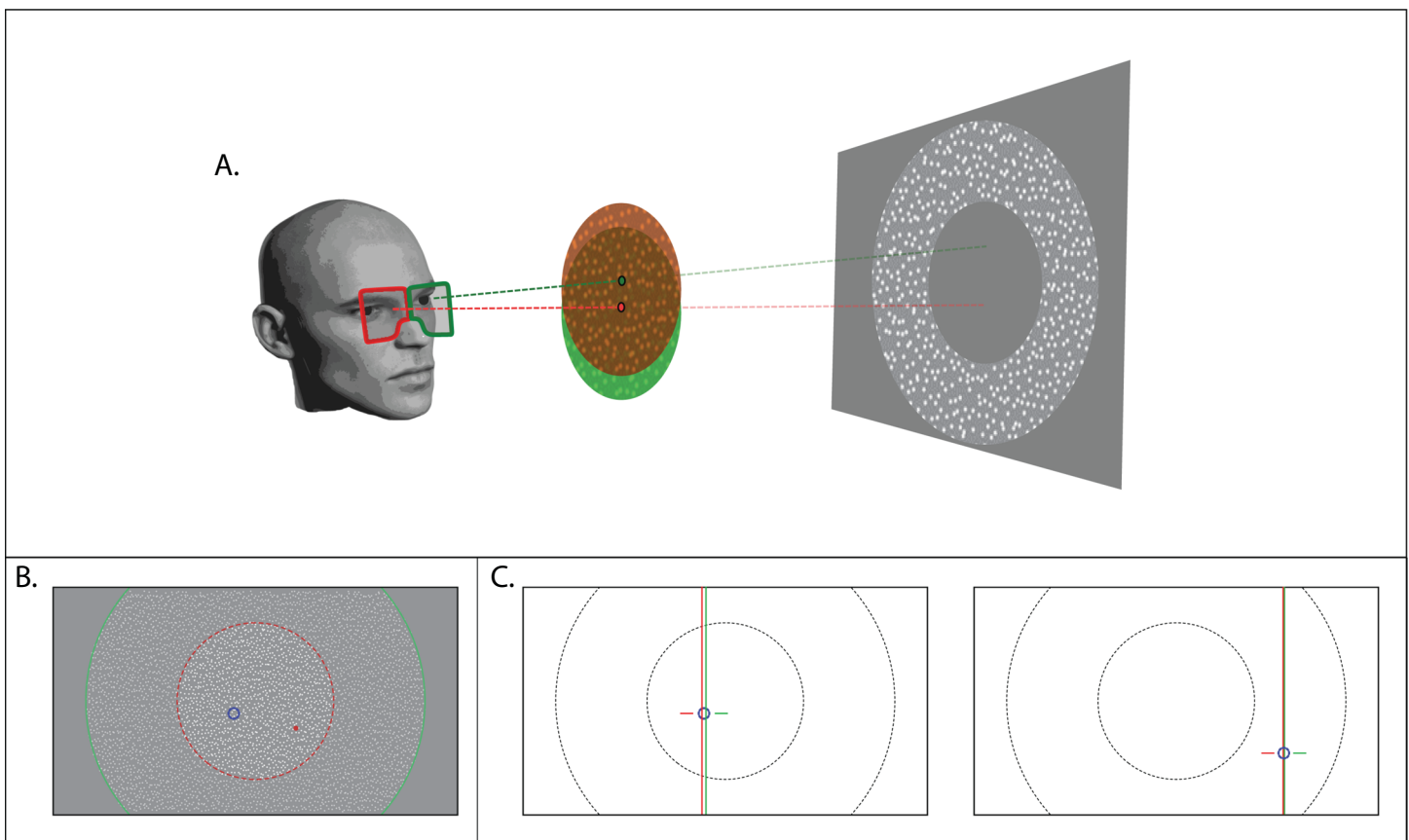


Figure 1. Experimental setup and stimuli. **(A)** The setup consisted of an active 3D display, shutter glasses, and a chin rest. The figure depicts how the visual stimulus was configured for having the background aligned and the foreground with some level of vertical misalignment. **(B)** The random dot stimulus used for the search task. The foreground (central part highlighted by the red dashed line) covered the central 8° radius of the visual field. The background (surround highlighted by the green solid line) covered a ring of maximum radius equal to 16° of visual angle. The search target (highlighted by the dotted blue line) covered a circular area with a radius of 0.5° of visual angle. The colored lines were not present during the experiment. The mouse cursor was shown as a red dot, used by the participant to move and click on the target. **(C)** The Nonius stimulus for measuring eye alignment. The Nonius stimulus appeared at the location and depth level of the last target. It consisted of a binocular vertical line (fusion lock) and two monocular horizontal line segments. The subject was asked to judge which of the horizontal lines was higher. Note that when the Nonius stimulus was shown in the foreground, the same horizontal disparity to the foreground was applied. The black circles highlighting the foreground and background areas, as well as the blue circle highlighting the search target, are only for visualization purposes and were not present during the experiment.

smaller than typical VR and pass-through AR devices (130°–150° diagonal field of view), but it is comparable to average see-through AR devices (45°–71° diagonal field of view). Considering angular resolution, our setup was able to provide better specifications than typical VR and pass-through AR devices (25–34 ppd) or see-through AR devices (30–45 ppd) (for more details, see [VRcompare, 2024](#)).

Participants

Thirty-three participants enrolled for Experiments 1 and 2. Two of the participants dropped out, and five did not meet the inclusion criteria, resulting in 26 participants (50% female, 50% male) taking part in the study. Six of the participants also took part in Experiment 3. All participants were naïve to the purpose of the study. Before the experiment was begun, visual acuity and stereopsis were measured using a Snellen chart and a Randot Stereotest, respectively. All participants had normal or corrected-to-normal visual acuity and stereopsis (≤ 40 arcsec). Their ages ranged from 25 to 40 years ($M \pm SD = 34.5 \pm 4.5$ years). All participants provided informed consent before the study started. The study was approved by an external institutional review board (WCG IRB, OHRP/FDA registration number IRB00000533) and IRB approval number 1-1479810-1 (protocol approval date December 14, 2021).

Visual search task: Where is wal-dot

The simulated AR environment was a random dot stereogram composed of two parts: a circular foreground that corresponded to the augmented content and a background annulus (encompassing the foreground) that corresponded to the real world (see [Figure 1B](#)). The foreground was subject to some amount of BVM, whereas the background was always vertically aligned. The depth of the background was fixed at the display distance. The foreground was shown with 20 arcmin of crossed horizontal disparity, thus appearing at approximately 650 mm for a nominal interpupillary distance of 64 mm. The foreground covered the central 8° radius of the visual field, and the background covered a ring of maximum radius equal to 16° of visual angle. Participants were instructed to locate a small cluster of white dots (target) popping out of either the background or the foreground. The target radius was 0.5° with an additional horizontal disparity of 10 arcmin. The horizontal disparity corresponded to distances of 690 mm and 615 mm when the target was presented in the background or the foreground, respectively. The target appeared at a random location at each trial, forcing the participant to continuously

switch between the misaligned foreground and the aligned background. The mouse pointer was shown as a red dot at the same depth as the environment it was in. Participants indicated the location of the target by moving the pointer to the search target and then clicking the mouse button. Response time was limited to 4 seconds after the target appearance. Sound feedback was provided to indicate whether the target was correctly found within the given time interval or missed. After the sound feedback, a gray screen was briefly flashed (0.5 second), and a new target appeared at a different location. The random dot pattern was created using the DrawDots command from Psychtoolbox, with anti-aliasing activated. Dots had a diameter of 0.15° (~8 pixels). To provide uniform coverage, the dots were arranged on a hexagonal grid with a density of 8 dots/deg², which resulted in a distance among dots of ~0.40°. With this density and noise, the number of dots falling within the target area varied between 4 and 9 dots. Before each trial, a different random dot pattern was generated by adding to the grid a random noise equal to half the size of the hexagon. Dots were white (350 cd/m²) against a gray background (80 cd/m²). Note that this task was specifically designed so that it can be only performed when stereopsis is properly functioning, to indicate when the visual system is not able to tolerate the BVM anymore.

Measuring eye posture

When the eyes are fixating on an object, the optical axes are generally not perfectly aligned on the target. This misalignment is referred to as fixation disparity. Fixation disparity can also be measured psychophysically using a Nonius stimulus (see [Figure 1C](#)), providing an accuracy of 1 to 2 arcmin ([Dhungel & Stevenson, 2022](#)), which is currently better than any standard vision-based eye tracker. The Nonius stimulus can be either vertical, to measure the horizontal disparity, or horizontal, to measure the vertical disparity. In the latter case, two horizontal lines (Nonius lines) are presented dichoptically (i.e., one to each eye), and the subject has to adjust their relative vertical position until the lines appear vertically aligned. The vertical distance between the two lines (i.e., Nonius offset) corresponds to the amount of vertical fixation disparity. That is the case when there are no other factors driving the eyes to a different posture. In the present task, the foreground may be subject to some BVM, and that can drive the vertical eye vergence to compensate (partially or fully) for the misalignment. Therefore, to measure the eye deviation from the physiological eye posture, we used an appropriately modified Nonius stimulus. The Nonius was inspired by [McKee and Levi \(1987\)](#) and had the following characteristics:

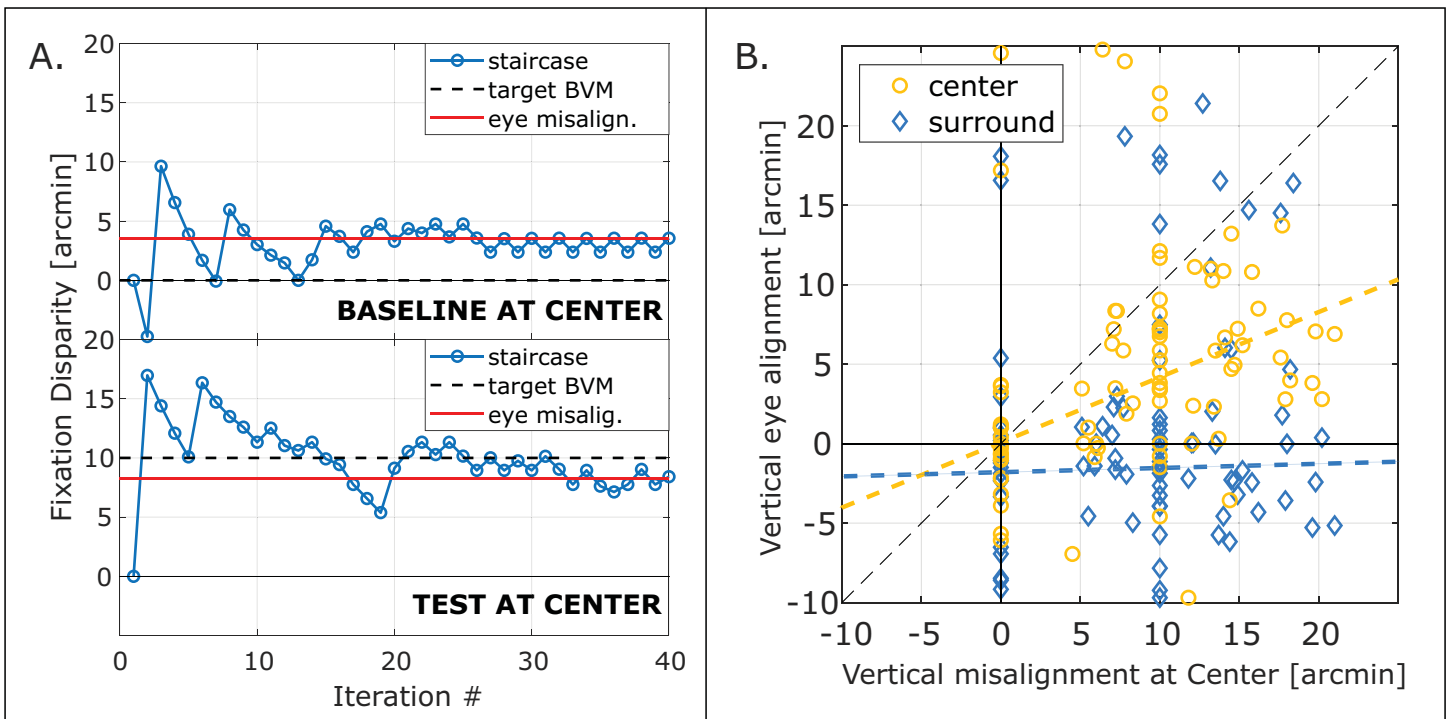


Figure 2. Eye alignment measurement. (A) Example traces of the measurement of vertical eye alignment for a single participant when looking at the central area. The measurements were performed in the baseline condition (top), when no misalignment was present, and in a test condition (bottom), where some misalignment was present (in the depicted case, 10 arcmin). The black dashed lines show the BVM that was present in the stimulus, the solid red lines show the measured values, and the blue solid lines depict the adaptive estimation procedure. (B) Overview of the eye alignment across participants. Eye alignment was measured when the participant was fixating at the central area (yellow circles) or the surround (blue diamonds). Each marker corresponds to a different session and subject, and the value was obtained as the difference between the baseline and the measured value. The dashed colored lines correspond to the best linear fit to the data. The individual data are shown in Supplementary Figure S1.

- Vertical fusion lock covered the whole field of view that was offered by the display.
- Nonius lines were 30 arcmin long and 4 arcmin thick, with a horizontal separation of 12 arcmin.
- Exposure duration was 66 ms (equal to four frames).

The vertical fixation disparity was measured during the search task by flashing the Nonius lines at regular intervals at the location of the search target. Before presenting the Nonius, the dots in the search target would turn black to signal to the participant that they should maintain their gaze on the search target. The color cue was present for 1 second to allow enough time for the participant to fixate on the target in case they were unable to locate it during the trial. In this configuration, BVM might be present in the foreground, influencing vertical eye alignment (see Figure 2A, bottom). The baseline for each participant was measured separately in a standalone task (Jaschinski, Bröde, & Griefahn, 1999), where no BVM was present (Figure 2A, top). In the standalone configuration, a black dot appeared against the gray

background for 1 second at a random location within the central area of the display (foreground). The participant was instructed to move their gaze at the location of the dot and anticipate the presentation of the Nonius that followed the dot offset. The baseline was then subtracted from the measurement during the search task to obtain the vertical eye alignment value. In the Nonius task, the participant had to report which line (left or right) was higher using the arrow keys on a keyboard. The vertical fixation disparity was estimated using the best PEST staircase method (Pentland, 1980) with 40 trials, as implemented in the Palamedes toolbox (Prins & Kingdom, 2018).

Experimental procedures

Experiment 1: Maximum vertical misalignment

During preliminary testing, we found large interindividual variability in the maximum value of BVM that could be tolerated before stereopsis was impaired. The first experiment aimed to measure this

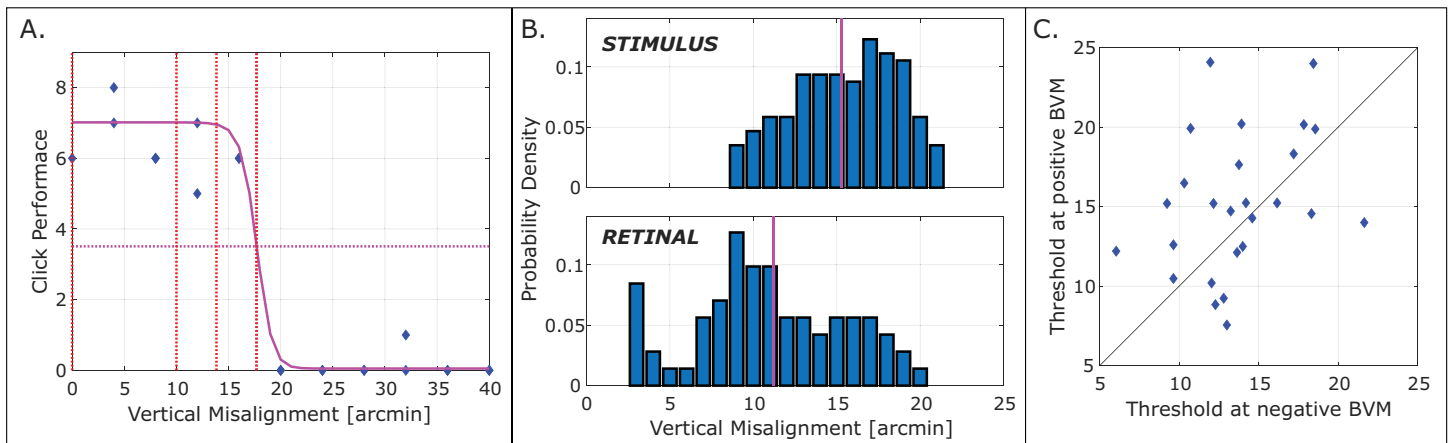


Figure 3. Maximum tolerable BVM. **(A)** Estimation of the maximum tolerable BVM for a single participant. The blue diamonds depict the search performance in targets found per minute (y -axis) at the central part of the stimulus, plotted against the corresponding BVM value. The solid pink line is the sigmoid fit to the data, and the dotted pink line corresponds to the 50% performance level used to specify the maximum tolerable BVM. The dotted red lines show the selection of the four BVM conditions that were tested in the main experiment. **(B)** Histogram of the tolerable stimulus BVM values (top) and tolerable retinal BVM (bottom) across participants. The retinal BVM is computed as the stimulus BVM minus the eye alignment measured by the Nonius test. The pink solid lines show the median of the distribution. **(C)** Scatterplot of BVM tolerance computed separately for positive (left-hyper) values of BVM, reported on the y -axis, and for negative (right-hyper) values of BVM, reported on the x -axis.

variability to use individual BVM tolerance values to tune the BVM levels tested for each participant in the subsequent experiment. For this purpose, we used the “where is wal-dot” search task, as it ensures that it can only be performed with properly functioning stereopsis. The task was split into 21 blocks of 30-second duration. During each block, a different value of BVM was applied to the central area of the stimulus. To cover a wide BVM range, we tested levels between -40 and 40 arcmin at 4-arcmin steps. The target switched position between the center and the surround at each trial throughout the session. Positive values of BVM occur when points in the left image are higher than corresponding points in the right image (left-hyper), and negative values are associated with points that are higher in the right than in the left image (right-hyper). Figure 3A provides an example for one participant of the individual performance data (targets found per minute) against the level of BVM, fitted with a sigmoid curve (Hill, 1913; Gibaldi, Barone, Gavelli, Malavasi, & Bevilacqua, 2015). The individual tolerance was estimated as the level of BVM that began to hinder stereopsis and task execution, and specifically as 50% of task performance. In the shown case, the estimated value is 18 arcmin.

Experiment 2: Eye posture and discomfort

The aim of the second experiment was to evaluate the effects of BVM on eye alignment and visual comfort during a relatively long exposure to BVM. Each session lasted approximately 30 minutes, and

participants completed four sessions in total. In each session, we tested a separate BVM condition. The BVM values were tuned for each participant based on their individual tolerance values from Experiment 1. All participants completed the zero (no BVM), 10 arcmin, and individual BVM tolerance levels. The fourth level was selected based on how close their individual tolerance was to 10 arcmin. Specifically, for a tolerance lower than 16 arcmin, the value was set at half the value; for a tolerance equal or higher than 16 arcmin, the value was set at 75% of the value (Figure 3A). Note that BVM tolerance was measured with both positive (i.e., left-hyper) and negative (i.e., right-hyper) values. However, to reduce the total duration of the experiment, we restricted the following testing to positive BVM values. Self-reported visual discomfort levels were measured for each session on a five-level scale using a subset of the visual discomfort questionnaire (VDQ) in Vinkers, Kaspiris-Rousellis, Halow, Maus, and Vlaskamp (2024). The 16-item questionnaire can be found in the Supplementary Appendix. To allow for the participant to recover from potential discomfort symptoms, each session was performed on a different day. The order of the BVM values was selected using a Latin square design to prevent potential biases in the participant responses. Each session adhered to the following procedure:

1. *Donning and verification*—The participant was positioned on the chin rest in front of the display, and the functionality of the 3D system was verified.

2. *Eye alignment baseline*—The user performed the standalone version of the eye alignment measurement. For the first five trials, large values of Nonius offset were displayed. A positive/negative audio feedback was given to the participant in case of correct/wrong response.
3. *Search task training*—The user performed a 30-second training of the “where is wal-dot” task. The time to find the target was not limited. If the participant pointed to a wrong location, a red circle highlighted the correct target position. The target kept switching between the center (foreground) and surround (background) locations.
4. *Complete task training*—Each subject completed a 2-minute training run at zero BVM, including both the search task and the eye alignment measurement. The 4-second response time limit was introduced, and the target switched between the center and the surround every 4 seconds. Before every switch, the Nonius stimulus was shown at the target location. The Nonius offset values were estimated separately for the center and surround locations using two adaptive staircases in parallel.
5. *Pre-VDQ*—After completing the training procedure, the participant was administered a digital version of the VDQ subset. The questionnaire was displayed on the 3D screen to maintain the participant position on the chin rest. The pre-VDQ was administered to capture the baseline comfort level of the participant and any potential discomfort symptoms due to the display system at zero BVM (e.g., shutter glasses flicker, reduced field of view, display brightness).
6. *Experimental session*—The experiment was equivalent to the complete training task, except for the three following parameters: (a) one of the four values of BVM was now applied to the central part (foreground) of the stimulus; (b) the duration of the session increased to approximately 10 minutes; and (c) the target switched between the center and surround approximately every 12 seconds and the switch was paced by the Nonius task.
7. *Post-VDQ*—The VDQ subset was administered again at the end of the experimental session to capture the visual discomfort that was potentially triggered by the BVM.

Experiment 3: Stimulus area effects

A potential factor influencing eye alignment and depth perception is the size of the “augmented” content (i.e., the size of the central area subject to BVM). Whereas the integration area for horizontal vergence is relatively small (approximately 5°) and horizontal vergence can be actively controlled by the subject (Popple, Smallman, & Findlay, 1998; Allison, Howard, & Fang, 2004), the integration area for vertical vergence is considerably larger (spanning a field of view up to

20°), and the subject has no control over the vergence response (Howard, Fang, Allison, & Zacher, 2000). The goal of the third experiment was to evaluate the effect of the central area size on eye alignment in the presence of fixed BVM. We used three central area radii (4°, 8°, and 12°) at a fixed BVM of 10 arcmin. The procedure was the same as in Experiment 2, with the exception of the session duration and the VDQ-related items (discomfort was not assessed due to the much shorter duration). The participants completed three sessions on separate days, one for each level of central area size, and each session lasted approximately 5 minutes.

Results

Experiment 1: Maximum vertical misalignment

Figure 3B shows the distribution of the tolerable BVM tolerance across participants. As described above, the levels of BVM present in the stimulus (top row) that could be tolerated before impairing stereopsis differed across individuals (median = 15 arcmin; range, 9–22), and that impairment was reflected in their task performance (Figure 3A). Because we were measuring eye alignment during the experiment, it was also possible to compute the effective amount of BVM at the retinas (bottom row) as the difference between the stimulus BVM and the measured eye alignment. The retinal BVM corresponds to the stimulus BVM but is reduced by the eye alignment, showing a smaller median but a larger variability (median = 11.3 arcmin; range, 1.1–18.8). Note that the tolerance was computed by the magnitude of BVM, neglecting if the misalignment was positive (left-hyper) or negative (right-hyper). Figure 3C reports the comparison between the tolerance computed for positive (y -axis) and negative (x -axis) BVM. Some asymmetry would be expected, due for example to vertical fixation disparity, although the two quantities were quite well correlated, with an average difference of -1.52 arcmin and $SD = 4.62$ arcmin.

Experiment 2: Eye posture and discomfort

Visual search performance

Figure 4A shows the search task performance for the central and surround areas, expressed as the number of targets found per minute (total targets found over the experiment duration). Task performance was measured separately for the central area (BVM was present) and surrounding area (no BVM). In both cases, the average performance decreased with increasing BVM. For the central area, the average task performance decreased slightly up to 10 arcmin of BVM (across participants, a

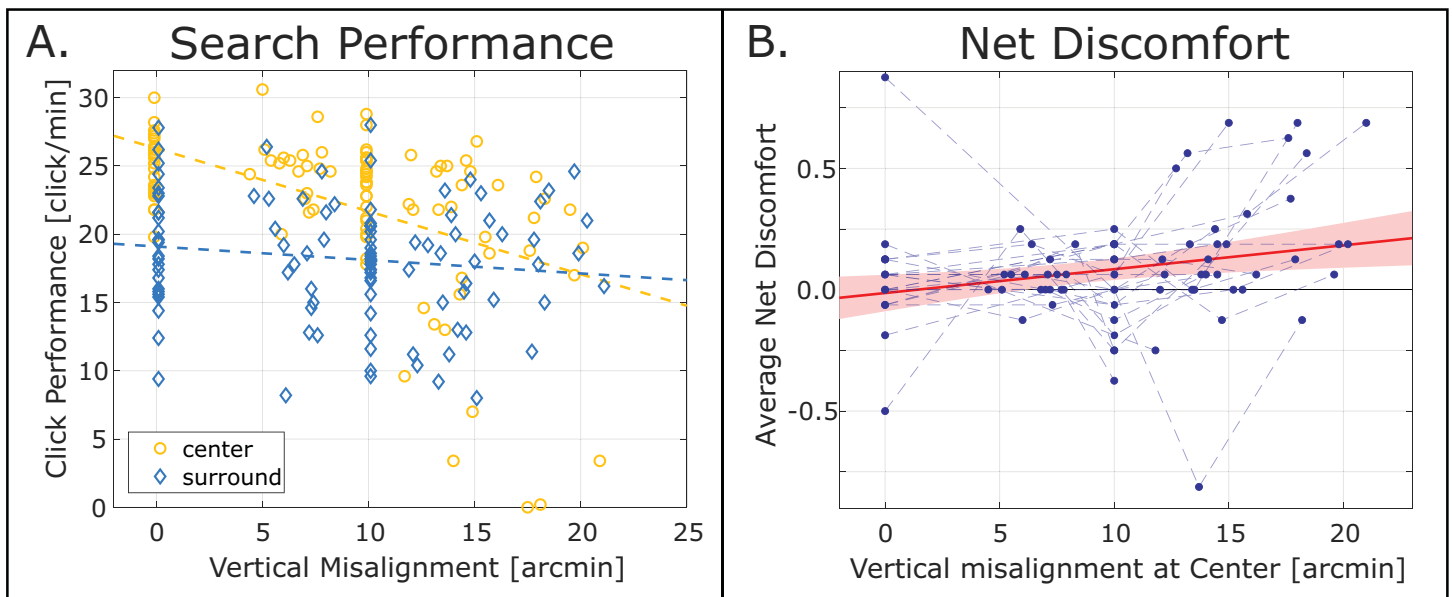


Figure 4. Search performance and net discomfort (difference between post- and pre-VDQ values) with increasing BVM. **(A)** Search task performance. Each marker corresponds to the average search performance for each participant at each session. Data are plotted against the radius of the central (misaligned) area (x-axis). Yellow circles represent the performance when the target was at the center, and blue diamonds represent the performance for the surround. The yellow and blue lines are the best linear fit to the data. **(B)** Net discomfort. Each circle represents the average net discomfort computed for a different session for one subject, as the difference between the post-VDQ and the pre-VDQ, averaged across all questions. The red line is the best linear fit to the data.

decrease of $M = 1.38 \pm 3.94$ targets/min compared with the baseline), before dropping more drastically close to the tolerance BVM values (decrease of $M = 6.64 \pm 9.01$ targets/min compared with the 10 arcmin BVM). For the surround area, performance was generally lower than the central part (difference of $M = 4.06 \pm 4.66$ targets/min), particularly at the smaller BVM values. That difference was implicit in the task, as the search area was larger and required wider eye movements to inspect. Note that, even though BVM was not present in the surrounding area, task performance still decreased subtly with the increasing central BVM (decrease of $M = 1.42 \pm 6.02$ targets/min from the baseline to 10 arcmin, and $M = 0.043 \pm 6.37$ targets/min between 10 arcmin and the tolerance value).

Eye posture

One of the primary goals of Experiment 2 was to evaluate the effect of BVM on eye posture. Figure 2A shows an example of the adaptive procedure for estimating the vertical eye alignment for one participant looking at the central area when no BVM was present (top, baseline), and with 10 arcmin of BVM (bottom, test condition). As described above, the eye posture that corresponds to the amount of vertical vergence performed by the eyes to compensate for the induced BVM can be obtained by subtracting the fixation disparity at baseline from the test condition. Figure 2B shows the measured amount of vertical vergence

across participants for the central (misaligned) and surround (aligned) areas. Evidently, when looking at the surround, the eyes tend to maintain their physiological alignment regardless of the amount of BVM present in the central area. However, when looking at the central misaligned area, the eyes tend to partially compensate for the stimulus misalignment. Specifically, the slope of the best linear fit was 0.42, indicating that the motor component of the visual system (vertical vergence) only partially took care of the BVM, and the rest was plausibly addressed by its sensory capabilities (binocular fusion). Of course that only describes the overall behavior across participants. To understand if the same approach is followed at the participant level, we also analyzed the individual data. Supplementary Figure S1 shows the data and best linear fits for each participant. Note that two of the participants (S9 and S17) were unable to perform the Nonius task correctly (pressing the same arrow throughout the task) and thus were excluded from this analysis. The most common pattern across participants was a physiological alignment of the eyes while looking at the surround, combined with different degrees of compensation for the BVM when looking at the center. For example, participants S7 and S15 fully compensated the BVM with eye posture (i.e., linear fit matches the identity line), whereas participants S1 and S6 did partially compensate for it but only for the larger values tested. It is also interesting to highlight a different strategy that was followed by participants S13 and S20, who

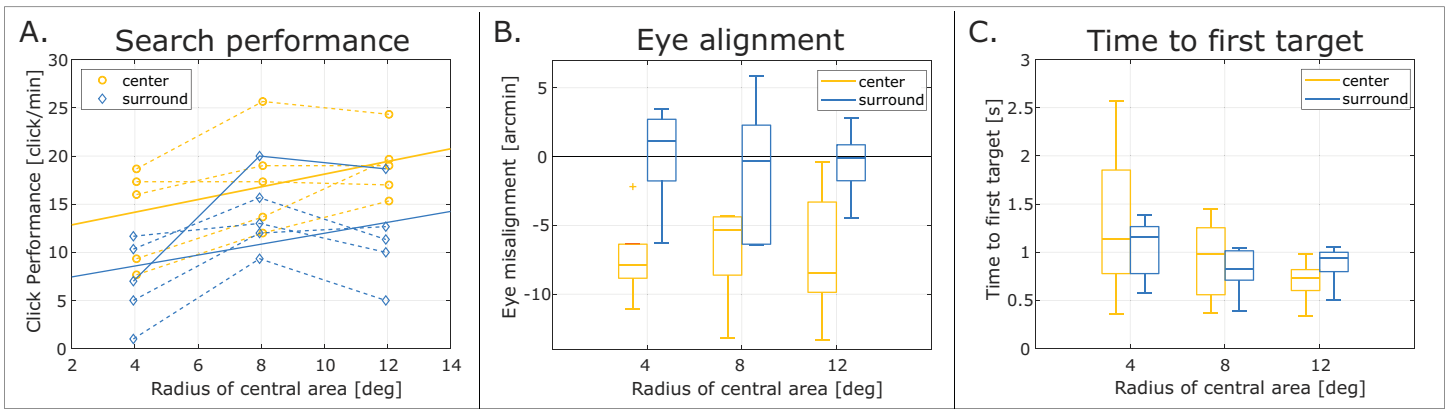


Figure 5. Search performance, eye alignment, and time to first target at varying stimulus sizes with a fixed BVM at the center of the 10 arcmin. **(A)** Search performance. Each marker is the average search performance (y -axis) computed at a different session for one subject. Data are plotted against the radius of the central (misaligned) area (x -axis). Yellow circles represent the performance when the target was at the center, and blue diamonds represent the performance for the surround. The yellow and blue lines are the best linear fit to the data. **(B)** Time to first target when looking at the misaligned central area (yellow) and at the aligned surround (blue). Each boxplot represents the time to the first target (y -axis) measured for the five subjects for different radii of the central area (x -axis). The median is marked by the horizontal tick, the box represents the first and third quartiles, and the whiskers delimit the minimum and maximum values. If present, outliers are marked by a cross. **(C)** Eye alignment measured when looking at the misaligned central area (yellow) and at the aligned surround (blue). Each boxplot represents the eye alignment (y -axis) measured for the five subjects for different radii of the central area (x -axis), and figure conventions are the same as in panel B.

exhibited partial BVM compensation together with a systematic bias in the direction of the BVM when looking at the (aligned) surround. These data would suggest that, when partially compensating for the BVM, the eyes acquire a skewed posture when also looking at the aligned surround, and that misalignment is driven by the BVM.

Visual discomfort

To estimate the effect of the increasing BVM on perceived visual discomfort, we examined the differences between the baseline (pre-VDQ) and post-VDQ responses. Figure 4B depicts the average net discomfort (the baseline score subtracted from the post-VDQ) across the questionnaire items for each participant, as a function of the level of BVM applied to the central area. Qualitatively, the net scores indicated only a slight increase in discomfort, particularly toward the tolerance BVM values for each participant. As nearly all individual responses (98%) were within one-level difference from the baseline, we decided to collapse them into a binary variable indicating whether an increase in the perceived symptoms was present at the different levels of BVM for each participant and questionnaire item. The re-encoded response was then modeled using a mixed logistic regression with the BVM as fixed effect and by-participant and by-item random intercepts and slopes (including the intercept–slope correlations). The model was fitted in R 4.3.2 (R Foundation for Statistical Computing,

Vienna, Austria) using the lme4 1.1.35.1 package (Bates, Mächler, Bolker, & Walker, 2015). For the average observer and item, an increase of 1 arcmin in BVM was associated with a 0.096 (95% confidence interval (CI), 0.016–0.171 by parametric bootstrap) increase in the log-odds of self-reporting an increase in visual discomfort symptoms (most likely one-level increase), with an intercept of -3.542 (95% CI, -4.487 to -2.680). Based on the item-level effects, the BVM estimate was driven primarily by items related to the vision subscale (e.g., unclear vision, difficulties in seeing sharp), whereas items related to head discomfort (e.g., feeling of pressure behind eyes, feeling of pressure in the head) were mainly the ones with responses closer to the zero BVM level. Note that we also expected some discomfort at zero BVM due to potential symptoms induced by the display system itself (e.g., shutter glasses flicker, reduced field of view).

Experiment 3: Stimulus area and eye posture

Visual search performance

Figure 5A shows the search task performance at different sizes of the central misaligned area. Generally, the data suggest an increase of task difficulty (lower performance) with the decreasing central area size, with a slope of 0.66 for the center and 0.57 for the surround. In the case of the surround, the result was expected given the increase of the search area (surround radius

remained constant as the central area size decreased). Counterintuitively, performance at the center showed a similar decrease, as well. One would expect that searching a 0.5° target within a 4° central area would be easier and performance would improve, as the target would always fall in the highly effective for visual search parafoveal area (Nuthmann, 2014). However, the data indicated that the task became more difficult as the central area decreased, with approximately 25% fewer targets found at 4° . Figure 5B shows the effect of the central area size on the time that lapsed from the switch between the two environments (center and surround) to when the first target was found. For each participant and session, we calculated the average time to the first target, across an average number of 20 switches. For the surround, there was no indication of an effect of the central area size on the time to first target, with an average time of approximately 1 second. However, in the central area, the time to the first target decreased systematically with the increasing central area radius, from 1.27 seconds at 4° radius to 0.74 second at 12° radius. This difference could explain the (counterintuitive) decrease in performance, which was likely due to the delay in obtaining effective stereopsis after each switch. Anecdotally, the task was also described by all participants as more difficult for the 4° area due to the time needed to perceive the depth difference in the stimulus and locate the target.

Eye posture

Figure 5B shows the variation of eye alignment with the size of the central misaligned area when participants were looking at center and the surround. Despite the differences in performance, the eye alignment did not show a trend as a function of the central area size. Also, the compensation to the BVM was qualitatively consistent with the previous experiment.

Discussion

Idiosyncrasies among real world, VR, and AR

In physiological conditions, the eyes of a person with normal stereopsis point to the same point in space. The eye posture is usually characterized by a very small angular deviation, known as fixation disparity, which is different for each individual. In this way, binocular disparity is close to zero at the fixation point and the surrounding region, well within the Panum's fusional area, allowing binocular fusion and accurate depth perception (Schor & Tyler, 1981; Wilcox & Allison, 2009). In VR, the situation can be quite different: the displays are likely subject to some misalignment, providing visual stimulation that is not coherent with

natural viewing (Banks, Read, Allison, & Watt, 2012; Souchet et al., 2023). Considering vertical misalignment, the eyes would be required to sit in an unnatural posture in order to compensate for it; otherwise, binocular vision would be impaired. The visual system is capable of quickly and effectively adapting eye posture and oculomotor control to unexpected situations (e.g., Schor & McCandless, 1995; Schor & McCandless, 1997; Kim et al., 2011). This capability is an evolutionary mechanism to maintain an effective visual performance across changes of the eye plant due not only to growing and aging but also to injuries and neurological problems (e.g., Herman, Blangero, Madelain, Khan, & Harwood, 2013; Maiello, Harrison, & Bex, 2016). In fact, large amounts of vertical disparity can be tolerated, up to 1.7° to 2.8° or even more (Kertesz, 1981; Sharma, 1992; Howard et al., 2000; Bharadwaj et al., 2007), before stereovision becomes impaired or diplopia arises. In this type of studies, the visual stimuli are bright with high contrast and are displayed in a central area of the field of view, whereas the periphery is generally not stimulated, as the experimental room is in complete darkness or covered with a black canvas. This visual stimulation is equivalent to what happens in VR, where the visual environment is uniform and vertical misalignment is coherent across the entire visual field, even if not natural. As a consequence, these adaptation mechanisms are engaged to effectively compensate for the amount of misalignment present (Schor & McCandless, 1995).

A different stimulation happens in AR, be it pass-through or see-through, where two conflicting visual environments may coexist. In pass-through devices, the real world is acquired by two cameras with parallel optical axes, and the visual information can be directly fed to the left and right displays. Therefore, just as in VR, the vertical misalignment of the displays will induce a vertical disparity pedestal that is coherent across the visual field. If the real-world cameras have some misalignment, this would add up to the real-world images only and would not affect the virtual content. As a result, real and virtual contents will be subject to different misalignments. A different possibility for pass-through AR would be to use the feed from the world cameras to first reconstruct a model of the 3D world and then render the model for the left and right displays. In that case, virtual and (reconstructed) real views would be coherent and subject only to the display misalignment. Whereas this approach would reconcile the misalignment idiosyncrasy between virtual and real contents, it comes at the price of an increased computational cost due to the reconstruction and rendering of the 3D model, thus likely increasing the motion-to-photon latency of the system and the chances of triggering visual discomfort and motion sickness (e.g., see reviews in Chang, Kim, & Yoo, 2020; Souchet et al., 2023).

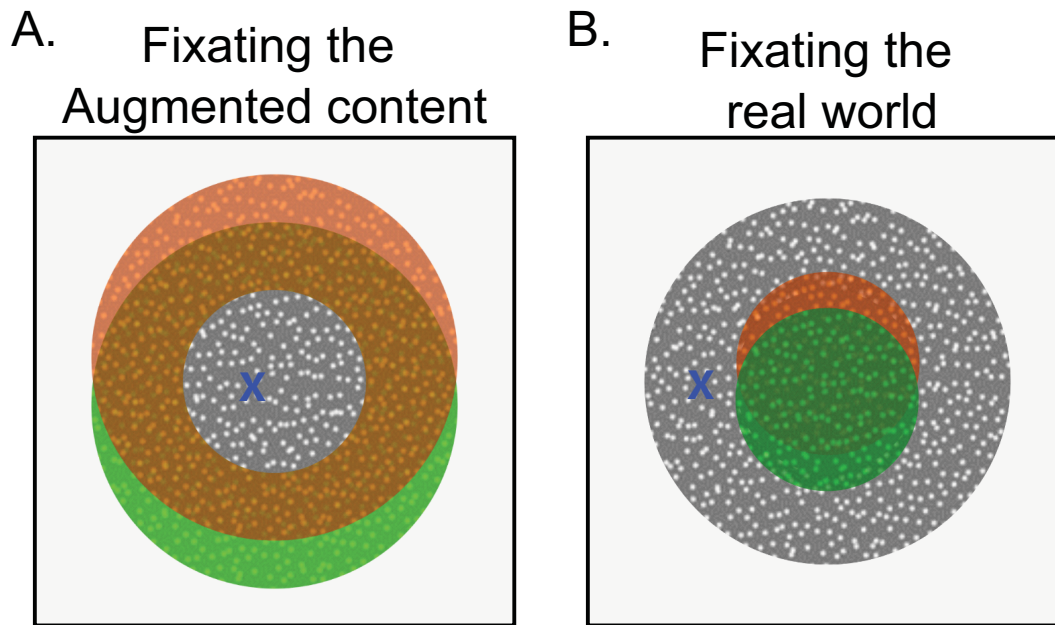


Figure 6. Representation of visual conflict in AR. The figure shows a qualitative representation of the images formed on the retinas when the eyes and images are vertically misaligned. When the fixation point (blue cross) falls on the augmented content and the eye posture would be totally compensating for the misalignment, the real world would form misaligned images on the retinas (A). In contrast, when the eyes are aiming at a point in the real world and would be physiologically aligned at the target, the augmented content would then form the misaligned images (B). For the sake of representation, content arriving to the right eye is colored in red, and green is used for the left eye, whereas if the two images are aligned they would form a gray image. Note that the two panels show the case of a complete compensation of vertical disparity by eye position, whereas the results shown in Figure 3 show that when looking at the misaligned content the compensation is only partial.

The situation in see-through AR is complementary. The real world would fall on the retinas of the user just like in natural viewing, forming geometrically aligned images and providing a natural experience to the user, and the augmented content would be subject to some misalignment with respect to the real-world images. In this configuration, when the eyes are accurately aiming at a point in the real world, the augmented content would form misaligned images on the two retinas (see Figure 6B). Besides, when the user is looking at the augmented content, a partial or total compensation of misalignment performed by eye posture would result in the real world projecting misaligned images (see Figure 6A). Therefore, irrespective of where the eyes are looking, some amount of conflict between the augmented and the real environments will always be present in the visual images for the visual system to cope with. A direct consequence is that this conflict prevents the adaptation of eye posture, thus limiting the amount of tolerable vertical misalignment.

The results of our experiments clearly highlight the effects of such conflict. In fact, the tolerable amount of vertical misalignment from the approximately 3° in VR is drastically reduced by a 10-fold factor in see-through AR, down to 15 arcmin (see Figure 3).

Vertical vergence: Ranges and characteristics

Considering this conflicting visual stimulation (see Figure 6), we can imagine two extreme approaches the visual system could use to cope with it. In one approach, the eyes would show no misalignment and stay physiologically aligned independent of the misalignment present at the central area of the visual field. In the opposite approach, the eyes would fully compensate for the misalignment even if the surround is pushing toward physiological eye alignment. In the experimental paradigm used in this work, the retinal eccentricity of the target would play a role in its detection. Because the participant is engaged in a free search, an eye tracker would be required to provide such information. Measuring the experienced retinal eccentricity of the target can provide a better assessment of the mechanisms underlying vertical disparity.

During natural vision, retinal disparities follow epipolar geometry, defined by a plane passing through the fixation point and the two centers of projection of the two eyes (Banks et al., 2012). To recreate physiological stereoscopic vision, stereoscopic displays must correctly recreate the epipolar geometry. If vertical misalignment were present in the stereoscopic display,

this would introduce non-epipolar and therefore non-natural vertical disparities. The ranges tolerated by the fusional mechanisms provide an insight into the underlying neural mechanisms.

Previous studies investigating binocular vision have shown that the largest vertical disparity triggering an effective fusional response is in the range of 3° to 6° . Out of this amount, more than 90% is compensated by vertical vergence, whereas the remaining 2% to 10% (specifically up to 8–15 arcmin) is taken care of by the sensory component of binocular fusion (Kertesz, 1981; Duwaer, 1982; Stevenson & Schor, 1997; Bharadwaj et al., 2007). Figure 2B clearly shows a similar trade-off between sensory and motor strategies to compensate for the misalignment present in the scene, but with specific differences in terms of range. On one hand, our data show that the sensory component tops at approximately 11 arcmin of vertical misalignment, in good agreement with previous data. In fact, the sensory component is not expected to change or adapt to stimulus amplitude (Luu & Abel, 2003), due to a stability of retinal correspondence (Cooper, Burge, & Banks, 2011). On the other hand, the motor fusional range is drastically reduced from 3° to 6° , as observed by these studies, compared to the 15 arcmin that our study evidenced. One plausible reason is that previous studies have used a VR-like environment, where only one environment is present, to investigate BVM. Differently, in our (simulated) AR environment, a visual conflict was constantly present in the scene, as shown in Figure 6. This conflict not only diminishes the reflexive response to compensate for vertical misalignment at fixation but also inhibits larger vertical vergence in order to avoid disrupting the binocular percept.

Another factor playing a role in vergence performance is the size of the misaligned content. Previous studies have shown that vertical vergence control is integrated over the central 20° of visual field (Howard et al., 2000), and vergence integration region increases with retinal eccentricity (Stevenson, Reed, & Yang, 1999). Considering the stimulus used in the present work, the two conflicting regions, aligned and misaligned, may produce an effect that is mediated by their size. Figure 5A clearly shows that search performance increases with the size of the central (misaligned) region. One might expect eye alignment to be driven more strongly by a larger central stimulus (Stevenson et al., 1999), resulting in an eye posture closer to a full compensation of BVM. Figure 5B shows that no effect on eye alignment is present, except for a larger variance at larger radii. Previous studies analyzed horizontal vergence responses under cue conflict conditions (Stevenson, Lott, & Yang, 1997; Sheliga, FitzGibbon, & Miles, 2007; Maxwell, Tong, & Schor, 2010), suggesting the presence of a *winner-take-all* strategy. This mechanism would help, for example, to privilege the images in the plane of fixation and

ignore the competing images that occur at other depth planes. Similarly, our results suggest that this type of strategy may also rule vertical vergence in the presence of conflicting cues. Accordingly, the decrease in search performance should have a different cause than vergence accuracy. Previous studies have shown that a larger stimulus region decreases the lag of vertical vergence (Howard et al., 2000). Figure 5C shows that, in our experiment, a smaller stimulus size negatively impacted the search performance, increasing the time to when the first target was found. This suggests that it is the speed of vergence rather than the accuracy of vergence that is having the greatest effect on task performance.

Tolerable misalignment versus discomfort

In line with prior works, an increase in the BVM was associated with an increased sensation of visual discomfort (Speranza & Wilcox, 2002; Kooi & Toet, 2004; Tyler et al., 2012). In comparison, the effect highlighted by our study can be considered rather small; however, it is worth noting that the tested BVM was restricted to a relatively short range (up to 40 arcmin) compared with the previous studies. Recall that, with our procedure, participants were unable to perform the task at larger BVM values, as the task required properly functioning stereopsis. In applications where that may not be necessary, allowing exposure to larger values of BVM could lead to the user experiencing more severe visual discomfort symptoms. Interestingly, the effect was more prominent toward individual BVM tolerance, suggesting that, even if the visual system is able to sustain more “extreme” eye postures, doing so comes at the cost of discomfort. How “extreme” the eye posture is depends on the individual tolerance values, which could potentially be used as an objective indicator of the individual susceptibility to vertical misalignment. Zhang et al. (2017) demonstrated this relationship more clearly by measuring the BVM tolerance using a VR-like setup on a standard 3D display. They tested a relatively large BVM range (up to approximately 2°) and found a strong effect on visual discomfort that was found to be related to the individual vertical fusion amplitude. As discussed above, in see-through AR, the amount of tolerable BVM is limited by the conflict between the real and the virtual content, down to a 10-fold factor. Although such an amount of BVM requires unnatural eye postures, it is still in the range of physiological vertical vergence and thus could mitigate the induced discomfort. In order to provide a numeric example, we can consider an average near-eye display with a typical eye relief (i.e., the distance between the eye and the closest optical element) of 30 mm. In the case of see-through AR displays, where the display technology mainly relies on waveguides to relay the

images, angular errors on the tilt axis translate directly to vertical angular error in the position of the active display. In this case, 1 mm of vertical shift between the display would result in $\sim 2^\circ$ of vertical misalignment, which is much larger than the average tolerance of 15 arcmin (corresponding to 0.125 mm of vertical shift). Those numbers make the case for accurate display alignment.

Static versus dynamic environment

Even if the goal of this study were to investigate the effects of vertical misalignment in see-through head-mounted devices, we decided not to use an actual head-mounted display (HMD) but rather to create a simulated environment that mimics one. The main motivation and advantage are that having the subject sitting on a chin rest in front of an external 3D display allows and grants high stimulation accuracy and stability over the course of the experiment. In the implemented setup we could simulate disparities down to 1.3 arcmin, and the chin rest allowed us to set and maintain the desired distance from the display. Also, for HMDs, perfect alignment between binocular images would not be possible. Moreover, it would be difficult to accurately measure the actual misalignment arriving on the retinas, and it would be even more difficult to maintain it over the course of the experiment due to device slippage.

Despite these advantages, this experimental setup did not allow for head movements or locomotion of the participant, thus limiting the naturalness of the experiment and its similarity to an actual use case of an HMD. Different studies that have investigated visual discomfort in VR setups (see reviews in [Chang et al., 2020](#); [Souchet et al., 2023](#)) indicate that the motion-to-photon latency as one of the main factors affecting comfort. The delay between the movements of the subject and the update of visual information is known to create a conflict between visual and vestibular streams, thus triggering symptoms equivalent to car or sea sickness (for an overview, see [Golding, 2016](#)). From this perspective, it is worth considering that in see-through AR only a part of the visual field is used to represent some virtual content, whereas the rest of the visual field is covered by the real world. Such a configuration could mitigate sickness symptoms due to the delay in the update of the *virtual* visual information, as it would affect only a limited portion of the field of view.

The present study identified a clear but limited effect of vertical misalignment on visual discomfort, triggering mainly vision symptoms. In interpreting this result, it is important to consider that the experimental design, considering the static configuration, might be implicitly limiting how discomfort manifests. The

ideal means to investigate the effect of misalignment in see-through AR would naturally be to perform a similar experiment using an HMD, which to the best of our knowledge has never been performed. For a proper design of such a study, we believe that two main factors should be taken into account. First, the design should try to mitigate possible effects due to visual discomfort stemming from other factors (e.g., amount and stability of motion-to-photon latency, calibration errors, device deformation) in order to isolate the effects of vertical misalignment. Second, the experimental procedure should include both static tasks, such as sitting at a table and performing manipulation tasks, and dynamic tasks involving locomotion, such as navigation, to investigate the possible interplay of vertical misalignment with a dynamic scene.

Conclusions

The present study investigated the effects of vertical misalignment in see-through AR and provided two main results. First, the burden of compensating for vertical misalignment is shared by the motor and sensory components of the visual system, as two aspects of the same mechanism. Second, the amount of tolerable misalignment, not just for comfortable vision but also for supporting binocular vision and depth perception, is quite limited due to the conflict between the (virtual) foreground and the (real) background. These results highlight the necessity of a system with high accuracy of display alignment or the requirement for a per-user calibration. This would ensure not only the visual comfort of the user but also the usability of the device itself and its capability to present stimuli in depth. In order to isolate the effects of vertical misalignment from other potential sources of visual discomfort, a study performed on an actual HMD rather than on an external 3D screen would be desirable.

Keywords: stereoscopic vision, head-mounted displays, vertical vergence, eye alignment, visual discomfort

Acknowledgments

Commercial relationships: A.G., Magic Leap (C); Y.L., Magic Leap (E); C.K.R., Magic Leap (E); M.S.M., Magic Leap (E); J.C.A.R., Magic Leap (C); B.N.S.V., Google (E); G.W.M., Magic Leap (E).

Corresponding author: Agostino Gibaldi.

Email: agibaldi@magicleap.com.

Address: Magic Leap Switzerland GmbH, Zürich, Switzerland.

References

- Aizenman, A. M., Koulieris, G. A., Gibaldi, A., Sehgal, V., Levi, D. M., & Banks, M. S. (2023). The statistics of eye movements and binocular disparities during VR gaming: Implications for headset design. *ACM Transactions on Graphics*, 42(1), 1–15.
- Allison, R. S., Howard, I. P., & Fang, X. (2004). The stimulus integration area for horizontal vergence. *Experimental Brain Research*, 156, 305–313.
- Banks, M. S., Read, J. C., Allison, R. S., & Watt, S. J. (2012). Stereoscopy and the human visual system. *SMPTE Motion Imaging Journal*, 121(4), 24–43.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Bharadwaj, S. R., Hoenig, M. P., Sivaramakrishnan, V. C., Karthikeyan, B., Simonian, D., Mau, K., . . . Schor, C. M. (2007). Variation of binocular-vertical fusion amplitude with convergence. *Investigative Ophthalmology & Visual Science*, 48(4), 1592–1600.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Cakmakci, O., & Rolland, J. (2006). Head-worn displays: A review. *Journal of Display Technology*, 2(3), 199–216.
- Chang, E., Kim, H. T., & Yoo, B. (2020). Virtual reality sickness: a review of causes and measurements. *International Journal of Human–Computer Interaction*, 36(17), 1658–1682.
- Collins, M. J., Brown, B., Bowman, K. J., & Caird, D. (1991). Task variables and visual discomfort associated with the use of VDTs. *Optometry and Vision Science*, 68(1), 27–33.
- Cooper, E. A., Burge, J., & Banks, M. S. (2011). The vertical horopter is not adaptable, but it may be adaptive. *Journal of Vision*, 11(3):20, 1–19, <https://doi.org/10.1167/11.3.20>.
- Dhungel, D., & Stevenson, S. (2022). Spatial–temporal contrast sensitivity of the eye alignment reflex. *Scientific Reports*, 12(1), 19480.
- Duwaer, A. (1982). Nonmotor component of fusional response to vertical disparity: A second look using an afterimage method. *Journal of the Optical Society of America*, 72(7), 871–877.
- Gavrilescu, M., Battista, J., Ibbotson, M. R., & Gibbs, P. (2015). Visual fatigue induced by optical misalignment in binocular devices: application to night vision binocular devices. In *Display Technologies and Applications for Defense, Security, and Avionics IX; and Head- and Helmet-Mounted Displays XX* (pp. 249–263). Bellingham, WA: SPIE.
- Gibaldi, A., & Banks, M. S. (2019). Binocular eye movements are adapted to the natural environment. *Journal of Neuroscience*, 39(15), 2877–2888.
- Gibaldi, A., Barone, D., Gavelli, G., Malavasi, S., & Bevilacqua, A. (2015). Effects of guided random sampling of TCCs on blood flow values in CT perfusion studies of lung tumors. *Academic Radiology*, 22(1), 58–69.
- Gold, T. (1971). Visual disparity tolerances for head-up displays. In *Electro-Optical System Design Conference 1971* (pp. 399–406). Chicago, IL: Industrial and Scientific Conference Management.
- Gold, T., & Hyman, A. (1970). *Visual requirements for head-up displays, final report. Phase I*. Washington, DC: Office of Naval Research.
- Golding, J. F. (2016). Motion sickness. *Handbook of Clinical Neurology*, 137, 371–390.
- Herman, J. P., Blangero, A., Madelain, L., Khan, A., & Harwood, M. R. (2013). Saccade adaptation as a model of flexible and general motor learning. *Experimental Eye Research*, 114, 6–15.
- Hibbard, P. B., van Dam, L. C. J., & Scarfe, P. (2020). The implications of interpupillary distance variability for virtual reality. In *2020 International Conference on 3D Immersion (IC3D)* (pp. 1–7). Piscataway, NJ: Institute of Electrical and Electronics Engineers.
- Hill, A. V. (1913). The combinations of haemoglobin with oxygen and with carbon monoxide. i. *Biochemical Journal*, 7(5), 471.
- Howard, I. P., Fang, X., Allison, R. S., & Zacher, J. E. (2000). Effects of stimulus size and eccentricity on horizontal and vertical vergence. *Experimental Brain Research*, 130, 124–132.
- Jacobs, D. H. (1943). *Fundamentals of optical engineering*. New York: McGraw-Hill.
- Jaschinski, W., Bröde, P., & Griefahn, B. (1999). Fixation disparity and Nonius bias. *Vision Research*, 39(3), 669–677.
- Kertesz, A. E. (1981). Effect of stimulus size on fusion and vergence. *Journal of the Optical Society of America*, 71(3), 289–293.
- Kim, E. H., Vicci, V. R., Granger-Donetti, B., & Alvarez, T. L. (2011). Short-term adaptations of the dynamic disparity vergence and phoria systems. *Experimental Brain Research*, 212, 267–278.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, 36(14), 1–16.
- Kooi, F. L., & Toet, A. (2004). Visual comfort of binocular and 3D displays. *Displays*, 25(2–3), 99–108.

- Luu, C. D., & Abel, L. (2003). The plasticity of vertical motor and sensory fusion in normal subjects. *Strabismus*, 11(2), 109–118.
- Maiello, G., Harrison, W. J., & Bex, P. J. (2016). Monocular and binocular contributions to oculomotor plasticity. *Scientific Reports*, 6(1), 31861.
- Maxwell, J., Tong, J., & Schor, C. M. (2010). The first and second order dynamics of accommodative convergence and disparity convergence. *Vision Research*, 50(17), 1728–1739.
- McKee, S. P., & Levi, D. M. (1987). Dichoptic hyperacuity: The precision of Nonius alignment. *Journal of the Optical Society of America A*, 4(6), 1104–1108.
- Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 342.
- Pelli, D. G. (1997). The Videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Pentland, A. (1980). Maximum-likelihood estimation: The best PEST. *Perception & Psychophysics*, 28, 377–379.
- Popple, A. V., Smallman, H. S., & Findlay, J. M. (1998). The area of spatial integration for initial horizontal disparity vergence. *Vision Research*, 38(2), 319–326.
- Prins, N., & Kingdom, F. A. A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. *Frontiers in Psychology*, 9, 1250.
- Schor, C. M., & McCandless, J. W. (1995). An adaptable association between vertical and horizontal vergence. *Vision Research*, 35(23–24), 3519–3527.
- Schor, C. M., & McCandless, J. W. (1997). Context-specific adaptation of vertical vergence to correlates of eye position. *Vision Research*, 37(14), 1929–1937.
- Schor, C. M., & Tyler, C. W. (1981). Spatio-temporal properties of Panum's fusional area. *Vision Research*, 21(5), 683–692.
- Self, H. C. (1986). *Optical tolerances for alignment and image differences for binocular helmet-mounted displays*. OH: Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base.
- Sharma, K. (1992). Vertical fusion amplitude in normal adults. *American Journal of Ophthalmology*, 114, 636–637.
- Sheliga, B., FitzGibbon, E., & Miles, F. (2007). Human vergence eye movements initiated by competing disparities: Evidence for a winner-take-all mechanism. *Vision Research*, 47(4), 479–500.
- Souchet, A. D., Lourdeaux, D., Pagani, A., & Rebenitsch, L. (2023). A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload. *Virtual Reality*, 27(1), 19–50.
- Speranza, F., & Wilcox, L. M. (2002). Viewing stereoscopic images comfortably: The effects of whole-field vertical disparity. In *Stereoscopic displays and virtual reality systems IX* (pp. 18–25). Bellingham, WA: SPIE.
- Stevenson, S., Lott, L., & Yang, J. (1997). The influence of subject instruction on horizontal and vertical vergence tracking. *Vision Research*, 37(20), 2891–2898.
- Stevenson, S., Reed, P. E., & Yang, J. (1999). The effect of target size and eccentricity on reflex disparity vergence. *Vision Research*, 39(4), 823–832.
- Stevenson, S., & Schor, C. M. (1997). Human stereo matching is not restricted to epipolar lines. *Vision Research*, 37(19), 2717–2723.
- Tyler, C. W., Likova, L. T., Atanassov, K., Ramachandra, V., & Goma, S. (2012). 3D discomfort from vertical and torsional disparities in natural images. In *Human Vision and Electronic Imaging XVII* (pp. 212–220). Bellingham, WA: SPIE.
- Vinkers, C. D. W., Kaspiris-Rousellis, C., Halow, S., Maus, G. W., & Vlaskamp, B. N. S. (2024). A visual discomfort questionnaire for use in research and applied settings. *Displays*, 83, 102737.
- VRcompare. (2024). VRcompare - The Internet's largest VR & AR headset database. Retrieved from <https://vr-compare.com/>.
- Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research*, 49(22), 2653–2665.
- Zhang, D., Nourrit, V., & de Bougrenet de la Tocnaye, J. L. (2017). 3D visual comfort assessment by measuring the vertical disparity tolerance. *Displays*, 50, 7–13.