Proceedings of the ASME 2026
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC/CIE2026
August 23-26, 2026, Houston, TX

IDETC/CIE2026-XXXX

SEEING ISN'T BELIEVING: THE VERTICAL–HORIZONTAL ILLUSION ACROSS SCREENS AND XR HEADSETS

Wenhao Yang[1,*], **Yueqing Li**[1], **Xinyu Liu**[1]

[1]Lamar University, Beaumont, TX, USA

ABSTRACT

The Vertical-Horizontal Illusion (VHI) systematically biases perceived geometric proportions, yet its behavior across modern digital devices remains understudied. This ecosystem now spans desktop monitors, smartphones, and an expanding class of extended-reality (XR) head-mounted displays (HMDs). This work introduces the Square Recognition Test (SRT), a unified cross-platform method for quantifying subjective squareness judgments and measuring device-dependent VHI biases. A within-subject study ($N = 24$) was conducted across PC, mobile, and Augmented Reality (AR) HMD conditions (flat- and curved-canvas). Results reveal a consistent tendency for participants to select slightly wider-than-tall rectangles as perceptually square across all device types, confirming the robustness of VHI in contemporary digital environments. Across all devices, PSE ranged narrowly from 1.02 to 1.04, suggesting that VHI magnitude is largely device-invariant for square judgments. Although numerical differences were observed, no statistically significant device effect was detected. To investigate mitigation strategies, two AR interface variants were evaluated, demonstrating that display geometry and rendering-surface curvature can influence perceptual stability. Curved-canvas AR rendering produced both lower bias and higher accuracy than flat-canvas AR, indicating that rendering geometry may be the more actionable lever for perceptual calibration in XR interface design.

Keywords: Visual perception, Vertical-Horizontal Illusion, geometric anisotropy, square recognition, cross-device evaluation, augmented reality, stereoscopic displays, user interface design

1. INTRODUCTION

Geometric–optical illusions constitute a well-established class of systematic distortions in human spatial perception. Among these, the *Vertical–Horizontal Illusion* (VHI) is one of the most extensively documented. The illusion refers to a robust anisotropy in perceived line length: a vertical extent is consistently judged to be longer than an objectively identical horizontal extent. This bias manifests across multiple canonical configurations, such as Figure 1, including: (1) T-shaped and L-shaped figures in which the vertical segment appears elongated [1]; (2) squares appearing taller and narrower than their true proportions; (3) the Helmholtz illusion [2], where a square composed of horizontal elements is perceived as taller than one composed of vertical elements; and (4) the Oppel–Kundt configuration [3], in which a filled interval is overestimated relative to an unfilled interval of equal length.

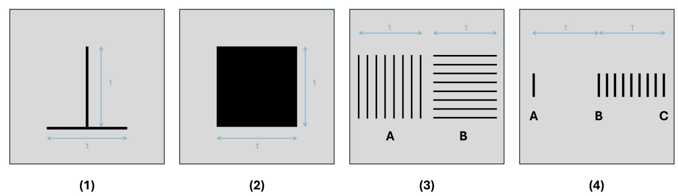


FIGURE 1: Canonical configurations demonstrating classical geometric–optical illusions related to the vertical–horizontal bias. (1) T-shaped and L-shaped figures. (2) Squares appearing taller and narrower than their true proportion. (3) The Helmholtz illusion [2]. (4) The Oppel–Kundt configuration [3].

These perceptual distortions also emerge in everyday contexts. Objects such as business cards or smartphones appear larger when oriented vertically rather than horizontally, which can be explained as Horizontal–Vertical Anisotropy [4]. In another example, observers reliably overestimate vertical distances while underestimating horizontal ones [5]. Consequently, VHI affects practical tasks such as drawing and aspect-ratio evaluation, and is particularly relevant to UI design, where vertical salience is routinely exploited to guide attention [4, 5].

Classic psychophysics attributes the VHI to anisotropies in the binocular visual field. Early work decomposed the illusion into two components: a vertical overestimation component and a bisection (dividing-line) component, whose relative strengths vary with figure configuration and viewing conditions [6, 7].

*Corresponding author: wenhao.yang@lamar.edu

Subsequent research formalized this account by showing that the horizontally elongated binocular field produces a "framing" effect that increases the apparent length of vertical extents, and that reducing binocular asymmetry reliably attenuates the illusion [8, 9]. Moreover, contemporary eye-tracking demonstrates that fixation stability and contextual framing cues modulate VHI magnitude, with greater retinal-image stability and stronger frames yielding larger effects [10]. Collectively, prior work shows that altering visual-field boundaries, surrounding frames, or gaze stability systematically amplifies or attenuates VHI [6, 8–11]. XR headsets introduce a new class of such manipulations: asymmetric FOV, lens-induced warping, and configurable rendering geometry. Whether these factors modulate VHI in the same way remains untested.

Today's human-computer interaction ecosystem spans diverse displays, from desktop monitors and smartphones to stereoscopic AR/VR head-mounted devices even in advanced industrial tasks [12–16], each differing markedly in field of view, optical distortion, pixel density, and available real-world cues. Modern HMDs use near-eye magnifying optics that introduce geometric warping and show wide variation in horizontal and vertical FOV across models [17, 18]. As interfaces shift from 2D to spatial XR, these hardware differences raise new questions about how classical perceptual biases behave under binocular, wide-FOV viewing. XR devices also introduce depth-cue inconsistencies, lens distortion, restricted FOV, and vergence-accommodation conflicts, all of which influence spatial judgments and visual stability [19]. Recent VR studies further show that optical distortions, FOV limits, and rendering characteristics systematically bias perceived geometry [17, 20]. Despite extensive research on VHI in traditional settings, its device-dependent expression in modern XR headsets and the extent to which interface design may mitigate such distortions remain largely unknown.

This work investigates two research questions: (1) *Does the Vertical Horizontal Illusion persist within modern XR head-mounted displays that use stereoscopic 3D rendering, and if so, to what extent?* and (2) *Can UI design interventions reduce the magnitude of VHI in XR interfaces?* The present work addresses these questions through three contributions:

1. A unified framework is introduced for measuring device-dependent VHI biases, including a perceptual square-matching task applied across mobile, monitor, and HMD displays in a within-subject study with 24 participants.

2. A set of AR interface variations is designed and evaluated to determine whether specific UI strategies can attenuate VHI in XR environments.

3. Empirical evidence and resulting design implications are provided to support the mitigation of perceptual anisotropies in future XR interface development.

2. RELATED WORK

2.1. The Vertical-Horizontal Illusion

Classic work by Künnapas demonstrated that the VHI comprises two separable components: vertical overestimation and bisection, each responding differently to figure orientation and segmentation [6]. He further attributed vertical overestimation to the horizontally elongated binocular visual field, which acts as a perceptual frame against which vertical extents appear longer [8]. This framing account is supported by eye-tracking evidence: VHI magnitude increases under highly stable fixation, consistent with greater reliance on visual-field boundaries for spatial comparison [10].

Contemporary models formalize these components, estimating orientation anisotropy around 6% and bisection errors around 16% [7], with illusion strength modulated by crossing/abutting junctions and local contour structure [11]. Large-scale real-world variants show even stronger anisotropies (15–25%), linked to gravitational rather than retinal reference frames [21].

2.2. Geometric Perception in XR Environments

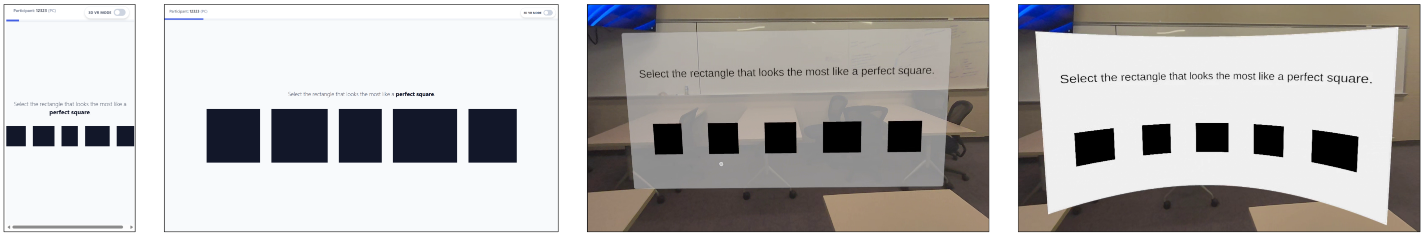
XR systems introduce geometric distortions beyond those found in conventional displays due to wide-FOV optics, stereo disparities, vergence-accommodation conflicts, and lens-induced warping. Optical analyses show that VR HMDs commonly produce magnification-driven distortions such as pincushion deformation [17], while device-level measurements highlight substantial variation in rendered horizontal and vertical FOV across headsets [18]. These physical differences are known to influence depth and size judgments: distances in early VR systems were typically underperceived, with perceived-to-actual ratios of 0.7–0.9 [22, 23], and horizontal FOV improvements yield more accurate distance estimation [24]. Related work shows that FOV asymmetries systematically alter spatial judgments in XR [25].

Given that VHI is sensitive to visual framing, these device-dependent FOV asymmetries and distortions plausibly modulate VHI strength, though this relationship has not been directly tested. Additional evidence of XR-specific geometric biases, including systematic size and taper-angle misjudgments in virtual environments [26], further motivates examining VHI under stereoscopic, wide-FOV, and lens-distorted viewing conditions.

2.3. Geometric Perception in Digital Interfaces

Perceptual distortions also appear in 2D graphical interfaces. Aspect-ratio judgments are biased by spatial context, such as surrounding rectangles, and by temporal adaptation to previously viewed aspect ratios [27]. Users also show systematic preferences for bounding-box geometry inconsistent with ground-truth proportions [28]. These biases affect layout alignment, symmetry assessment, and authoring tasks.

Device-level factors additionally modulate geometric perception. Vertical extents are robustly overestimated relative to horizontal ones [21], and sensitivity varies along visual-field meridians, with stronger anisotropy and adaptation along the horizontal meridian [29]. Eye-tracking studies indicate that stable fixation increases VHI magnitude [10], implying that fixed-focus interaction contexts may amplify biases. Across device classes, from ultrawide monitors to mobile vertical navigation patterns, interface framing may interact with classical perceptual anisotropies. This interaction underscores the need to measure VHI effects in contemporary digital ecosystems.



(1) Mobile **(2) PC Monitor** **(3) AR HMD with flat canvas** **(4) AR HMD with curved canvas**

FIGURE 2: User interface layouts across four display conditions: (1) a mobile device, (2) a PC monitor, (3) an AR head-mounted display (HMD) with a flat canvas, and (4) an AR HMD with a curved canvas.

3. METHODS

3.1. Task description

We developed a *Square Recognition Test* (SRT) system that operates across three platforms: mobile devices, personal computers (PCs), and AR HMDs. In each SRT trial, five black-filled rectangles were displayed horizontally on the screen. Their aspect ratios (width/height) were drawn from a candidate set

$$\mathcal{R} = \{0.80, 0.90, 0.95, 1.00, 1.05, 1.10, 1.20\}.$$

For each trial, five ratios were randomly sampled without replacement from $\mathcal{R}$ and presented in a randomized left-to-right order to mitigate positional bias. Thus, not every trial necessarily contained a perfect square (i.e., ratio = 1.00). Participants were instructed to select the rectangle that "looks most like a perfect square." Each session consisted of 10 trials. To minimize potential learning effects, no correctness feedback was provided at the end of each trial.

To control for size-related confounds, rectangle height was held constant across all trials; aspect ratio was manipulated solely by adjusting width. Each device employed its native primary interaction modality: keyboard-and-mouse interaction on PC, touch input on mobile devices, and hand-gesture interaction on AR HMDs. A brief training module was provided in the login interface on all platforms. In the AR condition, we implemented two types of user interfaces: a flat canvas and a curved canvas, enabling comparison of user performance across distinct spatial presentation geometries.

3.2. Devices and Participants

The study employed a within-subject design with four experimental conditions. Each participant completed all four conditions, yielding 24 participants per condition and 240 trials per condition. The presentation order of conditions was randomized and counterbalanced across participants to mitigate order effects. As shown in Figure 2, four device conditions were:

- **Mobile phone**: A 6.1" OLED touchscreen device held at roughly 35 cm.
- **PC**: Desktop setup with a 27" monitor viewed at approximately 60 cm.
- **AR HMD with a flat canvas**: A pass-through mixed-reality HMD (Meta Quest 3S) with a horizontal field of view of ~97° and a virtual focal distance of around 2 m.
- **AR HMD with a curved canvas**: The same pass-through mixed-reality HMD (Meta Quest 3S) and settings, configured with a curved rendering canvas with radius = 2 m, instead of a flat one.

All participants provided informed consent prior to the experiment. The study protocol was approved by the appropriate institutional review board [*Lamar IRB-FY26-141*].

3.3. Measures

PSE was the participant-level median chosen aspect ratio across their SRT trials, following the convention that the median of discrete choices estimates the point at which the participant is indifferent between "too tall" and "too wide". **Bias** was defined as PSE -1, with positive values indicating systematic preference for wider-than-tall rectangles as square. **Accuracy** was the proportion of trials with chosen ratio = 1.00. **Response time** (RT) was the participant-level median of per-trial durations in milliseconds. **Variability** was the within-participant SD of chosen aspect ratios.

3.4. Data Cleaning

Trials without a true-square option were excluded from accuracy analyses. However, these trials were retained for the estimation of PSE. RT outlier removal followed a per-participant IQR rule (RT > Q3 + 3× IQR was flagged). Trials with RT < 0.2 s accounted for fewer than 1% of responses across all devices.

3.5. Statistical Analysis

Normality of participant-level PSE distributions was assessed with the Shapiro-Wilk test; PSEs were non-normally distributed on at least one device (all $p < 0.05$), motivating non-parametric tests throughout. The primary between-device comparison used a Kruskal-Wallis test on PSE; pairwise device contrasts used Mann-Whitney U with effect size $r = Z/\sqrt{N}$. Descriptive uncertainty for device-level mean PSE was reported as mean ±1.96 SE (95 % CI based on the normal approximation). Accuracy CIs in figures used Wilson score intervals. For exploratory analyses of trial-order effects, we regressed mean bias on trial index within each device.

4. RESULTS

4.1. Across-Device Point of Subjective Equality (PSE)

Table 1 summarizes the mean participant-level PSEs. The values ranged from approximately 1.02 to 1.04. Across all four

TABLE 1: Participant-level performance across four device conditions. PSE = Point of Subjective Equality. Accuracy = proportion of valid trials in which the true square (1.00) was selected.

Measure	PC	Phone	AR-Flat	AR-Curved
Participants	24	24	24	24
Sample Size	240	240	240	240
PSE	1.0353	1.0438	1.0263	1.0188
Accuracy (mean)	0.412	0.323	0.252	0.406

devices, the Point of Subjective Equality (PSE) exceeded 1.00, indicating a consistent bias toward selecting rectangles that were slightly wider than tall as the most square-like. This pattern aligns with classical accounts of the VHI, in which horizontal extents tend to be perceptually underestimated relative to vertical ones.

Figure 3 shows the distribution of PSE, expressed as the median chosen aspect ratio, across the four device conditions. All conditions exhibit PSE values clustered around the unity ratio, indicating that participants generally perceived the reference shape as undistorted. However, the distributions differ in both spread and central tendency. The PC and Phone conditions show slightly broader distributions, with visible density extending toward both lower (0.9–0.95) and higher (1.1–1.15) ratios. In contrast, the AR-Flat and AR-Curved conditions display more compact distributions centered near 1.0–1.05.

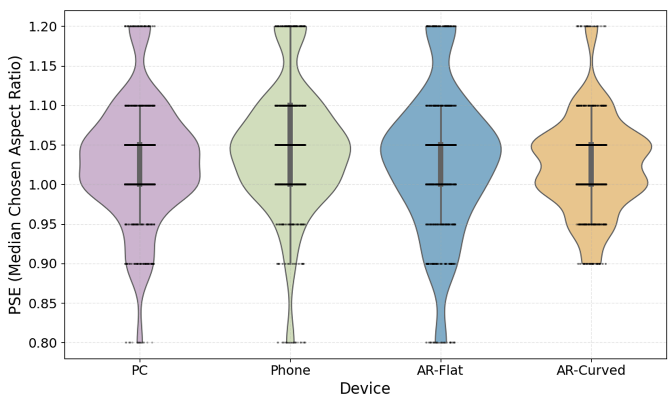


FIGURE 3: PSE Distribution by Device

Figure 4 illustrates the distribution of chosen aspect ratios for each device, showing how frequently each ratio level was selected. Across all devices, the modal choice occurs at or slightly above the objective ratio of 1.0. The ratio of 1.05 is particularly dominant for all conditions except PC, reaching the highest selection proportion in the AR-Curved (0.33) and Phone (0.30) conditions. Lower ratios (0.8–0.9) were rarely selected across all devices, each accounting for less than 8% of responses.

Although the overall shape of the distributions is similar across devices, some differences emerge in selection tendencies. The PC condition shows its peak at ratio = 1.0 (0.31), whereas the augmented reality displays (both Flat and Curved) shift their highest density toward 1.05. The AR-Curved condition also exhibits the steepest decline toward extreme values, with notably fewer selections at 1.2 compared with the Phone or AR-Flat conditions. These patterns are consistent with the PSE distributions, further supporting device-dependent differences in how participants judge aspect ratio.

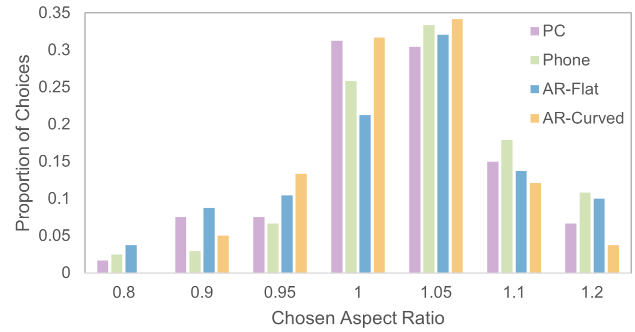


FIGURE 4: Distribution of 'most square-like' choices

4.2. Accuracy

Task accuracy varied across devices. Accuracy was computed on a per-participant basis as the proportion of trials in which the true-square option was selected (chosen ratio = 1.00), after excluding all trials without a valid correct answer. Formally,

$$\text{Accuracy} = \frac{N_{\text{correct}}}{N_{\text{t}} - N_{\text{void}}},$$

where N_{correct} is the number of correct true-square selections, N_{t} is the total number of completed trials, and N_{void} denotes trials lacking a valid correct option. For descriptive reporting, 95% Wilson confidence intervals were computed for device-level accuracy estimates.

True-square hit rate (chosen ratio = 1.00) was consistently low across devices (Table 1). A boxplot revealed the highest central tendency in the PC condition and comparatively lower accuracies in the AR conditions, see Figure 5. A repeated-measures Friedman test indicated a significant overall device effect, $\chi^2(3) = 8.216$, $p = 0.0418$, Kendall's $W = 0.119$ (small-to-moderate). Holm–Bonferroni corrected Wilcoxon signed-rank tests showed two significant pairwise differences: *PC* vs. *AR-Flat* ($p_{\text{Holm}} = 0.018$; large effect), and *AR-Flat* vs. *AR-Curved* ($p_{\text{Holm}} = 0.0336$; large effect). Other comparisons did not reach significance after correction (all $p_{\text{Holm}} \geq 0.26$). Overall, accuracy was not uniform across devices and tended to be lower in AR conditions relative to Phone and PC.

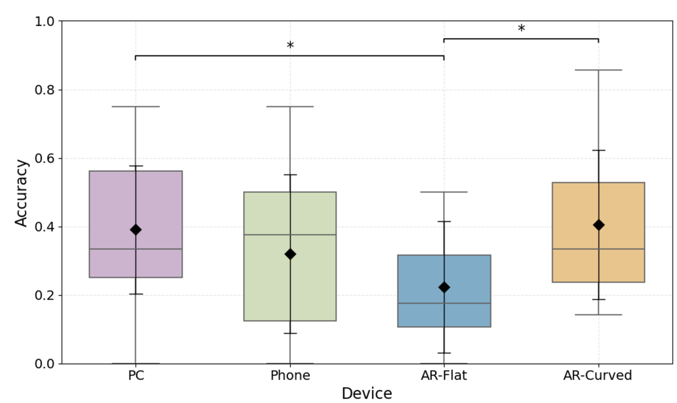


FIGURE 5: Accuracy by Device

4.3. Reaction Time

Reaction-time data were cleaned separately for each device using an interquartile-range (IQR) criterion, with a fixed lower bound of 0.2 s and an upper cutoff of $Q3 + 3 \times IQR$. Trials with RT < 0.2 s were excluded as invalid responses (e.g., accidental touches or premature inputs), whereas trials with RT $> Q3 + 3 \times IQR$ were excluded as abnormally long responses (e.g., interruptions or transient system delays). Device-specific values of $Q1$, $Q3$, IQR, and the resulting thresholds are reported in Table 2. After cleaning, the Phone condition exhibited the shortest median RT (≈1.68 s), while the other three conditions were slower, with PC showing the longest median RT.

TABLE 2: Thresholding Summary and Cleaned Descriptive Statistics (Transposed and Merged; Original n=240 per device; unit is second(s))

Metric	PC	Phone	AR-Flat	AR-Curved
Original sample size n_{before}	240	240	240	240
Cleaned sample size n_{after}	229	215	226	229
Removed count	11	25	14	11
Descriptive Statistics Before Cleaning				
Q1 (original)	2.908	1.298	2.004	2.120
Q3 (original)	8.866	3.377	7.672	7.600
Upper cutoff $Q3 + 3 \times IQR$	26.738	9.613	24.677	24.040
Descriptive Statistics After Cleaning				
Mean	6.129	2.253	5.612	5.032
Median	4.666	1.682	4.401	3.732
Q1	2.811	1.250	2.305	1.971
Q3	7.693	2.673	7.686	6.763
Minimum	0.223	0.497	0.204	0.429
Maximum	24.082	9.430	23.777	20.707

Figure 6 provides a visual summary of the cleaned reaction-time distributions across devices. As shown in the figure, the Phone condition clusters tightly around lower RTs, whereas the PC and AR displays exhibit broader and right-skewed distributions. A Friedman test revealed a robust overall device effect, $\chi^2(3) = 134.79$, $p < 10^{-28}$, Kendall's $W = .242$ (moderate). Holm–Bonferroni corrected Wilcoxon tests showed that *Phone* yielded significantly faster RTs than each of the other devices (*Phone* vs. *PC*: $p_{\text{Holm}} \approx 2.85 \times 10^{-21}$, $r \approx .708$; *Phone* vs. *AR-Flat*: $p_{\text{Holm}} \approx 6.63 \times 10^{-17}$, $r \approx .626$; *Phone* vs. *AR-Curved*: $p_{\text{Holm}} \approx 1.30 \times 10^{-16}$, $r \approx .619$). In addition, *AR-Curved* was significantly faster than *PC* ($p_{\text{Holm}} = .0027$, $r \approx .243$). The comparison between *AR-Flat* and *AR-Curved* did not reach significance after correction ($p_{\text{Holm}} = .0595$). These results indicate a clear device-dependent RT hierarchy, with Phone being the fastest, followed by AR-Curved and AR-Flat, and PC being the slowest among the tested devices.

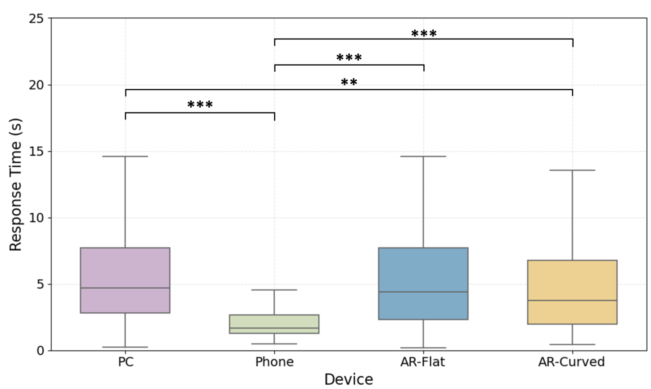


FIGURE 6: Response Time by Device (Cleaned, N=186 paired)

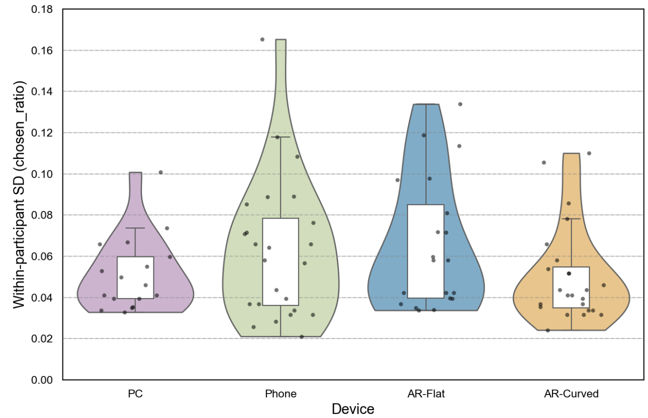


FIGURE 7: Within-participant variability across devices, quantified as the SD of chosen aspect ratios.

4.4. Variability

To assess the stability of participants' square judgments within each display condition, we quantified within-participant variability as the trial-level SD of the chosen aspect ratios. Figure 7 visualizes the resulting distributions across devices using a combined violin-box-strip plot. The violin contours illustrate the full variability distribution for each device, while the overlaid boxplots summarize the median and interquartile ranges. Individual points represent participant-level SD values, providing visibility into between-participant heterogeneity.

Across all conditions, participants exhibited relatively low SD values, indicating consistent trial-to-trial responding. However, subtle differences in distributional spread are observable across devices. In particular, responses in the AR conditions (AR-Flat and AR-Curved) show slightly broader dispersion compared to PC and Phone, suggesting marginally increased response variability in head-mounted AR displays. These descriptive patterns motivate deeper inferential analysis regarding device-related differences in perceptual stability, as reported in subsequent sections.

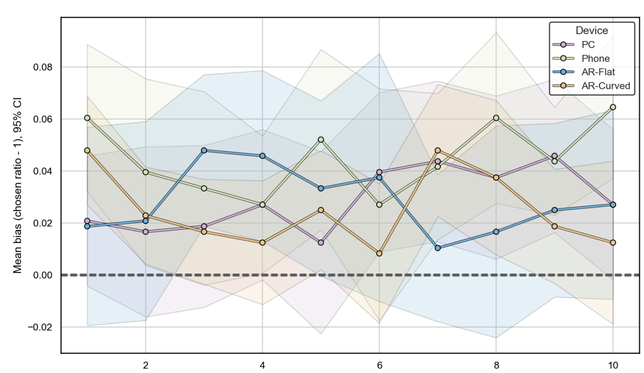


FIGURE 8: Mean bias (chosen ratio – 1) across trial index (1–10) for each device, with 95% confidence intervals.

4.5. Trial-Order Effects

To assess whether squareness judgments changed systematically over the 10-trial sequence, we examined trial-order effects on bias (chosen ratio – 1). As shown in Figure 8, mean bias trajectories were largely flat and strongly overlapping across devices, with only small, non-systematic fluctuations across trial index. Participant-level linear slopes (Figure 9), estimated by regressing bias on centered trial index within each participant-by-device block, clustered around zero for all devices. Moreover, none of the slope estimates approached statistical significance, suggesting minimal learning, fatigue, or short-term perceptual recalibration during the session. Together, these findings indicate that the device-related differences reported elsewhere are unlikely to be driven by trial-order drift.

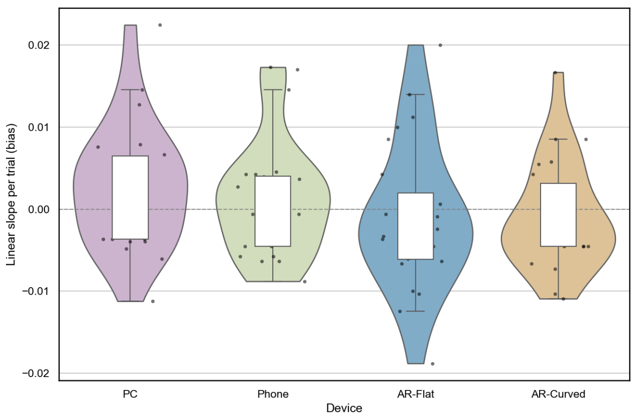


FIGURE 9: Participant-level linear slopes of bias (chosen ratio – 1) as a function of centered trial index, shown per device. Each violin illustrates the across-participant slope distribution; the overlaid box summarizes the median and interquartile range, and jittered dots show individual participants.

5. DISCUSSION

5.1. Robustness of VHI-Consistent Bias Across Devices

Across all tested platforms, participants consistently exhibited PSE values greater than 1.0, demonstrating that rectangles must be slightly wider than tall to appear perceptually square. This finding replicates classical psychophysical observations of VHI magnitude (typically 3-10%) and demonstrates its persistence in modern digital environments. Importantly, the illusion emerges even in a minimalistic judgment task devoid of contextual cues, suggesting that VHI reflects a deeply internalized spatial prior rather than a context-dependent perceptual artifact.

5.2. Interpreting the Absence of Significant Device Differences

Although numerical differences were observed (Phone > PC > AR), none achieved statistical significance. Three interpretations are plausible: (1) the VHI effect may be genuinely device-invariant for square judgments; (2) the current study may be underpowered to detect small-medium device effects; or (3) device-dependent distortions may be masked by substantial individual variability. Given prior evidence that perceived geometry is modulated by field-of-view asymmetries, binocular disparity, and rendering characteristics in XR systems, it is unlikely that devices have no influence at all. Instead, these results highlight the need for higher-precision psychophysical methods and larger samples.

5.3. Implications for Interface and XR System Design

Our results point to several practical design implications. Users tended to perceive slightly wider rectangles as more square, suggesting that nominally square UI elements may require perceptual adjustment. In AR, lower accuracy and higher variability indicate that optical distortion and viewing conflicts can exacerbate VHI. Notably, AR HMDs with a curved rendering canvas produced lower bias and more stable judgments than flat-canvas systems. These findings suggest that curvature-aligned rendering may help reduce VHI and improve perceptual fidelity in XR interfaces.

5.4. Stability of Behavior Across Task Repetitions

Trial-order analyses reveal no substantial learning or fatigue effects, indicating that the SRT task provides a stable measurement of subjective squareness within short sessions. This stability supports the feasibility of integrating SRT-like tests into practical UI evaluation workflows or automated calibration pipelines.

6. CONCLUSION

This study introduces the Square Recognition Test as a unified and lightweight method for quantifying subjective squareness perception across contemporary digital display platforms. By evaluating performance on PC, smartphone, and AR head-mounted displays, we demonstrate that the Vertical-Horizontal Illusion consistently influences geometric judgments in modern interface environments. Participants across all devices showed a systematic tendency to judge slightly wide rectangles as perceptually square, reinforcing the persistence of VHI beyond traditional laboratory contexts.

Although descriptive differences emerged between device categories, particularly between flat and curved AR canvases, no statistically significant device-dependent effects were detected in the current dataset. These results suggest that VHI functions as a robust perceptual prior that generalizes across diverse viewing conditions and interaction modalities. At the same time, the observed trends motivate future investigations employing larger sample sizes, higher-resolution psychophysical methods, and device-specific modeling to detect potentially subtle display-induced distortions.

Practically, the findings support the integration of perceptually calibrated aspect-ratio adjustments into UI design workflows and XR rendering systems. The curved-canvas condition already demonstrates that the surface curvature can shift PSE closer to veridical and reduce judgment variability, without any change to the underlying content. This indicates that perceptual improvements in XR systems can be achieved not only through hardware advances, but also through software-level adjustments, such as geometry-aware rendering aligned with user perception.

ACKNOWLEDGMENTS

Thanks to all of our participants.

REFERENCES

[1] Taylor, Crystal Marie. "Visual and haptic perception of the horizontal-vertical illusion." *Perceptual and motor skills* Vol. 92 No. 1 (2001): pp. 167–170.

[2] Helmholtz, H von et al. *Handbuch der physiologischen Optik*. Vol. 3. Leipzig Voss (1867).

[3] Wackermann, Jiří. "The Oppel-Kundt illusion." *Oxford compendium of visual illusion* (2017): pp. 303–307.

[4] Dopkins, Stephen and Galyer, Darin. "Horizontal–vertical anisotropy with respect to bias and sensitivity." *Journal of Vision* Vol. 20 No. 7 (2020): pp. 1–1.

[5] Hahnel-Peeters, Rebecka K, Idoine, Jessica L, Jackson, Russell E and Goetz, Aaron T. "Is the vertical-horizontal illusion a byproduct of the environmental vertical illusion?" *Evolutionary Psychology* Vol. 18 No. 4 (2020): p. 1474704920961953.

[6] K"unnapas, Teodor M. "An analysis of the "vertical-horizontal illusion."." *Journal of Experimental Psychology* Vol. 49 No. 2 (1955): pp. 134–140. DOI 10.1037/h0045229.

[7] Mamassian, Pascal and de Montalembert, Marie. "A simple model of the vertical–horizontal illusion." *Vision Research* Vol. 50 No. 10 (2010): pp. 956–962. DOI 10.1016/j.visres.2010.03.005.

[8] K"unnapas, Theodor M. "The vertical-horizontal illusion and the visual field." *Journal of Experimental Psychology* Vol. 53 No. 6 (1957): pp. 405–407. DOI 10.1037/h0048637.

[9] Prinzmetal, William and Gettleman, Laurie. "Vertical-horizontal illusion: One eye is better than two." *Perception & Psychophysics* Vol. 53 (1993): pp. 81–88. DOI 10.3758/BF03211717.

[10] Chouinard, Philippe A., Peel, Hayden J. and Landry, Oriane. "Eye-Tracking Reveals that the Strength of the Vertical-Horizontal Illusion Increases as the Retinal Image Becomes More Stable with Fixation." *Frontiers in Human Neuroscience* Vol. 11 (2017): p. 143. DOI 10.3389/fnhum.2017.00143.

[11] Mikellidou, Kyriaki and Thompson, Peter. "The vertical-horizontal illusion: Assessing the contributions of anisotropy, abutting, and crossing to the misperception of simple line stimuli." *Journal of Vision* Vol. 13 No. 8 (2013): p. 7. DOI 10.1167/13.8.7.

[12] Yang, Wenhao, Xiao, Qinqin and Zhang, Yunbo. "HA R 2 bot: A human-centered augmented reality robot programming method with the awareness of cognitive load." *Journal of Intelligent Manufacturing* Vol. 35 No. 5 (2024): pp. 1985–2003.

[13] Yang, Wenhao, Bai, Shi and Zhang, Yunbo. "RADAR: Robotics Assembly by Demonstration via Augmented Reality." *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*: pp. 7063–7070. 2024. IEEE.

[14] Chu, Chih-Hsing and Chen, Shau-Min. "Collaborative industrial product customization in real use environments through augmented reality: A case study of machine tool coolers." *Robotics and Computer-Integrated Manufacturing* Vol. 99 (2026): p. 103216.

[15] Chu, Chih-Hsing, Lai, Cong-Yun and Shih, Chih-Husan. "Computer-aided manufacturing in reality: augmented reality-enabled precision motion planning for robotic repair of complex shapes." *Journal of Manufacturing Systems* Vol. 85 (2026): pp. 193–206.

[16] Wu, Duidi, Zhao, Qianyou, Shen, Yuliang, Li, Junlai, Zheng, Pai, Qi, Jin and Hu, Jie. "A mixed reality-assisted human-to-robot skill transfer approach for contact-rich assembly via visuomotor primitives." *Robotics and Computer-Integrated Manufacturing* Vol. 99 (2026): p. 103208.

[17] Tong, Jonathan, Allison, Robert S and Wilcox, Laurie M. "Optical distortions in VR bias the perceived slant of moving surfaces." (2020): pp. 73–79.

[18] "HMD Geometry Database." (2026). URL https://risa2000.github.io/hmdgdb/.

[19] Feld, Nico and Pointecker, Fabian. "Perceptual Issues in Mixed Reality: A Developer-oriented Perspective on Video See-Through Head-Mounted Displays." (2025)URL https://www.researchgate.net/...

[20] Guan, Phillip, Mercier, Olivier, Shvartsman, Michael and Lanman, Douglas. "Perceptual Requirements for Eye-Tracked Distortion Correction in VR." *SIGGRAPH '22* (2022)URL https://dl.acm.org/doi/fullHtml/10.1145/3528233.3530699.

[21] Durgin, Frank H, Kwok, Chung Sze, Becker, Katelyn M and Phyu, Ya Min. "Large Bias in Matching Small Horizontal and Vertical Extents Separated in Depth in the Real World Is Similar for Upright and Supine Observers." *Vision* Vol. 9 No. 1 (2025): p. 11. DOI 10.3390/vision9010011.

[22] Renner, Rebekka S., Velichkovsky, Boris M. and Helmert, Jens R. "The perception of egocentric distances in virtual environments–A review." *ACM Computing Surveys* Vol. 46 No. 2 (2013): pp. 23:1–23:40. DOI 10.1145/2543581.2543590.

[23] Kelly, Jonathan W. "Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics." Preprint (2022). Accessed 2026.

[24] Kelly, Jonathan W., Doty, Taylor A., Ambourn, Morgan and Cherep, Lucia A. "Distance Perception in the Oculus Quest and Oculus Quest 2." *Frontiers in Virtual Reality* Vol. 3 (2022): p. 850471. DOI 10.3389/frvir.2022.850471.

[25] Masnadi, Sina, Pfeil, Kevin P., Sera-Josef, Jose-Valentin T. and LaViola Jr., Joseph J. "Effects of Field of View on Egocentric Distance Perception in Virtual Reality." *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2022. ACM, New York, NY, USA. DOI 10.1145/3491102.3517548.

[26] Zhang, Jian, Johal, Wafa and Knibbe, Jarrod. "Illusion Spaces in VR: The Interplay Between Size and Taper Angle Perception in Grasping." *CHI Conference on Human Factors in Computing Systems* (2025)DOI 10.48550/arXiv.2504.05791.

[27] Dickinson, J. Edwin, Green, Robert J., Harkin, Giorgia M., Tang, Matthew F. and Badcock, David R. "A new visual illusion of aspect-ratio context." *Vision Research* Vol. 165 (2019): pp. 80–83. DOI 10.1016/j.visres.2019.10.003.

[28] Strafforello, Ombretta, Kayhan, Osman S., Inel, Oana, Schutte, Klamer and van Gemert, Jan. "Aligning Object Detector Bounding Boxes with Human Preference." *arXiv* (2024)URL https://arxiv.org/html/2408.10844v1.

[29] Lee, Hsing-Hao and Carrasco, Marisa. "Visual adaptation stronger at the horizontal than the vertical meridian: Linking performance with V1 cortical surface area." *Proceedings of the National Academy of Sciences* Vol. 122 No. 29 (2025): p. e2507810122.