

Capabilities and limitations of the Apple Vision Pro as a portable psychophysics tool: Replicating crowding and polar angle asymmetries

Kamilla Volkova

Trinity College, Hartford, CT, USA



Kaila Dowd

Pacific Northwest National Laboratory,
Richland, WA, USA



Michael A. Grubb

Trinity College, Hartford, CT, USA



Many classic questions in vision science require presenting stimuli at precisely controlled locations in the visual field, a demand that has traditionally been met with specialized laboratory setups involving chinrests, high-resolution displays, and commercial eye trackers. Recent advances in portable mixed-reality hardware raise the possibility that some forms of visual psychophysics could be conducted on consumer devices outside these fixed laboratory environments. Here, we evaluate one such instrument, the Apple Vision Pro, by asking whether it can support psychophysical experiments on two well-studied aspects of peripheral vision: crowding and polar angle asymmetries. These phenomena provide useful benchmarks because both depend on precise control over stimulus position and their behavioral signatures have been well characterized. At the same time, the Apple Vision Pro poses important methodological challenges for vision research, including the inability to access continuous real-time eye-position data and the use of visionOS “points,” rather than degrees of visual angle to specify stimulus size and location. We outline strategies for working within these constraints and argue that the usefulness of the platform is not all-or-none, but depends on the scientific question being asked. Specifically, the Apple Vision Pro may be well suited to paradigms in which fixation can be verified at trial onset, stimulus presentation is brief enough to preclude eye movements, and relative positioning of stimuli is sufficient for the comparison of interest, while remaining less suitable for paradigms requiring continuous gaze monitoring, precise specification of absolute eccentricity, or direct quantitative comparison with conventional lab-based setups.

Introduction

Many classic findings in vision science depend on presenting stimuli at precisely defined locations in the visual field. This is especially true in studies of peripheral vision, where the retinal location of a stimulus is often central to the theoretical question at hand. A large body of research has focused on how vision changes as stimuli are presented at increasing distances from the fovea (i.e., as a function of eccentricity, for a review, see: [Olivers, 2025](#)). One prominent line of work concerns visual crowding. In such studies, target stimuli are presented at a fixed eccentricity in the visual periphery while the distance between the target and a pair of flankers is systematically manipulated (for reviews, see: [Pelli & Tillman, 2008](#); [Whitney & Levi, 2011](#)). To accurately assess how perception of the target changes as a function of the target-flanker distance, one must ensure that the target is presented at the same peripheral location across conditions. Similarly, in studies of polar angle asymmetries (i.e., differences in perception for identical stimuli presented the same distance from fixation, but at varying polar angles), researchers must ensure that the target is presented at the same isoeccentric location in all conditions (for a review, see: [Himmelberg, Winawer, & Carrasco, 2023](#)). Achieving this kind of control has typically required carefully designed laboratory setups with chinrests to control viewing distance, high-resolution displays to control timing and presentation of visual stimuli, and commercial eye trackers to monitor eye movements. Although such setups provide excellent experimental control, they also make many forms of visual psychophysics expensive, tied to a specific

Citation: Volkova, K., Dowd, K., & Grubb, M. A. (2026). Capabilities and limitations of the Apple Vision Pro as a portable psychophysics tool: Replicating crowding and polar angle asymmetries. *Journal of Vision*, 26(7):13, 1–9, <https://doi.org/10.1167/jov.26.7.13>.

<https://doi.org/10.1167/jov.26.7.13>

Received December 11, 2025; published July 22, 2026

ISSN 1534-7362 Copyright 2026 The Authors



location, and difficult to deploy outside specialized laboratory environments.

Can recent advances in portable consumer technology support psychophysical experiments that have traditionally depended on specialized lab infrastructure? Starting with the development of the stereoscope by Sir Charles Wheatstone in the early 19th century (Wheatstone, 1838), a steady progression of technological developments enabled the release of the Oculus Go in 2018, the first mass-produced virtual reality headset that could be used without an external computer or smartphone. There are now many such products available to consumers and vision scientists alike that might provide a reasonable and cost-effective solution to the first two problems (i.e., the ability to control viewing distance without a chinrest and the presence of high-resolution displays). More recently, several standalone headsets have incorporated built-in eye-tracking capabilities, making it possible to ask whether such devices might support portable psychophysics. In the present study, we evaluate one such instrument: the Apple Vision Pro (Apple, Inc., Cupertino, CA, USA).

Here, we report data from two experiments, one on visual crowding, the other on polar angle asymmetries, showing that the Apple Vision Pro reproduced the expected behavioral signatures from prior laboratory studies. The crowding experiment was designed to mimic the approach of two previous *Journal of Vision* articles (Grubb et al., 2013; Yeshurun & Rashal, 2010). Both studies induce and measure crowding by manipulating target-flanker distance with the method of constant stimuli. Both also use a task-irrelevant, spatially-uninformative peripheral onset to elicit the reflexive allocation of covert spatial attention (Carrasco, 2011). We replicate the key findings from those articles: target-flanker distance modulated task performance in the expected way and exogenous attention reduced the crowding critical distance. The polar angle asymmetry experiment was inspired by an extensive set of studies from Marisa Carrasco and colleagues, all of which show that visual perception for identical stimuli, presented equidistant from fixation, is modulated by polar angle (e.g., Abrams, Nizam, Carrasco, 2012). We replicate two classic findings in this literature using a simple upright/inverted judgement of a robot emoji: (1) the horizontal-vertical anisotropy (HVA), defined as an increase in sensitivity for isoeccentric stimuli presented on the horizontal, relative to the vertical, meridian, (2) the vertical meridian asymmetry (VMA), defined as an increase in sensitivity for isoeccentric stimuli presented on lower (south), relative to the upper (north), vertical meridian (Himmelberg et al., 2023).

As the subtitle to this article makes clear, however, using the Apple Vision Pro as a portable research tool comes with challenges and requires accepting some notable limitations. We take two primary limitations

in turn below and present our strategies for addressing them in the context of our experimental work. Highlighting these challenges and limitations here in the Introduction is meant to underscore the fact that the ability to use the Apple Vision Pro as a portable research tool depends heavily on the scientific question that one seeks to answer.

First, because of Apple's data privacy policies, researchers cannot simply query or record real-time eye position data. The workaround we developed was to exploit the Apple Vision Pro's gaze-and-pinch user interface: observers had to fixate a central + and initiate each trial with a pinch. This ensured that participants were fixating the center of the screen at the start of each trial. Importantly, the time between fixation verification and the onset of the briefly-presented target was fast enough in both experiments to preclude eye movements (100 ms in the crowding study, immediately in the polar angle asymmetry experiment). Furthermore, targets were equally likely to appear in either hemifield (crowding experiment) or at any cardinal location (polar angle asymmetry experiment). Thus we can be confident that eye movements cannot account for the ability to reproduce the expected behavioral signatures from prior laboratory studies.

Our strategy works well for the specific research question in the crowding study ("Does covert exogenous attention reduce the crowding critical distance?"), but it would fail if we were interested in covert *endogenous* attention. Why? Because we would have to provide reliable information about the location of the upcoming target and ample time to voluntarily allocate covert spatial attention. Under such conditions, it would be possible to make an eye movement to the target hemifield, which would be problematic for a study of peripheral crowding.

A second methodological challenge of using the Apple Vision Pro for visual psychophysics is an inability to program experiments using a commonly understood size unit like degrees of visual angle (DVA). Apple's visionOS specifies stimulus size and position in an abstract unit called a "point," which is a "unit of measurement that helps visual content remain consistent regardless of how it's displayed. In 2D platforms, a point maps to a number of pixels that can vary according to the resolution of the display; in visionOS, a point is an angular value that allows visual content to scale according to its distance from the viewer" (Apple). In both of our studies, the experimental stimuli were presented on a uniformly-colored background that matched the color of the entire surround (removing any extraneous depth information), and participants completed the studies seated in a chair (reducing variation in viewing geometry within a session). Thus we are confident that stimuli appeared at a consistent *relative* size and *relative* eccentricity within a given session.

For the specific research question in the polar angle asymmetry experiment (“Does polar angle modulate discriminability for stimuli at isoeccentric locations?”), ensuring that stimuli are presented the same distance from fixation is sufficient to make comparisons between locations. But should a researcher be interested in estimating performance at a *specific* eccentricity expressed in DVA or in making direct quantitative comparisons with measurements obtained using conventional displays, this limitation would prove more consequential.

Our take-home assessment for this *JoV Methods* article is that the current usefulness of the Apple Vision Pro for vision science is not all-or-none but depends on the scientific question at hand. The gaze-and-pinch interface can be used to ensure central fixation, after which stimuli can be immediately presented. If target presentation is brief enough to preclude an eye movement, one can be reasonably assured that stimuli will appear at their intended locations in the visual field. Furthermore, if relative positioning of stimuli is sufficient for the theoretical comparison of interest, then the platform can support meaningful inference, as demonstrated by our ability to successfully reproduce the exact outcomes predicted by prior laboratory-based research on crowding and polar angle asymmetries. That said, for paradigms requiring continuous eye-position monitoring, precise specification of absolute stimulus location in DVA, or direct quantitative comparison with conventional display-based setups, the current limitations of the Apple Vision Pro remain substantial.

Methods, visual crowding experiment

Overview

The study closely mimicked Experiment 2 from Yeshurun and Rashal (2010) and was completed using the Apple Vision Pro headset. We developed the study as an application for visionOS, distributed on the same apparatus for each participant. The application was developed for visionOS 2.5 in Swift programming language using Xcode. The experiment consisted of 800 trials in an approximately 60 minute session. We note that the last three experimental trials for one participant are missing due to the battery pack being accidentally disconnected; so for this participant, 797 trials were recorded instead of 800.

Participants

Fourteen participants completed the study; all were students at Trinity College. The study received ethical

approval by the Trinity College Institutional Review Board, and informed consent was obtained for each participant. During recruitment, participants were told “Because you cannot wear glasses inside the VisionPro, you must be able to clearly see a computer screen without corrective lenses. Contact lenses are okay.” Participants were compensated with \$20, and all completed the study. That said, if a participant had to stop the experiment because of discomfort from the headset, they would have been paid a prorated amount based on how much time had elapsed (time in minutes/60 × \$20).

Instruction, calibration, and practice

Participants were shown a series of slides that explained the study and the task. Then, for each participant, the Vision Pro was put into Guest Mode to calibrate their eyes and hands. The Vision Pro then proceeded with the standard, Apple-designed calibration process; this includes a video on using eyes and hand gestures and a calibration by gazing and tapping at colored circles on the screen in different lighting conditions. Before the experiment began, participants completed 32 practice trials, which were identical to the experimental trials (see below).

Experimental details

The experiment consisted of 800 trials, organized as sixteen blocks of 50 trials each. Experimental stimuli were presented on a mid-grey screen (size, 1280 × 651 [in units of Apple Vision Pro “points”]). The Vision Pro was put into the Immersive Mode so that the area surrounding the experimental screen was uniformly covered in grey. A white central fixation cross (diameter of 40 points) was presented throughout the experiment. Participants had to fixate the cross and make a finger tap to initiate each trial (see Figure 1A). A green circular exogenous cue (diameter of 20 points) appeared for 50 ms on the horizontal meridian, 200 points from the center cross; the location of the cue was equally likely to be on the left or the right and provided no information about the location of the upcoming stimulus array. After an inter-stimulus interval of 50 ms, the stimulus array was rendered for 100 ms on the horizontal meridian, 340 points to the right or to the left of fixation. The stimulus array consisted of the three elements: a white target T in one of four configurations (0°, 90°, 180°, 270° rotation) and two flankers (each the letter H, which were randomly and independently presented with 0° or 90° rotation) at a fixed target-to-flanker distance (30.6, 64.6, 98.6, 132.6, 166.6, 200.6, 234.6, and 268.6 points). T and H had a size of 30.6 points. After a 400 ms inter-stimulus-interval, participants reported

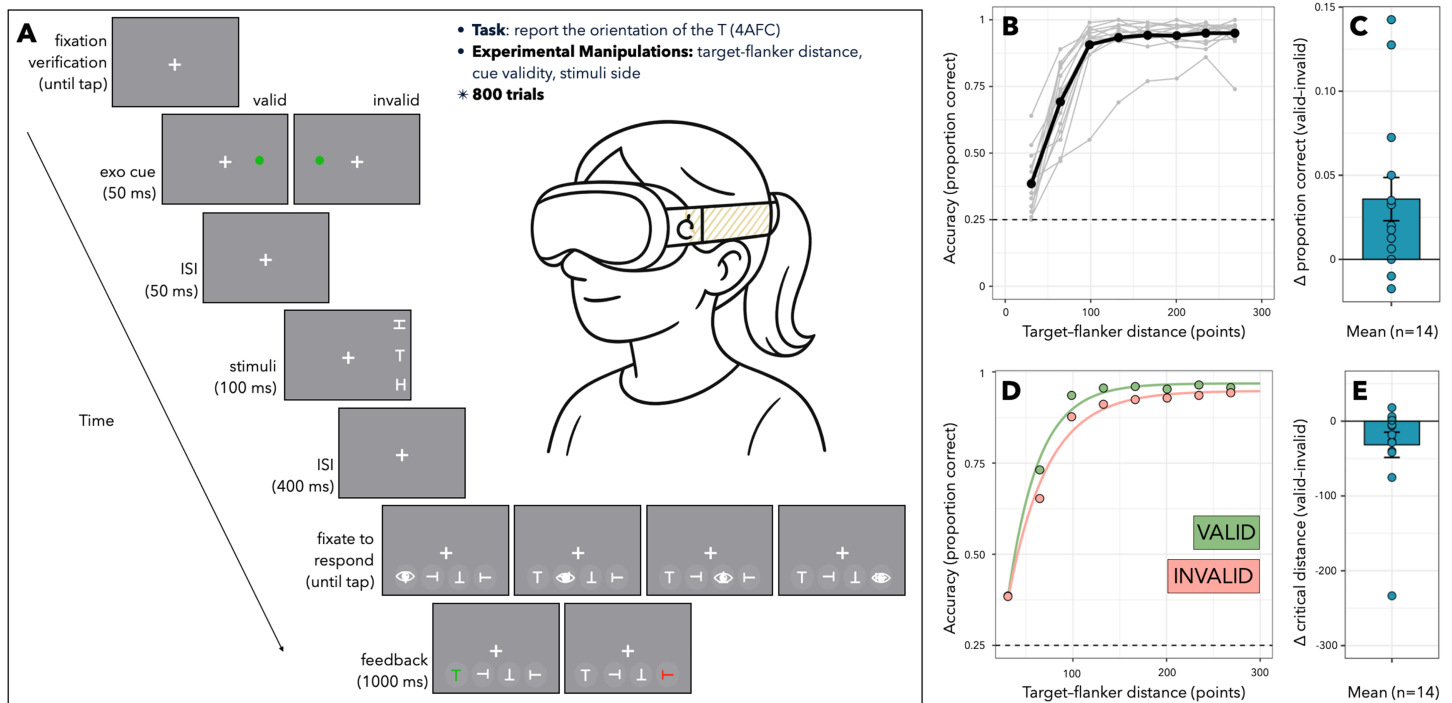


Figure 1. Visual crowding experiment. **(A)** Trial sequence. **(B)** Accuracy as a function of target–flanker distance: black line, mean across participants; gray lines, individual participants; dashed line, chance performance. **(C)** Attentional cueing effect: bar height, mean change in proportion correct (valid minus invalid); points, individual participants; error bar, standard error of the mean across participants. **(D)** Accuracy as a function of target–flanker distance: green points, mean across participants valid trials; pink points, mean across participants invalid trials; solid lines, exponential model fit to group data for illustrative purposes. **(E)** Attentional cueing effect: bar height, mean change in critical distance (valid minus invalid); same format as panel C.

the orientation of the target T by fixating one of the four response options and making a finger tap. They then received 1000 ms of feedback: their chosen option would turn green if correct, red if incorrect.

Data analysis

We calculated proportion correct for each participant in each experimental condition (two cue types and eight target–flanker distances). Cues were *valid* when they appeared in the same visual hemifield as the subsequent target and *invalid* when they appeared in the opposite visual hemifield. We then used JASP (JASP Team, 2024) to conduct a 2×8 repeated-measures ANOVA on proportion correct data.

Model fitting

We plotted proportion correct as a function of target–flanker distance for both cue types and fitted an exponential model to the data (as in: Grubb et al., 2013; Yeshurun & Rashal, 2010):

Model 1:

$$pc = a(1 - e^{-s(d-i)}), \quad d > i$$

where pc is proportion correct, a is the asymptote, s is the scaling factor, d is the target-to-flanker distance, and i is the x-intercept. We used nonlinear least squares (nls function in R), which determined the most suitable parameters for each participant in each cue condition. The estimated parameter values were then used to calculate the critical distances:

$$cd = i - \frac{\ln(0.10)}{s}$$

where cd is the critical distance, i is the x-intercept, and s is the scaling factor.

For each model fit, we calculated a pseudo- R^2 value as a measure of goodness of fit using the formula $1 - (RSS/TSS)$, where RSS is the residual sum of squares and TSS is the total sum of squares. The average value was 0.9514, with a standard deviation of 0.0499 and a range of [0.7990, 0.9975].

Apparatus

Both experiments used the Apple Vision Pro from Apple with a screen refresh rate of 90 Hz. Apple reports that the device contains 23 million pixels across the two displays, but does not provide an official $X \times Y$ pixel resolution in its technical specifications. The headset weighs 26.4–28.2 ounces (750–800 grams). Full technical specifications can be found here: <https://www.apple.com/apple-vision-pro/specs/>.

Stimulus sizes and locations are reported in visionOS points throughout the article. Although an exact conversion from points to degrees of visual angle would be useful for many vision science applications, Apple does not provide the information needed to make this conversion precisely. Thus the point values reported here should be interpreted as the native coordinate units used by the application. This limitation does not affect the within-device comparisons reported here, which depend on relative stimulus positioning rather than absolute specification of eccentricity in degrees of visual angle.

Methods, polar angle asymmetries experiment

Overview

The experiment for Apple Vision Pro was built on visionOS 2.5 in Swift language using Xcode. The experiment consisted of 608 trials in an approximately 60-minute session.

Participants

Ten participants completed the study; all were students at Trinity College. The study received ethical approval by the Trinity College Institutional Review Board, and informed consent was obtained for each participant. During recruitment, participants were told “Because you cannot wear glasses inside the VisionPro, you must be able to clearly see a computer screen without corrective lenses. Contact lenses are okay.” Participants were compensated with \$20 and all completed the study. That said, one of the participants did a version with only 400 trials.

Instruction and calibration

Participants were shown a series of slides that explained the study and the task. Then, for each participant, the Vision Pro was put into Guest Mode

to calibrate their eyes and hands. The Vision Pro then proceeded with the standard, Apple-designed calibration process; this includes a video on using eyes and hand gestures and a calibration by gazing and tapping at colored circles on the screen in different lighting conditions.

Experimental details

The experiment consisted of 608 trials, organized in one block. Experimental stimuli were presented on a black screen (size, 1200×1200 points). The Vision Pro was put into the Immersive Mode so that the area surrounding the experimental screen was uniformly covered in black. A white central fixation circle (diameter of 17 points) was overlaid with a white cross (diameter of 34 points). Participants had to fixate the circle with the cross and make a finger tap to initiate each trial (see Figure 2A). Four 103×103 points TV frame images appeared 340 points from the center at 0° , 90° , 180° , and 270° for east, north, west, and south, respectively. Their background was filled with mid-gray, and they remained in the same locations through the time of the experiment. A 57.8×57.8 -point robot emoji appeared for 100 ms, in either an upright or inverted configuration, in one of the four locations. Eight combinations of robot location (0° , 90° , 180° , 270°) and orientation (upright, inverted) were presented in a random order. Then, a white line (length of 17 points) appeared from the center of the screen prompting participants to report the orientation of the robot at the indicated location by fixating one of the two response options and making a finger tap. The point of the response cue was to remove any spatial uncertainty about where the robot emoji had been presented. Participants then received 500 ms of feedback: the central fixation would turn green if correct, red if incorrect. Then, a 1000 ms trial countdown indicated how many trials remained.

Maintaining task difficulty

The contrast of the robot emoji was set to 0.5 at the beginning of the experiment, and it adapted to the participant's performance. After two complete iterations of the eight unique trials types, proportion correct for these 16 trials was computed, and the new contrast level was calculated with the following formula:

$$c_{i+1} = c_i \left(2 - \frac{a_i}{0.75} \right)$$

where c_i is the current contrast, a_i is proportion correct, and c_{i+1} is the new contrast used for the next 16 trials. Thus, when proportion correct was 0.75, contrast did

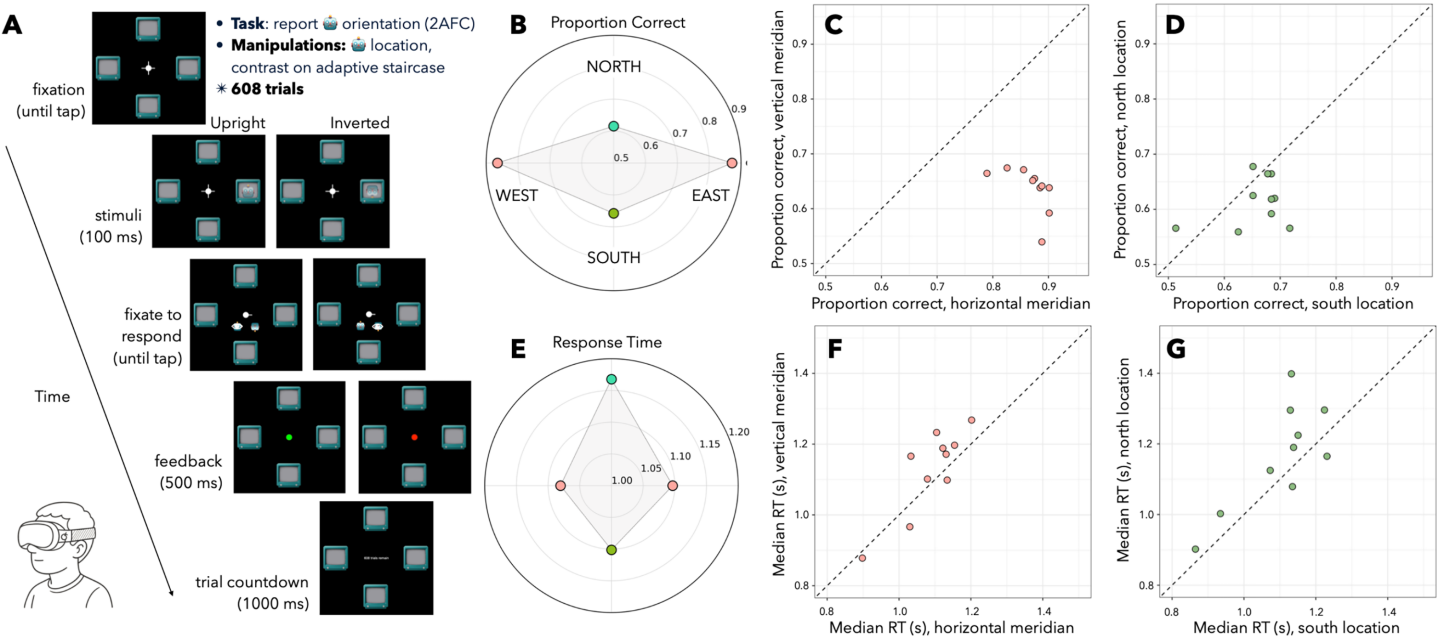


Figure 2. Polar angle asymmetries experiment. (A) Trial sequence. (B) Accuracy as a function of target location in proportion correct. (C) Horizontal-Vertical Anisotropy, proportion correct: points, individual participants. (D) Vertical Meridian Asymmetry, proportion correct: points, individual participants. (E) Median response time (correct trials only) as a function of target location in seconds. (F) Horizontal-Vertical Anisotropy, RT: points, individual participants. (G) Vertical Meridian Asymmetry, RT: points, individual participants.

not change; when accuracy exceeded 0.75, the current contrast was multiplied by a number less than 1, thereby reducing the contrast for the next set of 16 trials and making the task more difficult; when accuracy was less than 0.75, the current contrast was multiplied by a number greater than 1, thereby increasing the contrast for the next set of 16 trials and making the task easier. Contrast was constrained to remain between 0.05 and 1. In short, this procedure allowed task difficulty to be adjusted for each participant while ensuring that every contrast level was presented equally often at each location.

Data analysis

We calculated proportion correct and median response times for correct trials for each participant in each experimental condition separately (two meridian conditions [horizontal and vertical] and four location conditions [south, north, east, west]).

Bayesian model fitting

On each trial, a given participant either reported “upright” or “inverted” configuration. We modeled the trial-level decision ($\text{saidUpright} = 1$ or saidUpright

$= 0$) as the outcome of a single draw from a binomial distribution with probability p (Model 2, line 1; see Table 1). The probability of a *saidUpright* response was allowed to vary according to standard signal detection theory, with two parameters: d' representing perceptual sensitivity to upright signals, and b representing a decisional bias to report one option over the other (Model 2, line 2). In this formulation, *targetStatus* was coded as 1 when an upright signal was present and -1 when it was absent (i.e., when the emoji was inverted). Line 3 decomposes sensitivity into a participant-specific baseline term (δ_{0i}) and two fixed-effects designed to capture the HVA (cH) and

1	$\text{saidUpright}_t \sim \text{Bernoulli}(p_t)$
2	$p_t = \Phi(\text{targetStatus}_t \times d_t/2 - b_t)$
3	$d_t = \delta_{0i} + \delta_H \times cH_t + \delta_V \times cV_t$
4	$b_t = \beta_{0i} + \beta_H \times cH_t + \beta_V \times cV_t$
Prior distributions	
5	$\delta_{0i} \sim \mathcal{N}(1.35,1)$
6	$\beta_{0i} \sim \mathcal{N}(0,1)$
7	$\delta_H \sim \mathcal{N}(0,1)$
8	$\delta_V \sim \mathcal{N}(0,1)$
9	$\beta_H \sim \mathcal{N}(0,1)$
10	$\beta_V \sim \mathcal{N}(0,1)$

Table 1. Model 2.

VMA (cV). For the HVA contrast, cH was coded as $+0.5$ when the emoji appeared at horizontal locations (east, west) and -0.5 when it appeared at vertical locations (north, south). For the VMA contrast, cV was coded as $+0.5$ when the emoji appeared at the south location, -0.5 when it appeared at the north location, and zero when it appeared along the horizontal meridian. Line 4 allows the decisional bias term to vary in the same way.

We used the *ulam* function in the *rethinking* package in R to estimate posterior distributions. We ran four chains, each with 1000 iterations. The largest R-hat convergence diagnostic value for any parameter was 1.000445, which indicates successful mixing of the chains.

Use of generative AI

A large language model (ChatGPT, OpenAI) was used as a supplementary coding tool during this project to assist with debugging code in the Swift programming language during development of the visionOS experimental applications and during the development of model-based analyses in R. In both cases, the large language model was used in a manner analogous to traditional use of online documentation and search engines for troubleshooting: it suggested code snippets or modeling strategies, which the authors then implemented, tested, and modified as needed. The same tool was also used iteratively during the writing process to increase clarity and to provide feedback on structure and flow. All scientific citations were selected by the authors, who take full responsibility for the content of this work.

Results, visual crowding experiment

Accuracy for reporting the orientation of the target T was modulated by both the distance of the flankers and the reflexive allocation of spatial attention. We assessed proportion correct data using a 2×8 ANOVA, with target-to-flanker distance (in units of Apple Vision Pro “points”: 30.6, 64.6, 98.6, 132.6, 166.6, 200.6, 234.6, and 268.6) and exogenous cue condition (valid, invalid) as repeated-measures factors. We observed a main effect of target-to-flanker distance ($F(7, 91) = 140.61$, $p < 0.001$), with accuracy increasing to asymptotic performance as target-to-flanker distance increased (Figure 1B). We also observed a main effect of cue condition ($F(1, 13) = 7.70$, $p = 0.016$), with greater accuracy on valid relative to invalid trials (Figure 1C). The target-to-flanker distance $\times$ cue condition interaction did not reach the conventional significance

threshold ($F(7, 91) = 1.85$, $p = 0.086$). However, the pattern of means suggested that attention mainly improved performance at intermediate target-to-flanker distances (Figure 1D, points). To quantify this pattern, we next fit an exponential model to the data (Figure 1D, solid lines) and used the resulting parameters to derive a model-based estimate of the crowding critical distance in each attention condition.

When spatial attention was reflexively allocated to the same hemifield as the target T (i.e., valid trials), the critical distance was reduced, relative to when the cue appeared in the opposite hemifield (i.e., invalid trials). Consistent with previous research (e.g., Grubb et al., 2013; Yeshurun & Rashal, 2010), the critical distance was defined as the target-to-flanker distance at which orientation discrimination accuracy reached 90% of the asymptotic value. We estimated the critical distance, separately for each participant and for each attention condition, and then computed the within-participant change (Figure 1E). A t -test revealed that estimates of the critical distance were significantly smaller in the valid, relative to invalid conditions ($t(13) = -1.87$, $p = 0.0418$, one-tailed in predicted direction). Given that one of the estimates was more than three standard deviations from the group mean, we verified that this outlier didn’t exert undue influence by using a Wilcoxon signed-rank test ($p = 0.0148$, one-tailed in predicted direction). In short, we successfully induced and measured visual crowding, while also replicating previous research showing that covert exogenous spatial attention reduces the critical distance.

Results, polar angle asymmetries experiment

Despite being equidistant from fixation, polar angle modulated both accuracy (Figure 2B) and response times (Figure 2E). Participants were better and faster at reporting the orientation of the robot emoji when it appeared on the horizontal meridian (east and west locations) relative to the vertical meridian (south and north locations). We calculated proportion correct (Figure 2C) and median response time for correct trials only (Figure 2F) for each meridian. Paired t -tests revealed that accuracy was better ($t(9) = 10.78$, $p < 0.001$, one-tailed in predicted direction) and response times were faster ($t(9) = -1.85$, $p = 0.0488$, one-tailed in predicted direction) when the target appeared on the horizontal meridian, ruling out the possibility of a speed-accuracy trade-off. Within the vertical meridian, target presentation in the southern location resulted in better performance than the northern location. We calculated proportion correct (Figure 2D) and median response time for correct trials only (Figure 2G) for each location along the vertical meridian. Paired

t -tests revealed that accuracy was better ($t(9) = 2.27$, $p = 0.0247$, one-tailed in predicted direction) and response times were faster ($t(9) = -2.18$, $p = 0.0286$, one-tailed in predicted direction) when the target appeared at the southern location, again ruling out the possibility of a speed-accuracy tradeoff. In short, we replicated both of the classic polar angle asymmetries.

A model-based, signal detection theory analysis confirmed robust polar angle asymmetries that could not be explained by response bias. To complement the proportion correct analyses with a trial-level approach, we next modeled the binary responses (“upright” vs. “inverted”) using a Bayesian mixed-effects signal detection theory model, with participant-level random intercepts on sensitivity and response bias and fixed effects of target location (see Model 2 in Methods). In this model, each observer’s sensitivity d' was decomposed into an individual baseline plus additive components for the horizontal-vertical anisotropy (horizontal vs. vertical meridian, δ_H) and the vertical meridian asymmetry (south vs. north along the vertical meridian, δ_V); response bias was allowed to vary with these same contrasts (β_H , β_V). The mean of the posterior distribution for the HVA sensitivity parameter was strongly positive (δ_H mean = 1.55, 95% credible interval: [1.40, 1.70]), indicating higher d' when the target appeared on the horizontal relative to the vertical meridian. The mean of the posterior for the VMA sensitivity parameter was also positive (δ_V mean = 0.22, 95% credible interval: [0.03, 0.41]), indicating higher d' for southern than for northern locations on the vertical meridian. The mean of the posterior distribution for the HVA bias parameter was centered near zero (β_H mean = -0.01 , 95% credible interval: [-0.09 , 0.06]), whereas the mean of the posterior distribution for the VMA bias parameter was reliably nonzero (β_V mean = 0.29, 95% credible interval: [0.19, 0.38]). Thus, when the data were analyzed at the trial-level using d' instead of proportion correct, and even when response bias was allowed to vary with location, we recover the same pattern of location-specific perceptual advantages, demonstrating that the polar angle asymmetries are robust to a different analysis and alternative method of inference.

Discussion

Recent technological advances in high-resolution, mixed-reality displays with built-in eye tracking have raised the possibility that some forms of visual psychophysics could be conducted on portable consumer devices rather than only in specialized laboratory settings. In this project, we evaluated one such platform, the Apple Vision Pro, by asking whether we could reproduce expected results from two well-

studied aspects of peripheral vision: crowding and polar angle asymmetries. We chose these phenomena because their behavioral signatures are well characterized (Grubb et al., 2013; Himmelberg et al., 2023; Yeshurun & Rashal, 2010), and we knew exactly what successful replication would look like in both cases. Indeed, for both phenomena, we replicate classic effects in two small samples of Trinity College students. In the first experiment ($n = 14$), we induced visual crowding, used a model-based approach to quantify the critical distance, and showed that covert exogenous spatial attention modulated this critical distance. In the second experiment ($n = 10$), we measured both canonical polar angle asymmetries (the horizontal-vertical anisotropy and the vertical meridian asymmetry) and collected data of sufficient quantity and quality to fit a Bayesian signal detection model to trial-level responses demonstrating that these asymmetries arise from differences in perceptual sensitivity rather than response bias. Taken together, these results indicate that the Apple Vision Pro can support meaningful inference under certain circumstances.

In this moment of continued replication anxiety (e.g., Tyner et al., 2026), it is worth stressing that many canonical findings in vision science are remarkably robust. That robustness makes such effects especially valuable as benchmarks for evaluating new methods because they provide clear empirical predictions against which a new platform can be judged. In that sense, vision science is especially well positioned to assess emerging portable technologies such as the Apple Vision Pro in a principled way. If such platforms can reliably reproduce well-established behavioral signatures, they may allow the field to bring the laboratory to a wider range of settings (e.g., schools, assisted living facilities, and public venues) and thereby broaden the populations and contexts in which psychophysical data can be more readily collected.

Methodologically, these findings should be interpreted in light of two important constraints discussed in the *Introduction*: the inability to access continuous real-time eye position data and the inability to specify stimulus size and location in conventional units such as degrees of visual angle. In the present experiments, we were able to work within these constraints by verifying central fixation at trial onset, presenting targets quickly enough to preclude eye movements, and relying on comparisons for which relative positioning was sufficient. For paradigms that require continuous gaze monitoring, exact specification of absolute eccentricity, or direct quantitative comparison with conventional display-based setups, these limitations remain substantial.

Developing experiments for the Apple Vision Pro required a meaningful amount of programming expertise. In the present case, however, both visionOS applications were developed by the first and second

authors as part of their undergraduate honors projects at Trinity College. Each student was completing a double major that paired Computer Science with Neuroscience or Psychology, and both were able to bring substantial prior programming experience to the project despite having little background in Apple development. Each experiment ultimately required approximately 170–200 hours of design, development, debugging, and testing, placing the workload at roughly the scale of a semester-long undergraduate research credit or the expected engaged-learning hours of a typical college course. We think this is worth noting because it suggests that, although VR-based experimentation is not plug-and-play, it is nonetheless accessible to small research groups, including undergraduate-centered labs, especially when those groups can draw on collaborators with computer science training.

Even with the limitations outlined here, portable mixed-reality devices such as the Apple Vision Pro may still be valuable for certain kinds of behavioral research. Their chief advantage is not that they can replace conventional laboratory setups in all cases, but that they make it possible to bring experiments to participants in settings where traditional psychophysical infrastructure would be difficult to deploy. More broadly, self-contained headsets may support the development of paradigms that can be administered in more naturalistic environments. At the same time, the Apple Vision Pro is only one example of a broader class of portable mixed-reality devices (e.g., HTC Vive Pro Eye and the Meta Quest Pro), and its relatively high cost and inability to accommodate conventional eyeglasses mean that the device will be most useful when its specific capabilities justify these practical constraints. At present, however, such promise is best understood as conditional: devices such as the Apple Vision Pro are most useful when the research question benefits from a portable, self-contained platform and can be addressed without continuous gaze monitoring or precise specification of absolute stimulus position in degrees of visual angle.

Keywords: crowding, polar angle asymmetry, exogenous attention, virtual reality

Acknowledgments

Supported by NSF CAREER #2141860 to MAG.

Commercial relationships: none.

Corresponding author: Michael A. Grubb.

Email: michael.grubb@trincoll.edu.

Address: Department of Psychology, Life Sciences Center, Trinity College, Hartford, CT 06106, USA.

References

- Abrams, J., Nizam, A., & Carrasco, M. (2012). Isoeccentric locations are not equivalent: The extent of the vertical meridian asymmetry. *Vision Research*, 52(1), 70–78.
- Apple. Images, <https://developer.apple.com/design/human-interface-guidelines/images>.
- Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Research*, 51(13), 1484–1525, doi:10.1016/j.visres.2011.04.012.
- Grubb, M. A., Behrmann, M., Egan, R., Minshew, N. J., Heeger, D. J., & Carrasco, M. (2013). Exogenous spatial attention: evidence for intact functioning in adults with autism spectrum disorder. *Journal of Vision*, 13(14), 9, doi:10.1167/13.14.9.
- Himmelberg, M. M., Winawer, J., & Carrasco, M. (2023). Polar angle asymmetries in visual perception and neural architecture. *Trends in Neurosciences*, 46(6), 445–458, doi:10.1016/j.tins.2023.03.006.
- Olivers, C. N. L. (2025). Selective attention and eccentricity: A comprehensive review. *Neuroscience & Biobehavioral Reviews*, 178, 106368, doi:10.1016/j.neubiorev.2025.106368.
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129–1135, doi:10.1038/nn.2187.
- Tyner, A. H., Abatayo, A. L., Daley, M., Field, S., Fox, N., Haber, N. A., . . . Errington, T. M. (2026). Investigating the replicability of the social and behavioural sciences. *Nature*, 652(8108), 143–150, doi:10.1038/s41586-025-10078-y.
- Wheatstone, C. (1838). *Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision*. London: Royal Society Publishing.
- Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15(4), 160–168, doi:10.1016/j.tics.2011.02.005.
- Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision*, 10(10), 16, doi:10.1167/10.10.16.