Perspective Cues Make Eye-specific Contributions to 3-D Motion Perception

Lowell W. Thompson[1], Byounghoon Kim[1], Zikang Zhu[1], Bas Rokers[2], and Ari Rosenberg[1]

Abstract

■ Robust 3-D visual perception is achieved by integrating stereoscopic and perspective cues. The canonical model describing the integration of these cues assumes that perspective signals sensed by the left and right eyes are indiscriminately pooled into a single representation that contributes to perception. Here, we show that this model fails to account for 3-D motion perception. We measured the sensitivity of male macaque monkeys to 3-D motion signaled by left-eye perspective cues, right-eye perspective cues, stereoscopic cues, and all three cues combined. The monkeys exhibited idiosyncratic differences in their biases and sensitivities for each cue, including left- and right-eye perspective cues, suggesting that the signals undergo at least partially separate neural processing. Importantly, sensitivity to combined cue stimuli was greater than predicted by the canonical model, which previous studies found to account for the perception of 3-D orientation in both humans and monkeys. Instead, 3-D motion sensitivity was best explained by a model in which stereoscopic cues were integrated with left- and right-eye perspective cues whose representations were at least partially independent. These results indicate that the integration of perspective and stereoscopic cues is a shared computational strategy across 3-D processing domains. However, they also reveal a fundamental difference in how left- and right-eye perspective signals are represented for 3-D orientation versus motion perception. This difference results in more effective use of available sensory information in the processing of 3-D motion than orientation and may reflect the temporal urgency of avoiding and intercepting moving objects. ■

INTRODUCTION

To resolve sensory ambiguities and improve sensitivity, the brain integrates different cues within and across modalities (Dakin & Rosenberg, 2018; Rohde, van Dam, & Ernst, 2016; Seilheimer, Rosenberg, & Angelaki, 2014; Landy, Banks, & Knill, 2011). For example, transforming 2-D retinal images into 3-D percepts that can guide behaviors such as catching a frisbee involves the integration of perspective and stereoscopic cues (Chang, Thompson, et al., 2020; Thompson, Ji, Rokers, & Rosenberg, 2019; Rokers, Fulvio, Pillow, & Cooper, 2018; Hillis, Watt, Landy, & Banks, 2004; Knill & Saunders, 2003). Perspective cues originate from the projection of the 3-D world onto the 2-D retinae (Knill, 1998; Stevens, 1981). Stereoscopic cues are based on differences between the left and right retinal images (Howard & Rogers, 1995). Ideal observer models provide a framework for describing how such cues are integrated to yield coherent percepts (Landy et al., 2011).

These models widely assume that separate cues support the creation of independent sensory estimates by distinct neuronal populations with predominantly different noise sources (Seilheimer et al., 2014; Ma, Beck, Latham, & Pouget, 2006). This assumption likely holds more across sensory modalities than within modalities, particularly if the same neurons extract stimulus features from separate cues. Such nonindependence in sensory processing is information limiting and therefore restricts the perceptual gains of cue integration (Chang, Thompson, et al., 2020; Fulvio, Ji, Thompson, Rosenberg, & Rokers, 2020; Oruç, Maloney, & Landy, 2003). Previous studies suggest that 3-D object orientation estimates based on perspective and stereoscopic cues are effectively independent. That work further suggests that perspective cues from the left and right eyes are indiscriminately pooled, despite them being distinct because of the horizontal offset between the eyes (Chang, Thompson, et al., 2020; Hillis et al., 2004; Knill & Saunders, 2003). If the two eyes' perspective cues instead made (partially or fully) independent contributions to orientation perception, then perceptual sensitivity to stimuli containing both cues would be greater than observed. Thus, to account for why 3-D orientation sensitivity is not maximized, the canonical model of 3-D cue integration includes a lack of independence between the left- and right-eye perspective cue representations.

Here, we sought to test if such information-limiting processes are a general property of cue integration when the same cue type is detected by multiple corresponding sensors (e.g., the two eyes, ears, or hands) by extending that work to the domain of 3-D motion processing. Humans can successfully judge the direction of 3-D motion based on perspective and stereoscopic cues, but efforts to assess cue integration have had limited

[1]University of Wisconsin-Madison, [2]New York University Abu Dhabi

© 2021 Massachusetts Institute of Technology

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

success (Fulvio et al., 2020; Thompson et al., 2019). To overcome this challenge, we collected large within-subject data sets from two macaque monkeys that were well trained to discriminate toward versus away motion. Across conditions, the direction of motion was either signaled by a single cue type (left-eye perspective, right-eye perspective, or stereoscopic cues) or all three cues. The motion coherence was varied to estimate biases and sensitivities.

In contrast to previous 3-D orientation perception results, we found that sensitivity to 3-D motion was greater than predicted by the canonical cue integration model. Instead, 3-D motion sensitivities were best explained if left- and right-eye perspective cues made at least partially independent contributions to perception. Thus, although both sensory domains integrate perspective and stereoscopic cues, they differ in how left- and right-eye perspective signals are represented. For 3-D motion, the ocular identity of perspective cues from the two eyes remains at least partially preserved (indicating greater independence than in the orientation domain). This allows perspective cues from both eyes to be exploited to achieve greater motion sensitivity through cue integration than has been observed for 3-D orientation processing. This more effective use of available sensory information for 3-D motion processing may be related to the relevance and temporal urgency of avoiding and intercepting moving objects. These differences in 3-D orientation and motion processing reveal that neurocomputational processes, which impose information-limiting constraints on cue integration, are domain specific and imply novel factors that can shape perception in binocular vision, audition, and bimanual touch.

METHODS

Animal Preparation

All surgeries and experimental procedures were approved by the Institutional Animal Care and Use Committee at the University of Wisconsin-Madison and were in accordance with National Institutes of Health guidelines. Given previous challenges in assessing 3-D motion cue integration and standard practices in the field, we prioritized obtaining a large number of samples from two male rhesus monkeys (*Macaca mulatta*; Monkey J: 6.2 years of age, 7.6 kg in weight; Monkey C: 5.8 years of age, 7.5 kg). Neither monkey previously participated in any other study. A Delrin ring for stabilizing the head during experiments was attached to the skull under general anesthesia (Chang, Doudlah, et al., 2020; Chang, Thompson, et al., 2020; Rosenberg, Cowan, & Angelaki, 2013). After recovery, the monkeys were trained to sit in a primate chair and to fixate visual targets within 2° version and 1° vergence windows for liquid reward. Eye positions were monitored optically at 1000 Hz (EyeLink 1000 plus, SR Research).

Behavioral Control and Stimulus Presentation

Experimental control was performed using the open-source REC-GUI system (RRID:SCR_019008; Kim et al., 2019). Stimuli were presented on a 24-in. Acer GN246HL LED monitor (1920 × 1080 pixels, 120 Hz) at a viewing distance of 57 cm. Stereoscopic presentation was achieved by temporally interlacing left- and right-eye images using an NVIDIA 3-D Vision 2 wireless glasses kit. We modified the glasses for a macaque interocular distance. The crosstalk (averaged across eyes) when the maximum or minimum stimulus luminance was presented to the "closed" eye and the background was presented to the "open" eye was low: 1.85% and 0.88%, respectively (Woods, 2012). The stimuli were created in MATLAB R2015a using Psychtoolbox 3 (Kleiner et al., 2007) and rendered with anti-aliasing using an NVIDIA Quadro K4000 graphics card on a Windows 7 workstation.

Visual Stimuli

The stimuli depicted motion toward or away from the midpoint between the monkey's eyes, commonly referred to as the "cyclopean eye." Twenty-two dots, 11 light (23.70 cd/m^2) and 11 dark (0.09 cd/m^2), were initialized with pseudorandom positions within a 3-D volume. The volume was oriented toward the cyclopean eye, spanned ±1° of horizontal disparity (42.5 cm depth range), and corresponded to a 3° diameter circular aperture on the screen (Figure 1A). The background was gray (8.45 cd/m^2). Stimuli were presented at 25 locations in a 5 × 5 grid centered on the cyclopean eye. The grid spacing was 3.42° horizontal and 3.19° vertical. All tested visual field locations were less eccentric than the blind spot, which is found on the horizontal meridian at an eccentricity of ~15°–17° in macaques (Komatsu, Kinoshita, & Murakami, 2002). Four stimulus conditions were included.

In the combined cue condition, the direction of motion was signaled by perspective and stereoscopic cues (Figure 1B, top; Supplemental Movie: https://osf.io/8wxk7/). The perspective cues included optic flow and changes in retinal density and size. Because of the horizontal offset of the eyes, the optic flow patterns for each eye had foci of expansion/contraction that were offset in opposite directions on the screen (Thompson et al., 2019; Cormack, Czuba, Knöll, & Huk, 2017). Thus, the perspective cues were eye specific, and the left- and right-eye images had opposite net horizontal motion directions. At the nearest depth, a dot subtended 0.18°. At the depth of fixation (57 cm), a dot subtended 0.1°. At the farthest depth, a dot subtended 0.04°. Across depth, the average dot size was 0.107°. The stereoscopic cues included changing disparity and interocular velocity differences, both of which contribute to 3-D motion perception (Allen, Haun, Hanley, Green, & Rokers, 2015; Rokers, Czuba, Cormack, & Huk, 2011; Nefs, O'Hare, & Harris, 2010; Brooks, 2002; Shioiri, Saisho, & Yaguchi,

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

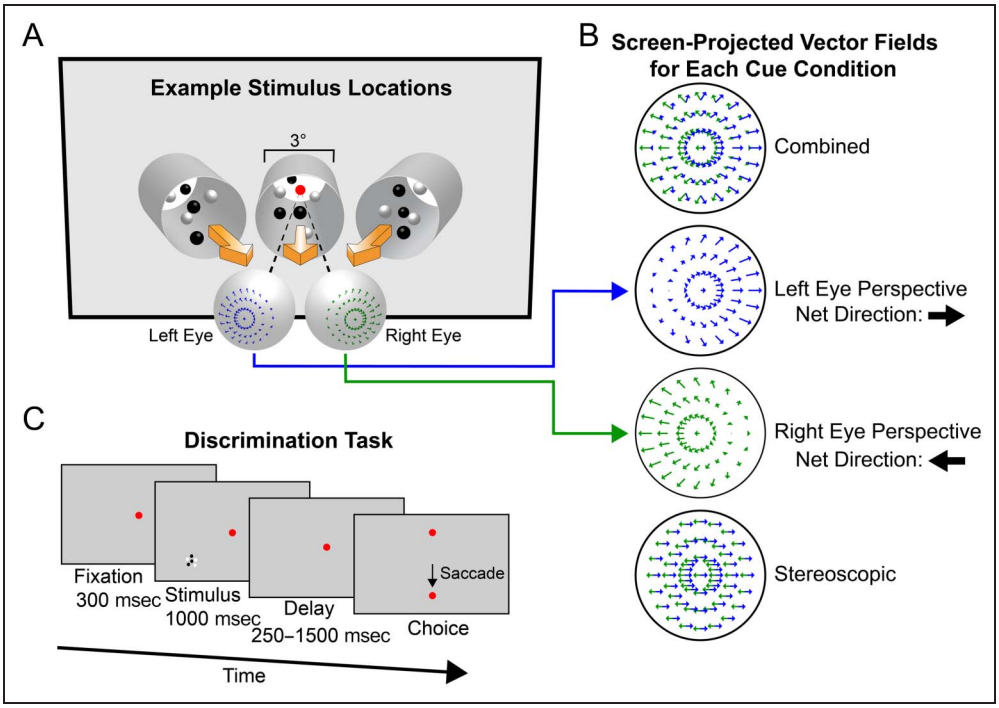


Figure 1. Stimuli and experimental design. (A) The stimuli depicted dots that moved toward or away from the cyclopean eye. Toward trajectories are illustrated for three different visual field locations. The dots were confined to a volume spanning ±1° of horizontal disparity (a depth range of 42.5 cm at the 57 cm viewing distance), which corresponded to a 3° diameter circular aperture on the screen. The red dot at the center of the middle location represents the fixation target. For clarity, the stimuli shown here are reductive. Actual examples are shown in a Supplemental Movie (https://osf.io/8wxk7/). (B) The four cue conditions illustrated as screen-projected vector fields for toward motion. Combined cue stimuli contained left-eye (blue) and right-eye (green) perspective cues as well as stereoscopic cues. Perspective cue stimuli were monocular presentations of the combined cue stimuli. Left- and right-eye perspective cues had equal and opposite net 2-D horizontal motions. Stereoscopic cue stimuli consisted of dot pairs that horizontally translated in opposite directions in the two eyes. The screen projections are reflected compared with the retinal projections in A because of the optics of the eyes. (C) Discrimination task. The direction of motion (toward or away) was indicated by a saccade (black arrow) to one of two choice targets (lower or upper, respectively).

2000). At all visual field locations, the average retinal speed of the dots was 4.38°/sec ± 1.0°/sec standard deviation (*SD*).

In the two perspective cue conditions (one for each eye), the appropriate half-images of combined cue stimuli were presented to a single eye (Figure 1B, middle two panels; Supplemental Movie: https://osf.io/8wxk7/). Both eyes saw the fixation target. In the stereoscopic cue condition, the stimuli were rendered using orthographic projection and had a fixed dot size (0.1°) to eliminate any perspective cues that could signal 3-D motion (Fulvio et al., 2020; Thompson et al., 2019; Figure 1B, bottom; Supplemental Movie: https://osf.io/8wxk7/). Following previous work (Chang, Thompson, et al., 2020; Fulvio et al., 2020; Thompson et al., 2019), the dot size in the stereoscopic cue condition was set equal to the dot size in the other conditions at the depth of fixation and was minimally different from the average dot size. The change in disparity over time was equated in the combined cue and stereoscopic cue conditions. As such, the left- and right-eye dot pairs translated horizontally at equal and opposite speeds (4.2°/sec at all visual field locations). Because the change in disparity over time was equated across conditions, there was a small difference in retinal speed (0.18°/sec) between the stereoscopic cue stimuli and the average speed of the other stimuli. However, at these speeds, the difference would have little-to-no impact on 3-D motion sensitivity (Cooper, van Ginkel, & Rokers, 2016). Moreover, the difference is substantially smaller than the speed tuning bandwidths of middle temporal (MT) area neurons (Priebe, Lisberger, & Movshon, 2006; Nover, Anderson, & DeAngelis, 2005), which likely contribute to the computation of 3-D motion. Importantly, the rank order of observed cue sensitivities was monkey specific (e.g., Monkey J's averaged single-cue sensitivities ranked greatest to least were stereoscopic cues, left-eye perspective cues, and right-eye perspective cues; see Figure 3), further indicating that this small difference in stimulus speeds could not account for cue-related differences in 3-D motion sensitivity. Because 3-D trajectories in which the direction of motion is toward or away from the cyclopean eye produce 2-D retinal motion patterns that are balanced (i.e., equal and opposite) in the two eyes, the net horizontal motions of the combined cue stimuli were consistent with the stereoscopic cue stimuli.

Seven motion coherences were used. Signal dots moved toward or away from the monkey, whereas noise dots were reassigned to pseudorandom positions within the volume. The proportion of signal to noise dots was: 0, 0.09, 0.18, 0.36, 0.45, 0.64, and 1 (0/22, 2/22, 4/22, 8/22, 10/22, 14/22, and 22/22 signal/total dots, respectively). To discourage the tracking of individual dots, on each stereoscopic frame pair (i.e., every 0.017 sec) each dot was pseudorandomly selected to be either a signal or noise dot (Fulvio et al., 2020; Thompson et al., 2019).

To encourage the monkeys to judge the 3-D direction of motion, we used a cloud of dots rather than a plane. With a plane, the monkeys could potentially judge where the stimulus ended (or was most recently) relative to the

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

fixation plane. To conserve the number of dots on the screen at all times, dots that exited the front or back of the volume wrapped to the opposite side. With a plane, such wrapping creates a strong apparent 3-D motion signal in the undesired direction. A dot cloud reduces this effect because the dots wrap at different times. To further reduce apparent motion in the undesired direction, the dots were given new pseudorandom (x, y) locations and inverted polarity (e.g., a dark dot became light) when they wrapped. For toward motion in the combined cue and perspective cue conditions, dots can potentially exit the volume at any depth because of the expanding optic flow pattern. This would create uncontrolled motion noise that would be unmatched across cue conditions. Moreover, depending on the procedure for assigning where the dots are redrawn, this could cause density gradients to form during a trial that, if detected, would reveal the direction of motion (e.g., this would occur if exiting dots were always redrawn at the back of the volume). To eliminate these confounds, all dots were pseudorandomly selected from a large bank (n = 10,000) of precomputed trajectories that traversed the full volume without exiting the aperture prematurely. Importantly, this ensured that the dot disparities were uniformly distributed throughout the volume for the entire stimulus duration regardless of the motion direction. Thus, it was not possible to perform the task based on the distribution of static disparities, dot sizes, or dot density on the screen.

Discrimination Task

Each trial proceeded as follows (Figure 1C). First, the monkey acquired fixation on a red circular target (0.28°, 4.45 cd/m^2), which was located at the center of the screen and aligned with the cyclopean eye. After 300 msec of fixation, a stimulus was presented for 1 sec while fixation was maintained. To discourage the animals from anticipating the end of the trial, the fixation target then remained for a pseudorandom duration between 250 msec and 1.5 sec selected from a truncated exponential growth function (mean = 1.1 sec; Chang, Thompson, et al., 2020; Kiani, Hanks, & Shadlen, 2008; Roitman & Shadlen, 2002). The fixation target then disappeared, and two red choice targets (each 0.28°, 4.45 cd/m^2) appeared 7.4° above/below the fixation target. The direction of motion was indicated as "toward" or "away" by making a saccade to the lower or upper target, respectively. A liquid reward was given for correct responses. Responses to 0% coherence trials were rewarded pseudorandomly with 50% probability. The intertrial interval was 1.5 sec. A trial was aborted without reward if fixation was broken before the appearance of the choice targets or if a choice was not made within 500 msec.

Task Training and Experimental Procedure

Task training began with combined cue stimuli at full coherence, centered on the fixation target. Initially, only the correct choice target was presented. The distractor target was slowly introduced by increasing its contrast (Chang, Doudlah, et al., 2020; Chang, Thompson, et al., 2020). After an accuracy of ~90% was reached with full distractor contrast, the three single-cue conditions at full coherence were introduced simultaneously. After an accuracy of ~90% was reached for each of these conditions, lower motion coherence levels were sequentially introduced for all of the cue conditions simultaneously. After all coherences were included, we trained at each grid location until sensitivity stabilized. The total training period, including chair, fixation, and task training, was ~1 year per monkey. Data collection then began.

In each experimental session, a visual field location was pseudorandomly selected from the 25 grid positions (Figure 2). Stimuli were pseudorandomly presented in a block structure. A block included one completed trial for each combination of cue condition and directionally signed motion coherence (4 × 13 = 52 stimuli). Data collection continued until sensitivity estimates converged for all cue conditions and visual field locations, such that the addition of the three most recent blocks impacted sensitivity estimates by ≤5%. The average number of completed trials at each location was 2871 (SD = 1107) for Monkey J (71,785 trials total) and 3802 (SD = 1734) for Monkey C (95,053 trials total).

Data Analysis

Behavioral Performance

Performance was quantified for each cue condition and location. The proportion of "toward" responses was calculated as a function of the directionally signed motion coherence. The Psignifit toolbox was used to fit a cumulative Gaussian to each case (Wichmann & Hill, 2001):

$$g(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \quad (1)$$

Here, $g(x)$ is a cumulative Gaussian, x is the directionally signed motion coherence, μ is the response bias, and $\frac{1}{\sigma}$ is the sensitivity.

Cue Integration

The integration of left-eye perspective cues, right-eye perspective cues, and stereoscopic cues was evaluated using cue integration theory, incorporating the possibility that the neural representations of the cues were not independent (Chang, Thompson, et al., 2020; Fulvio et al., 2020; Oruç et al., 2003). All fitting was performed using nonlinear least squares regression (lsqnonlin in MATLAB). Four physiologically plausible models were tested.

Model 1. The "two-cue" model assumed that perspective cues and stereoscopic cues are independent and that

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

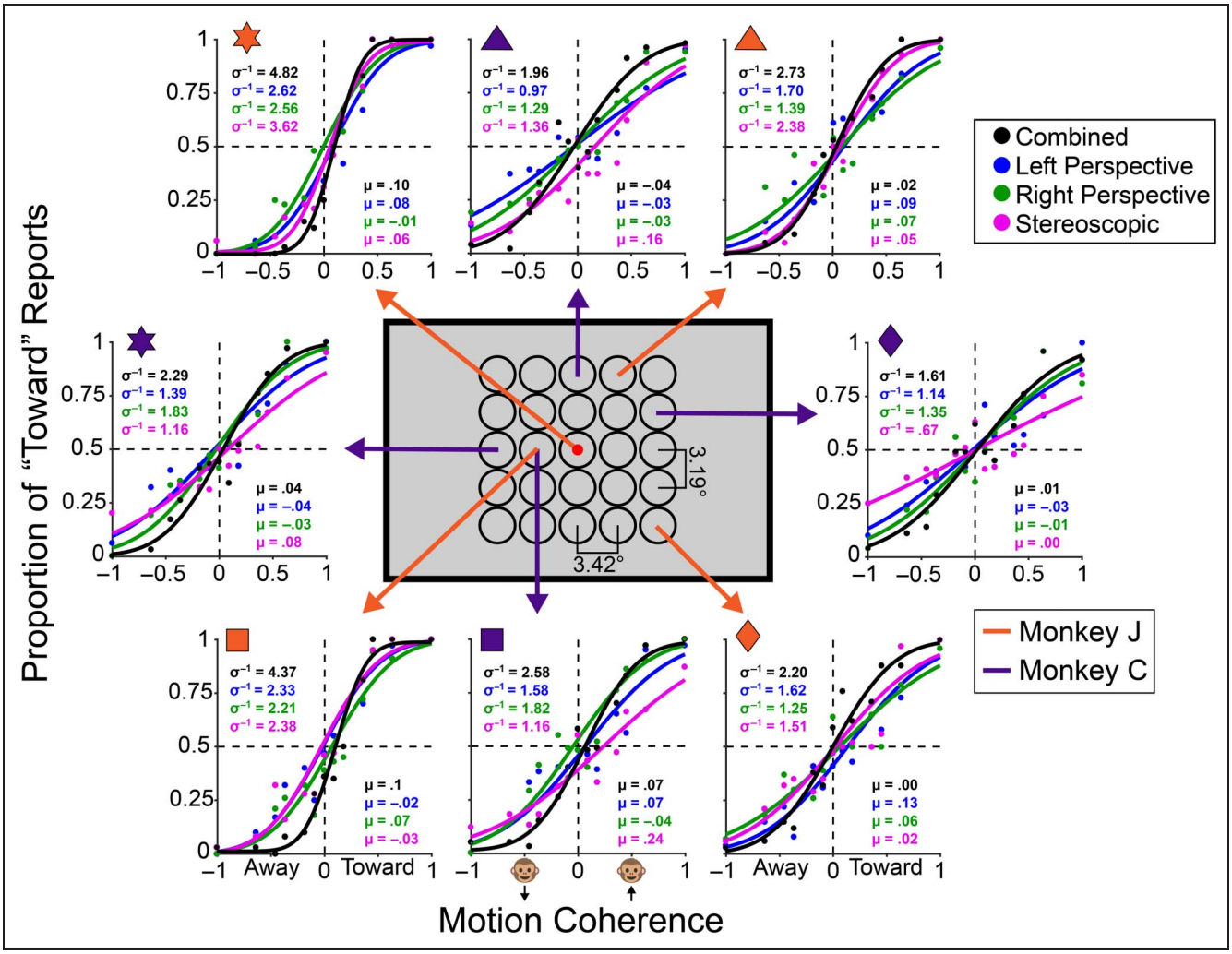


Figure 2. 3-D motion discrimination. The center depicts the visual display with each of the 25 tested visual field locations marked by a black circle. The red dot at the center corresponds to the fixation target. Surrounding plots show performance for each cue condition at eight example locations (Monkey J: orange arrows; Monkey C: purple arrows). Cue conditions: combined cue (black), left-eye perspective (blue), right-eye perspective (green), and stereoscopic (magenta). Data points show the proportion of "toward" reports for each of the 13 tested motion coherences. Solid curves are cumulative Gaussian fits. Vertical dashed lines mark 0 motion coherence. Horizontal dashed lines mark chance performance. Biases and sensitivities are reported in the insets. Performance was more accurate if the bias (μ) was closer to 0 and more precise if the sensitivity (σ^{-1}) was larger. Steeper psychometric curves reflect greater sensitivity. The colored symbols in the top left corners of the plots are used to mark these examples in Figures 3 and 4.

left- and right-eye perspective cues are indiscriminately pooled (i.e., a fully dependent representation) such that a single perspective signal contributes to perception:

$$\frac{1}{\sigma_{\hat{C}}} = \sqrt{\frac{1}{\sigma_P^2} + \frac{1}{\sigma_S^2}} \tag{2}$$

Here, $\frac{1}{\sigma_{\hat{C}}}$ is the predicted combined cue sensitivity; $\frac{1}{\sigma_P}$ is the single perspective cue sensitivity, which was estimated after pooling the responses to left- and right-eye perspective cue stimuli; and $\frac{1}{\sigma_S}$ is the stereoscopic cue sensitivity.

Model 2. The "three-cue" model assumed that all three cues are independently represented:

$$\frac{1}{\sigma_{\hat{C}}} = \sqrt{\frac{1}{\sigma_{P_L}^2} + \frac{1}{\sigma_{P_R}^2} + \frac{1}{\sigma_S^2}} \tag{3}$$

Here, $\frac{1}{\sigma_{P_L}}$ and $\frac{1}{\sigma_{P_R}}$ are the left- and right-eye perspective cue sensitivities, respectively.

Model 3. The "correlated perspective cues" model assumed that perspective cues and stereoscopic cues are

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

independent and that left- and right-eye perspective cue representations are partially dependent:

$$\frac{1}{\sigma_{\hat{C}}} = \sqrt{\frac{\frac{1}{\sigma_{P_L}^2} + \frac{1}{\sigma_{P_R}^2} - \frac{2\rho}{\sigma_{P_L}\sigma_{P_R}}}{1-\rho^2} + \frac{1}{\sigma_S^2}} \quad (4)$$

The degree of dependency between the left- and right-eye perspective cue representations is determined by $0 \leq \rho < 1$ (Oruç et al., 2003). Larger values of ρ indicate greater dependency. At $\rho = 0$, the correlated perspective cues model reduces to the three-cue model (all cues are independently represented). As ρ approaches 1, the left- and right-eye perspectives cue representations become fully dependent. A single value of ρ was determined separately for each monkey by fitting Equation 4 to all of the sensitivities measured across the visual field ($n = 25$ locations). In a second analysis, we examined whether ρ varied systematically across the visual field by separately fitting ρ at each visual field location and for each monkey.

Model 4. The "all cues correlated" model assumed dependencies between all three cues and therefore required three parameters to capture each of the pairwise dependencies. This model did not outperform the correlated perspective cues model and the equations are extensive (Oruç et al., 2003), so we do not reproduce them. The values of the three parameters were determined by fitting the model to all of the sensitivities measured across the visual field for each monkey ($n = 25$ locations).

Generalized Linear Model

Differences in sensitivity across the visual field, controlling for cue condition, were tested using the model:

$$\frac{1}{\sigma} \sim 1 + \sin(\theta) + \cos(\theta) + E + \lambda_1 + \lambda_2 + \lambda_3 \quad (5)$$

Here, a visual field location was described by its angular position (θ) and eccentricity (E). Three vectors (λ_1, λ_2, and λ_3) defined a set of "contrast codes" that specified the four cue conditions (Judd, McClelland, & Ryan, 2017).

Experimental Controls

Crosstalk Control

Although the crosstalk in our 3-D display was low, we wanted to test if "ghost images" in the nonstimulated eye may have affected our measurements of left- and right-eye perspective cue biases and sensitivities. We therefore remeasured the performance with the left- and right-eye perspective cue stimuli with either the left or right eye shutter physically covered by an opaque patch. The patched eye was pseudorandomly selected on different sessions. Performance was measured at the central visual field location (fixation, 0° eccentricity). This control was performed after all main experiment and other control experiments were completed (Monkey J: $n = 3314$ trials, $n = 5$ sessions; Monkey C: $n = 6031$ trials, $n = 6$ sessions).

Utrocular Control

We additionally wanted to rule out the possibility that the monkeys performed the task using utrocular information (Blake & Cormack, 1979; Smith, 1945). To test this possibility, we presented appropriate half-images of stereoscopic cue stimuli at full coherence (i.e., leftward or rightward translating dots) to a single eye at either four (Monkey J; eccentricities: 6.8°, 7.2°, 7.5°, and 9.3°) or three (Monkey C; eccentricities: 0° and two at 7.2°) locations. Thus, one eye saw 22 dots (11 light, 11 dark) translating left or right at 4.2°/sec. Both eyes saw the fixation target. The stimulated eye and direction of motion were counterbalanced. These utrocular control trials were interleaved into the main protocol in either seven (Monkey J: $n = 275$ trials) or four (Monkey C: $n = 229$ trials) sessions. To avoid training a behavioral response contingency based on these trials, responses to these few and infrequently presented stimuli were pseudorandomly rewarded.

Stereoscopic Cue Control

For the stereoscopic cue stimuli, it is possible that the monkeys perceived a conflict between the stereoscopically defined 3-D motion and absence of perspective-defined 3-D motion. If a conflict were perceived, it would be expected to increase with the number of dots and to cause stereoscopic cue sensitivities to be underestimated. On this basis, it has previously been suggested that stereoscopic cue sensitivity would decrease as a function of dot number if a conflict were perceived (Hillis et al., 2004). We therefore measured stereoscopic cue sensitivity using 6, 11, 22, 33, and 44 dots with the same coherences used in the main experiment (except for six dots, in which case all possible coherences were shown). Sensitivity was measured at a single visual field location (Monkey J: 6.8° eccentricity, $n = 10{,}692$ trials, $n = 9$ sessions; Monkey C: 3.2° eccentricity, $n = 8{,}695$ trials, $n = 6$ sessions).

Vergence Control Analysis

To test if the 3-D directional reports were related to small vergence eye movements that did not violate the online vergence window, we calculated the vergence velocity during each stimulus presentation (Sanada & DeAngelis, 2014) and used multivariate logistic regression to estimate the effect of vergence on the responses in each session:

$$\ln\left(\frac{P}{1-P}\right) \sim 1 + C + V \quad (6)$$

Here, P is the probability of reporting the direction as "toward," C is the directionally signed motion coherence, and V is vergence velocity.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

RESULTS

Quantifying 3-D Object Motion Perception

The primate visual system detects 3-D object motion as a pair of eye-specific 2-D patterns of retinal motion. Although macaque monkeys are a principal model system for studying the neural basis of 3-D vision, the cues upon which they rely to judge 3-D object motion is currently unknown. We therefore characterized the ability of two macaque monkeys to discriminate the direction of 3-D motion from these signals using stimuli that consisted of a cloud of dots, which moved either toward or away from the cyclopean eye (Figure 1A). Previous studies showed that humans perceive 3-D motion from each eyes' perspective cues as well as stereoscopic cues (Fulvio et al., 2020; Thompson et al., 2019). However, that work could not assess how the cues together shape 3-D motion perception. To determine the contributions of each of these cues to 3-D motion perception, we presented stimuli from four cue conditions in which 3-D motion was signaled by left-eye perspective cues, right-eye perspective cues, stereoscopic cues, or all three cues combined (Figure 1B).

The monkeys were trained to report the direction of 3-D motion in a "toward" versus "away" discrimination task (Methods). A trial was initiated by fixating on a target presented at the center of the screen for 300 msec (Figure 1C). A 3-D motion stimulus then appeared within a 3° circular aperture for 1 sec while fixation was maintained on the target. After a variable delay, the fixation target disappeared, and two choice targets appeared. The direction of motion was reported by making a saccade to one of the choice targets (toward = lower target, away = upper target). A liquid reward was provided for correct responses.

To further assess how perception varied across the visual field, we presented the stimuli at 25 nonoverlapping visual field locations (Figure 2, center). The location was varied across sessions but constant within a session. Performance was assessed using a motion coherence paradigm in which the proportion of dots that moved toward or away from the monkey was varied (Fulvio et al., 2020; Thompson et al., 2019). At a coherence of +1, all dots moved toward the monkey. At a coherence of −1, all dots moved away from the monkey. Psychometric curves showing the proportion of trials in which the direction of motion was reported as "toward" as a function of the directionally signed motion coherence are shown for each cue condition at eight example visual field locations in Figure 2.

Biases in 3-D Motion Perception Are Small but Cue Dependent

As the first step in characterizing the monkeys' 3-D motion perception, we evaluated potential systematic biases in their responses (Figure 2, biases are reported in the lower right inset for each example location). Across all cue conditions and visual field locations (n = 100 per monkey), there was little-to-no bias. A tendency to report the direction of motion as "toward" or "away" would result in a negative or positive bias, respectively. The mean biases were 0.04 ± 0.08 *SD* for Monkey J and 0.02 ± 0.11 *SD* for Monkey C. Although the biases were small, the mean bias was significantly different from 0 for Monkey J (Wilcoxon signed-rank test, $p < .001$), indicating a slight tendency to report "away." The biases were not significantly different from 0 for Monkey C ($p = .45$).

To further test if the biases were cue dependent, we separately examined the biases for each cue condition across all visual field locations. For the left-eye perspective cue stimuli, the mean bias was 0.02 ± 0.10 *SD* for Monkey J and 0.03 ± 0.08 *SD* for Monkey C ($n = 25$). Neither bias was significantly different from 0 (Wilcoxon signed-rank test, $p \geq .15$), indicating that the perceived direction of motion based on left-eye perspective cues was generally accurate. For the right-eye perspective cue stimuli, the mean bias was 0.04 ± 0.10 *SD* for Monkey J and not significantly different from 0 ($p = .10$). The mean bias for Monkey C was −0.07 ± 0.08 *SD* and significantly different from 0 ($p < .001$), indicating a tendency to report "toward." However, because the bias was still relatively small on average (see Figure 2 for visual reference), the perceived direction of motion based on right-eye perspective cues was also generally accurate. For the stereoscopic cue stimuli, the mean bias was 0.05 ± 0.06 *SD* for Monkey J and 0.12 ± 0.12 *SD* for Monkey C. Both biases were significantly different from 0 ($p < .001$), indicating that stereoscopic cues were associated with away biases for both monkeys. These cue-dependent differences in biases may reflect separate neural processing of the cues, consistent with reports that some macaque MT neurons are sensitive to stereomotion (Czuba, Huk, Cormack, & Kohn, 2014; Sanada & DeAngelis, 2014) but show little evidence for optic flow selectivity (Nakhla, Korkian, Krause, & Pack, 2021; Lagae, Maes, Raiguel, Xiao, & Orban, 1994). Lastly, for the combined cue stimuli (which contained perspective and stereoscopic cues), the mean bias for Monkey J was 0.05 ± 0.06 *SD* and significantly different from 0 ($p = .001$). Notably, this away bias was not present when the perspective cues were presented alone and was therefore attributable to the stereoscopic cues. The mean combined cue bias for Monkey C was −0.02 ± 0.08 *SD* and not significantly different from 0 ($p = .46$). For Monkey C, performance with the combined cue stimuli was therefore consistent with the away bias for stereoscopic cues being cancelled out by an opposite toward bias for right-eye perspective cues and is suggestive of cue integration.

Sensitivity to 3-D Motion Is Cue Dependent

Having found that the monkeys were generally accurate in their judgments of 3-D motion direction, we next needed to assess how their sensitivities depended on the defining visual cues to assess cue integration (Figure 2,

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

sensitivities are reported in the top left inset for each example location). To do so, we performed pairwise comparisons of the sensitivities between cue conditions. We first compared the left- and right-eye perspective cue sensitivities (Figure 3A). Monkey J was significantly more sensitive to left- than right-eye perspective cues (mean ratio = 1.11; Wilcoxon signed-rank test, $p = .028$). Monkey C was instead significantly less sensitive to left- than right-eye perspective cues (mean ratio = 0.80, $p < .001$). Importantly, because the presented 3-D motion directions were toward and away from the cyclopean eye, the left- and right-eye perspective cues were matched in each eye (Methods). The finding that both monkeys had significantly different sensitivities to these cues may reflect separate neuronal processing of left- and right-eye perspective cues to 3-D motion, which sharply contrasts with previous 3-D orientation perception results, which found that sensitivities to left- and right-eye perspective cues were indistinguishable (Chang, Thompson, et al., 2020).

We next compared the stereoscopic and perspective cue sensitivities (Figure 3B). Monkey J was significantly more sensitive to stereoscopic cues than the left- or right-eye perspective cues (mean ratios: 1.18 and 1.29, respectively; both ps $\leq .01$). Monkey C was similarly sensitive to stereoscopic and left-eye perspective cues (mean ratio = 0.97, $p = .10$), but significantly less sensitive to stereoscopic than right-eye perspective cues (mean ratio = 0.74, $p < .001$). Together with the comparison of left- and right-eye perspective cue sensitivities, these results reveal idiosyncratic differences in relative sensitivity to different 3-D motion cues, a well-documented feature of human 3-D motion processing (Fulvio et al., 2020; Thompson et al., 2019; Allen et al., 2015; Nefs et al., 2010).

We lastly compared the combined cue and single-cue sensitivities (Figure 3C). For both monkeys, the combined cue sensitivities were always greater than the single-cue sensitivities (Wilcoxon signed-rank test, all six ps $\leq 1.2 \times 10^{-5}$). The mean ratios of combined to single-cue sensitivities were: left-eye perspective (Monkey J: 1.66; Monkey C: 1.96), right-eye perspective (Monkey J: 1.82; Monkey C: 1.50), and stereoscopic (Monkey J: 1.45; Monkey C: 2.11). Importantly, this finding indicates that regardless of differences in how much the monkeys relied on each of the available 3-D motion cues, the cues were nevertheless integrated to achieve more precise 3-D motion perception. We systematically evaluate how the cues were integrated in the final section.

A Foveal Advantage in 3-D Motion Sensitivity

Previous studies with humans found that 3-D motion sensitivity varies across the visual field in idiosyncratic ways (Thompson et al., 2019; Barendregt, Dumoulin, & Rokers, 2014, 2016). We therefore wanted to evaluate if similar variability was observed for the two monkeys. First, we quantified how sensitivity depended on eccentricity, using a generalized linear model (GLM; Equation 5, omitting the cosine and sine covariates) to control for cue

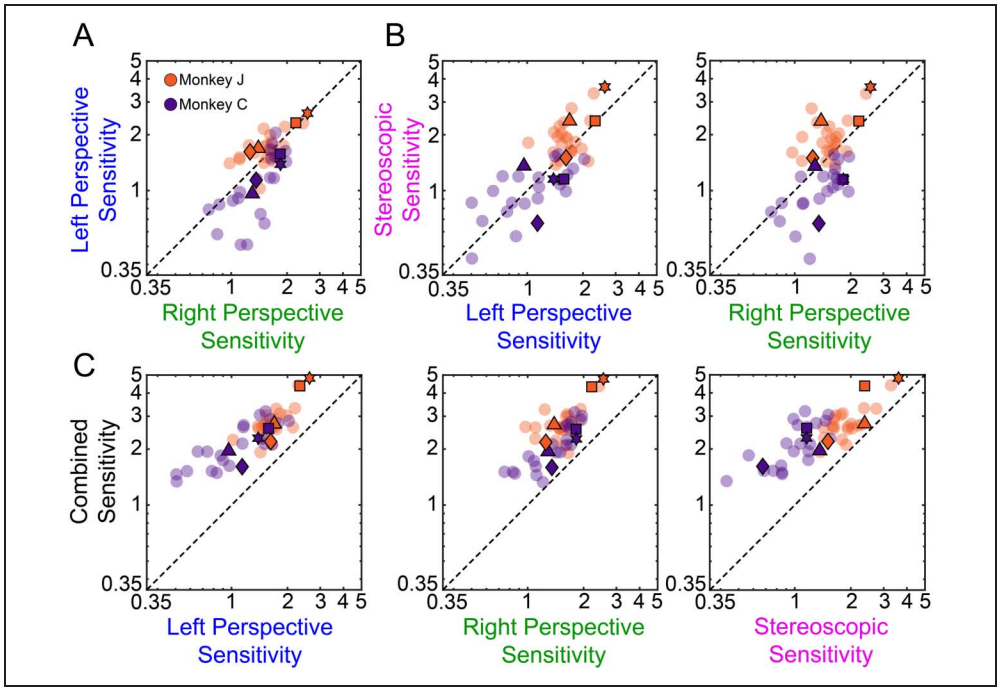


Figure 3. Sensitivity to 3-D motion is cue dependent. Each point corresponds to a single visual field location ($n = 25$ per monkey), plotted on a log scale. Dashed black lines are identity lines. (A) Comparison of left- and right-eye perspective cue sensitivities. For Monkey J, left-eye sensitivities were generally greater than right-eye sensitivities. For Monkey C, right-eye sensitivities were generally greater than left-eye sensitivities. (B) Comparison of stereoscopic and perspective cue sensitivities. For Monkey J, stereoscopic cue sensitivities were generally greater than both left- and right-eye perspective cue sensitivities. For Monkey C, stereoscopic cue sensitivities were similar to the left-eye perspective cue sensitivities but generally lower than the right-eye perspective cue sensitivities. (C) Comparison of combined cue and single-cue sensitivities. Combined cue sensitivities were always greater than the single-cue sensitivities. Opaque symbols correspond to the examples in Figure 2.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

condition (Figure 4A). Sensitivity significantly decreased with eccentricity for both monkeys, indicating a foveal advantage (Monkey J: $b = -0.15, F(1, 95) = 76.18, p < .001$; Monkey C: $b = -0.11, F(1, 95) = 65.34, p < .001$). Moreover, the interaction between eccentricity and cue condition was not statistically significant for either monkey, indicating that the reduction in sensitivity with greater eccentricity was similar for perspective, stereoscopic, and combined cue stimuli (ANCOVA; Monkey J: $F(3, 92) = 2.42, p = .071$; Monkey C: $F(3, 92) = 0.82, p = .49$). To more thoroughly visualize how sensitivity varied across the visual field, we removed the cue-dependent differences in sensitivity by separately z-scoring the sensitivities for each cue condition across all visual field locations. We then averaged the z-scored sensitivities across the four cue conditions at each location and used spline interpolation to generate sensitivity heat maps across the visual field (Figure 4B).

The sensitivity heat maps highlighted the falloff in sensitivity with eccentricity for both monkeys and further suggested that the decrease might not be isotropic for Monkey J. To test if the sensitivities (not z-scored) depended on the angular position of the stimulus, we included the cosine and sine components of the polar angle as covariates in the GLM (Equation 5; Fisher, Lewis, & Embleton, 1987). The cosine component reflects variability in sensitivity along the horizontal axis and was not significant for either monkey (Monkey J: $b = -0.02, F(1, 93) = 0.15, p = .70$; Monkey C: $b = 0.05, F(1, 93) = 1.03, p = .31$). Thus, sensitivity was similar to the left and right of fixation. The sine component reflects variability in sensitivity along the vertical axis. We found that this component was significant for Monkey J ($b = 0.14, F(1, 93) = 5.85, p = .017$), but not Monkey C ($b = -0.05, F(1, 93) = 1.06, p = .31$). The positive regression weight for Monkey J indicates that performance was better in the upper visual field than the lower visual field, as seen in the z-scored sensitivity plot (Figure 4B, left). To further test if the upper/lower visual field sensitivity difference in Monkey J was cue dependent, we repeated this analysis separately for each cue condition. The effect was most prevalent for the stereoscopic cues ($b = 0.27, F(1, 21) = 0.85, p = .035$). None of the other cue conditions were statistically significant, but the p value in the combined cue condition approached significance ($p = .074$) whereas the left- and right-eye perspective cue p values did not (both ps $\geq .88$). Thus, the anisotropy in sensitivity across the visual field in Monkey J was most attributable to stereoscopic cues. These analyses thus reveal a foveal advantage in 3-D motion sensitivity for both perspective and stereoscopic cues, have some similarities to human data showing idiosyncratic differences in how 3-D motion sensitivity varies across the visual field, and further demonstrate a dissociation of perspective and stereoscopic cue sensitivities.

Experimental Controls

Before systematically evaluating how the monkeys perceptually integrated the 3-D motion cues, we wanted to rule out several potential confounds. First, we wanted to verify that crosstalk in the 3-D display did not generate "ghost images" of the left- or right-eye stimuli in the nonstimulated eye that could be used to perform the task. To test this, we assessed performance for the left- and right-eye perspective cue stimuli with either the left or right eye's shutter covered with an opaque patch

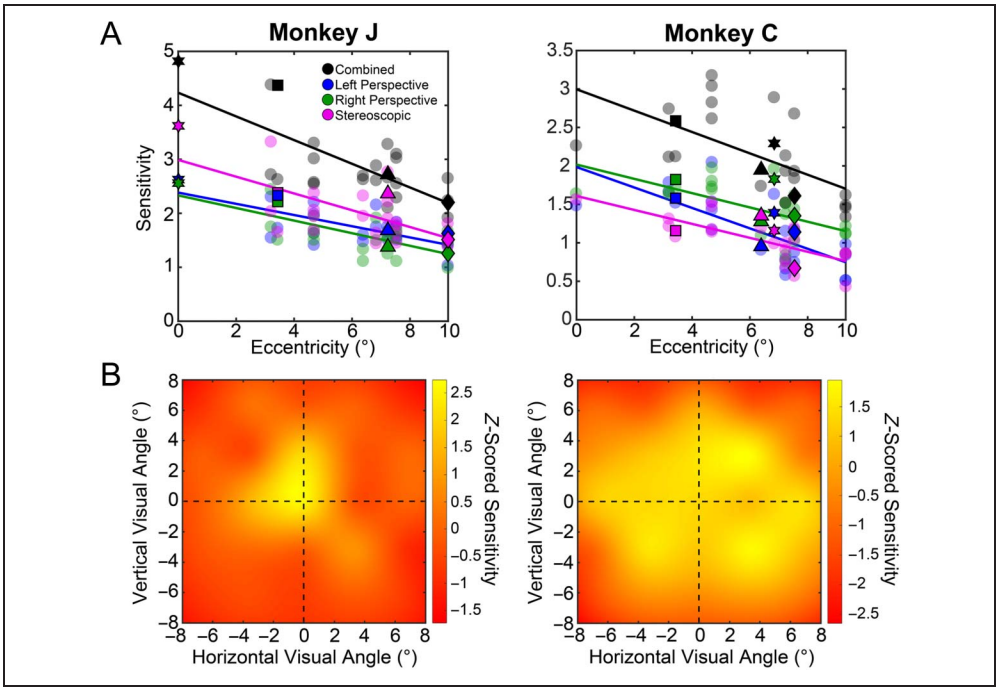


Figure 4. Sensitivity to 3-D motion varies across the visual field. (A) Sensitivity at each tested visual field location (points) plotted as a function of eccentricity for each cue condition (colors) and fit with a GLM (solid colored lines). Opaque symbols correspond to the examples in Figure 2. (B) Heat maps showing z-scored sensitivities across the visual field (averaged over cue conditions and spline interpolated). Yellow hues indicate greater sensitivity. Dashed lines mark the horizontal and vertical meridians, their intersection marks the point of fixation.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

(the patched eye was varied across sessions; Methods). Performance during these control sessions is shown in Figure 5. Importantly, when the stimuli were presented to the patched eye, the psychometric data were conspicuously flat. Indeed, for all four data sets (2 monkeys × 2 patched eyes) in which 3-D motion was presented to the patched eye, the fitted psychometric function parameters had implausibly large biases (all four ≤ −1.14, whereas −1 is the smallest possible coherence) and very low sensitivities that were smaller than any measured in the main experiment (all four ≤ 0.35). Furthermore, none of the correlations between the data and fits were significant (all four $p \geq .07$). These findings thus indicate that if any ghost images were detectable by the nonpatched eye, then they were insufficient to perform the task.

Second, we wanted to confirm that the monkeys did not use utrocular information (i.e., knowledge of which eye was stimulated) to discriminate the direction of 3-D motion using a simple heuristic. To clarify the concern, note that if they could discern which eye was stimulated by which 2-D pattern of retinal motion, then the task could be performed using a contingency table (e.g., "look down if the right eye sees net leftward motion"). To test this, we interleaved "utrocular control trials" into a small subset of the main experiments (Methods). These trials consisted of appropriate half-images of the stereoscopic cue stimuli (full coherence) that were presented to the left or right eye only. Importantly, neither monkey performed above chance on these trials (Binomial test; Monkey J: $p = .72$, $n = 275$ trials; Monkey C: $p = .091$, $n = 229$ trials). This result suggests that the monkeys did not use a utrocular heuristic to discriminate the direction of 3-D motion.

Third, we considered the possibility that sensitivity to the stereoscopic cue stimuli was underestimated because of a potential cue conflict between the stereoscopically defined 3-D motion and absence of perspective-defined 3-D motion. If a conflict were perceived, it would increase with the number of dots defining the stimulus and would appear as a decrease in sensitivity with increasing dot number (Hillis et al., 2004). We therefore performed a dot number control experiment in which sensitivity was measured using stereoscopic cue stimuli defined by 6, 11, 22, 33, or 44 dots (Methods). For Monkey J, we found a small positive (importantly, not negative) effect of dot number on sensitivity ($b = 0.006$, $F(1, 41) = 4.57$, $p = .039$), consistent with an increase in visual signal strength. For Monkey C, sensitivity did not significantly depend on the number of dots ($b = -0.001$, $F(1, 28) = 0.08$, $p = .77$). These results thus suggest that stereoscopic cue sensitivity was not affected by this potential cue conflict, consistent with recent 3-D orientation perception data (Chang, Thompson, et al., 2020).

Lastly, we wanted to rule out the possibility that perception was significantly affected by vergence eye movements. Following a previous study (Sanada & DeAngelis, 2014), we compared the vergence velocities during toward and away motions in the stereoscopic cue condition. For each session, we performed an ANOVA and found that significant differences between the toward and away vergence velocities were relatively rare (Monkey J: 10%; Monkey C: 6%). More critically, the direction of the vergence eye movements was not consistently related to the direction of stimulus motion, and the across-trial average difference between toward and away vergence velocities was very small (Monkey J: 0.004°/s; Monkey C: 0.006°/sec) compared with the 2° disparity

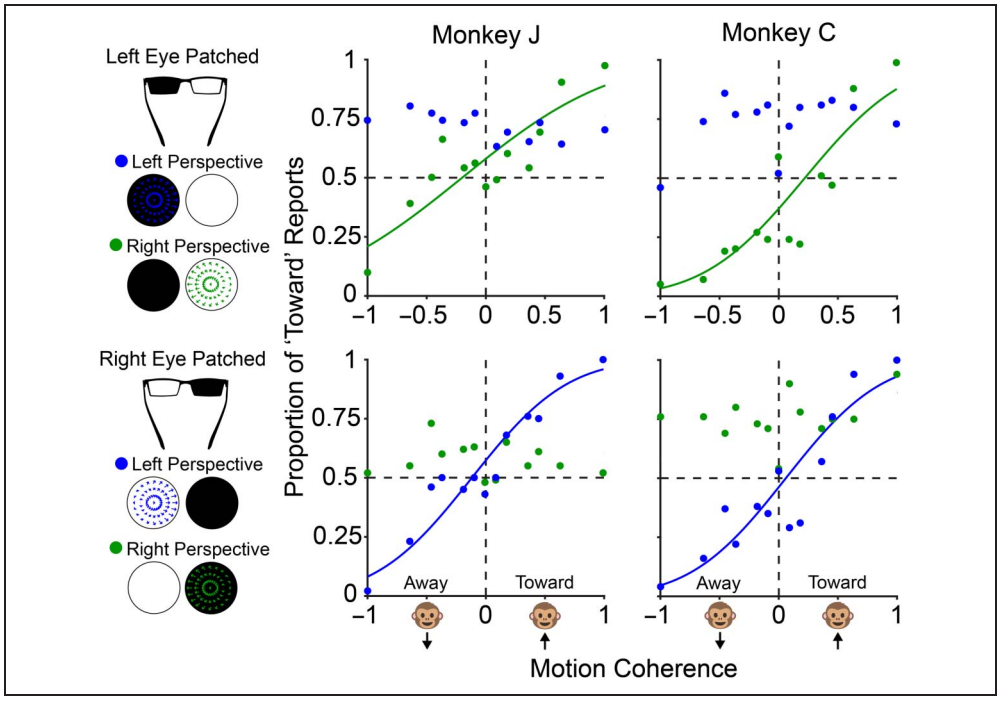


Figure 5. Crosstalk control. The left margin shows schematics of the viewing conditions (left or right eye patched) along with screen projected vector fields for the left- and right-eye perspective cue stimuli (following Figure 1). The four subplots to the right (plotted as in Figure 2) show the perspective cue performance for each monkey (columns) with either the left (top row) or right (bottom row) eye patched. Solid curves are cumulative Gaussian fits to the unpatched eye data. The patched-eye fits could not be significantly fit with a cumulative Gaussian.

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

volume. We also used multivariate logistic regression to estimate the effect of vergence on the directional reports (Methods). For both monkeys and all main experiment sessions, we found that the directionally signed motion coherence was a significant predictor of the reported motion direction when vergence velocity was not included as a covariate (all ps $\leq 1.88 \times 10^{-5}$). Importantly, when vergence velocity was included as a covariate, it was significantly related to the directional reports in only a minority of the sessions (Monkey J: 5.7%; Monkey C: 11.3%) and never affected the significance of the main effect of coherence. Together, these results suggest that the vergence eye movements were well controlled and had a negligible impact on perception.

Separate Contributions of Left- and Right-eye Perspective Cues to 3-D Motion Perception

The combined cue sensitivities were always larger than the single-cue sensitivities (Figure 3C), implying that the monkeys perceptually integrated the cues. We therefore wanted to determine how the left-eye perspective, right-eye perspective, and stereoscopic cues contributed to 3-D motion perception. To do so, we compared the observed combined cue sensitivities to predicted sensitivities from four cue integration models that reflect different physiologically plausible scenarios for how the cues might be represented and combined (Chang, Thompson, et al., 2020; Fulvio et al., 2020; Oruç et al., 2003). To visualize how well the models explained the observed sensitivities, we normalized each model prediction by the corresponding observed combined cue sensitivity (i.e., predicted/observed sensitivity ratios were calculated) and averaged over all visual field locations. Statistics were directly performed using the observed and predicted sensitivities.

The first model assumed that 3-D motion perception relies on independent perspective and stereoscopic cue representations, and that left- and right-eye perspective cues are indiscriminately pooled into a single fully dependent representation of perspective information (Chang, Thompson, et al., 2020; Fulvio et al., 2020; Oruç et al., 2003). We therefore refer to this model as the "two-cue" model (Equation 2) and estimated the pooled perspective cue sensitivity by fitting a cumulative Gaussian to the responses to all of the perspective cue trials regardless of the stimulated eye. We tested this model first because it accounts well for the integration of perspective and stereoscopic cues for 3-D orientation perception in both humans (Hillis et al., 2004; Knill & Saunders, 2003) and monkeys (Chang, Thompson, et al., 2020), and its ability to account for 3-D motion perception is unclear (Fulvio et al., 2020). In sharp contrast to previous 3-D orientation perception results, we found that the model systematically underestimated both monkeys' 3-D motion sensitivities (Figure 6; compare bar 1: observed combined cue sensitivity and bar 2: two-cue predicted sensitivity). The mean ratio of predicted to observed sensitivities was 0.92 for Monkey J and 0.77 for Monkey C. For both monkeys, the predicted and observed sensitivities were significantly different (Wilcoxon signed-rank test; both ps $\leq 6.6 \times 10^{-4}$). We also considered a variation of the two-cue model in which only the larger of the left- and right-eye perspective cue sensitivities contributed to combined cue perception (as might occur if perspective signals from the two eyes competed in a winner-take-all circuit). This version of the model also significantly underestimated both monkeys' sensitivities (both ps $\leq .032$). These results thus imply that both left- and right-eye perspective cues were utilized to estimate the direction of 3-D motion and that their representations were not fully dependent.

The second model assumed that left-eye perspective, right-eye perspective, and stereoscopic cues make fully independent contributions to 3-D motion perception. We therefore refer to this model as the "three-cue" model (Equation 3). Whereas the two-cue model sets a lower bound on combined cue sensitivity based on maximum likelihood estimation, the three-cue model sets an upper bound (Chang, Thompson, et al., 2020; Fulvio et al., 2020). We found that this model systematically overestimated the combined cue sensitivity of Monkey J but accurately predicted the sensitivity of Monkey C (Figure 6; compare bar 1: observed sensitivity and bar 3: three-cue predicted sensitivity). For Monkey J, the mean ratio of the predicted to observed sensitivities was 1.10, and the sensitivities were significantly different (Wilcoxon signed-rank test, $p < .001$). For Monkey C, the mean ratio was 1.01, and the sensitivities were not significantly different ($p = .82$). Together with the two-cue model, these results imply that left- and right-eye perspective cues make at least partially separate contributions to 3-D motion perception and that the degree of dependency differs across individuals. Importantly, this demonstrates that 3-D motion processing makes more effective use of available sensory signals than has been observed for 3-D orientation processing.

We therefore wanted to assess the degree of dependency between the representations of left- and right-eye perspective cues for each monkey. To do so, we implemented a third model, which assumed that perspective and stereoscopic cues are independent but that left- and right-eye perspective cue representations are partially dependent. The degree of dependency is determined by a correlation parameter, ρ (Equation 4; Oruç et al., 2003). We therefore refer to this model as the "correlated perspective cues" model and fit a single value of ρ for each monkey using the sensitivities measured across all 25 visual field locations. The combined cue sensitivities of Monkey J were best accounted for by $\rho = 0.36$, indicating a moderate degree of dependency. In contrast, the combined cue sensitivities of Monkey C were best accounted for by $\rho = 0.04$, indicating that the left- and right-eye perspective cue representations were nearly independent, consistent with the accurate prediction of the

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

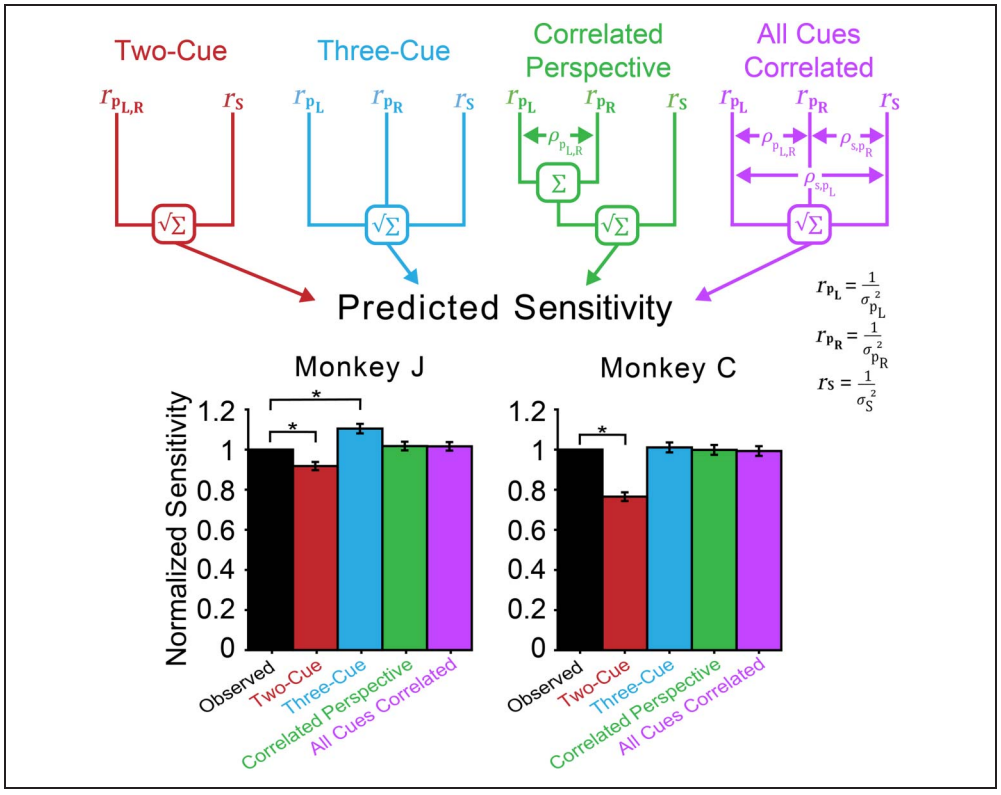


Figure 6. Comparison of observed combined cue sensitivities and predictions of the four cue integration models. Schematics (top) illustrate how each model combines the reliabilities (squared sensitivities) of the left-eye perspective (r_{P_L}), right-eye perspective (r_{P_R}), and stereoscopic cues (r_S). Bar plots (bottom) show sensitivities normalized by the observed combined cue sensitivities and averaged over visual field locations for Monkeys J (left) and C (right). Error bars show standard errors of the mean. Black brackets indicate models whose predictions significantly differed from the observed combined cue sensitivities (Wilcoxon signed-rank test, $p < .05$). Bar 1 (black): Observed combined cue sensitivity (normalized to one for each visual field location). Bar 2 (red): The two-cue model assumed left- and right-eye perspective cues are indiscriminately pooled ($r_{P_{L,R}}$) and integrated with an independent stereoscopic cue representation. This model underpredicted the combined cue sensitivities of both monkeys. Bar 3 (blue): The three-cue model assumed that all three cues provided independent estimates of 3-D motion. This model overpredicted the combined cue sensitivity of Monkey J but accurately predicted the combined cue sensitivity of Monkey C. Bar 4 (green): The correlated perspective cues model assumed that left- and right-eye perspective cue representations were partially dependent ($\rho_{P_{L,R}}$) and integrated with an independent stereoscopic cue representation. This model accurately predicted the combined cue sensitivities of both monkeys. Bar 5 (purple): The all cues correlated model assumed that all three cues were partially dependent on one another (with three pairwise ρ parameters). This elaborate model did not outperform the correlated perspective cues model.

three-cue model. The finding that ρ differed between the monkeys is consistent with individual differences in the dependency of linear perspective and texture gradient cues for slant perception in humans (Oruç et al., 2003). Moreover, the difference in ρ values could not be explained by a difference in the magnitude of left- versus right-eye perspective cue sensitivities, because these were not significantly different for the two monkeys (Wilcoxon signed-rank test, $p = .35$). The sensitivities predicted by this model were highly consistent with the combined cue sensitivities of both monkeys (Figure 6; compare bar 1: observed sensitivity and bar 4: correlated perspective cues predicted sensitivity). The mean ratio of the predicted to observed sensitivities was 1.02 for Monkey J and 1.0 for Monkey C. For both monkeys, the predicted and observed sensitivities were not significantly different (Wilcoxon signed-rank test, $p \geq .51$). Thus, this result shows that the combined cue 3-D motion sensitivities of both monkeys could be explained by a single model in which perspective and stereoscopic cue representations are independent and left- and right-eye perspective cue representations range between partially dependent and nearly independent.

We also considered that the dependency between the left- and right-eye perspective cues might systematically depend on the location within the visual field (e.g., it might change with eccentricity). Because the correlated perspective cues model flexibly captures the degree of (in)dependence between the left- and right-eye perspective cue representations, this possibility can be tested by examining if ρ varies systematically across the visual field. For each monkey, we therefore fit ρ at each of the 25 visual field locations and used a general linear mixed effects model to test for dependencies on eccentricity and angular position (treating monkey as a random effect). Importantly, we found that ρ did not significantly depend on eccentricity ($b = -0.02$, $F(1, 46) = 0.77$, $p = .39$), the horizontal (cosine) component of the angular position ($b = 0.06$, $F(1, 46) = 93$, $p = .34$), or the vertical (sine) component of the angular position ($b = 0.02$, $F(1, 21) = 0.13$, $p = .72$). Thus, we found no indication that the dependency between left- and right-eye perspective cues systematically varied across the visual field.

Together, the above findings also rule out that the left- and right-eye perspective cues rivaled, such that perspective information from only one eye was represented at a time. If they had rivaled, then the overall contribution of perspective cues to combined cue perception would have been a linear combination of the left- and right-

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

eye sensitivities weighted by their relative dominance. In that case, the overall contribution of perspective cues to perception would have been intermediate to the individual sensitivities. Instead, the overall contribution was greater than either of the individual sensitivities.

We lastly considered the possibility that a more elaborate model might better account for the monkeys' combined cue sensitivities. To test this, we implemented a model in which all of the 3-D motion cues are at least partially dependent upon one another. This model therefore included all possible pairwise correlations between the single cues, and we refer to it as the "all cues correlated" model (Oruç et al., 2003). Despite having two more free parameters than the correlated perspective cues model, the two models performed similarly (Figure 6; compare bar 4: correlated perspective cues and bar 5: all cues correlated).

To determine which single model provided the most parsimonious account of the data, we performed model comparisons using Akaike's information criterion (AIC). For each model, we computed an Akaike weight (a normalization of all of the models' AIC values such that they sum to 1; Burnham & Anderson, 2010). Larger values indicate greater model support. The largest AIC weight was for the correlated perspective cues model (0.79), which outperformed the all cues correlated model (0.16), the three-cue model (0.05), and the two-cue model (1.5×10^{-5}). Thus, 3-D motion perception was best described by a model in which 3-D motion estimates were independently computed from stereoscopic and perspective cues, and left- and right-eye perspective cues made at least partially independent contributions to perception. This finding contrasts sharply with previous human and monkey 3-D orientation perception results (Chang, Thompson, et al., 2020; Thompson et al., 2019; Hillis et al., 2004; Knill & Saunders, 2003), which were best described by the two-cue model. The current findings thus reveal a core difference in how perspective signals detected by the left and right eyes are combined for 3-D orientation versus 3-D motion perception, such that motion processing makes more effective use of available sensory information.

DISCUSSION

We investigated the contributions of left-eye perspective, right-eye perspective, and stereoscopic cues to 3-D motion perception in macaque monkeys. Consistent with previous human data, we observed idiosyncratic differences in 3-D motion cue sensitivity. These idiosyncrasies may reflect that estimating 3-D motion from 2-D retinal signals is an ill-posed problem with multiple potential solution strategies. As such, visual experience and physiological differences (e.g., in binocular integration) may result in individuals developing different relative sensitivities to the available cues. Importantly, despite such idiosyncracies, both monkeys perceptually integrated the cues to achieve more precise 3-D motion perception. Most importantly, we investigated a fundamental yet largely neglected question: How is the perspective information sensed by each eye integrated with stereoscopic information to achieve robust 3-D perception? To assess this, we tested four physiologically plausible cue integration models which differed in the extent to which the representations of each of the cues were dependent upon one another.

Model comparisons revealed that perception was best explained if 3-D motion estimates were independently computed from perspective and stereoscopic cues, and the representations of left- and right-eye perspective cues were at least partially independent. This finding implies that the identity of left- and right-eye perspective signals is preserved in the neuronal processing of 3-D motion, whereas it is not for 3-D orientation processing (Chang, Thompson, et al., 2020; Hillis et al., 2004; Knill & Saunders, 2003). By not preserving the ocular identity of perspective signals to 3-D orientation, substantial information loss occurs resulting in poorer sensitivity. In contrast, 3-D motion processing preserves ocular identity at least partially, mitigating that information loss. Intriguingly though, control experiments revealed that the ocular origins of the two perspective signals did not extend to the perceptual level because the monkeys did not use utrocular (eye-of-origin) information to perform the task. These findings thus suggest that 3-D motion perception reflects neuronal activity at or after the level of cue integration.

Together with previous 3-D orientation studies, the current results revealed an important parallel in the processing of different 3-D features. Across domains, sensitivity to combined cue stimuli is consistent with the integration of independently represented perspective and stereoscopic cues. This is supported by our finding that stereoscopic cues, but not perspective cues, were consistently associated with away biases that were either abolished or preserved in the combined cue responses depending on the monkeys' perspective cue biases. Likewise, an anisotropy in Monkey J's sensitivity across the visual field was linked to stereoscopic cues but not perspective cues. These results may reflect that distinct computations are required to estimate 3-D information from perspective and stereoscopic cues and that those computations are performed before cue integration.

Where might these computations be performed for 3-D motion processing? Recordings from macaque monkeys suggest that some MT neurons may be selective for 3-D motion based on stereoscopic cues (Czuba et al., 2014; Sanada & DeAngelis, 2014) but not optic flow (Nakhla et al., 2021; Lagae et al., 1994). However, another study concluded that MT neurons are not genuinely selective for motion in depth (Maunsell & Van Essen, 1983), and neuroimaging results suggest that the fundus of the STS (FST), more than MT, processes stereomotion (Héjja-Brichard, Rima, Rapha, Durand, & Cottereau, 2020).

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

Numerous studies implicate the medial superior temporal areas (MSTd and MSTl/v; Sasaki, Angelaki, & DeAngelis, 2019; Duffy, 1998; Eifuku & Wurtz, 1998; Tanaka, Sugita, Moriya, & Saito, 1993) and ventral intraparietal area (VIP; Sunkara, DeAngelis, & Angelaki, 2015, 2016; Maciokas & Britten, 2010) in 3-D motion processing based on perspective cues and likely stereoscopic cues. A recent study suggests that self- and object-motion signals (from MSTd and MSTl/v, respectively) likely converge in VIP to estimate object motion in world coordinates (Sasaki, Anzai, Angelaki, & DeAngelis, 2020). In each of these areas, it will be important to test if selectivity for perspective cues depends on the stimulated eye. If the representation of toward/away motion is cue invariant, then depending on which eye is stimulated, neurons should respond best to optic flow patterns with opposite net 2-D motion directions, consistent with previous theoretical work (Sabatini & Solari, 2004; Poggio & Talbot, 1981). Indeed, some FST neurons respond similarly to opposite 2-D motion directions (Rosenberg, Wallisch, & Bradley, 2008) and may therefore have the requisite input to compute 3-D object motion.

The current results also revealed a fundamental difference in the representation of left- and right-eye perspective cues across 3-D feature domains. For 3-D orientation, sensitivity to combined cue stimuli is consistent with the two perspective signals being indiscriminately pooled. In contrast, for 3-D motion, the current findings suggest that the signals remain at least partially separate until the stage of cue integration, as recently hinted at by preliminary human data (Fulvio et al., 2020). Supporting this possibility, our within-animal comparisons revealed differences in sensitivity to balanced left- and right-eye perspective cues. Those differences could reflect neural and/or optical factors (Elliott et al., 2009), but regardless of their origin, the finding that cue integration yields greater increases in sensitivity for 3-D motion than orientation implies a difference in how perspective cues are represented across these feature domains. This previously unrecognized difference, which results in more effective use of available sensory information for 3-D motion than orientation processing, may reflect evolutionary pressures associated with avoiding and intercepting moving objects. It further highlights that, even within the same sensory modality, differences can exist in the implementation of fundamental computations such as cue integration and show that previously identified information-limiting processes may not be a general property of cue integration when the same cue type is detected by multiple corresponding sensors, such as the left and right eyes, ears, or hands.

Our findings revealed more effective processing of 3-D motion than previously found for 3-D orientation. But where might neural correlates of this difference be found? Existing data suggest that these features are processed by parallel streams, with a V3A ➔ posterior intraparietal area ➔ caudal intraparietal area pathway for 3-D orientation processing (Chang, Doudlah, et al., 2020; Elmore, Rosenberg, DeAngelis, & Angelaki, 2019; Alizadeh, Van Dromme, Verhoef, & Janssen, 2018; Van Dromme, Premereur, Verhoef, Vanduffel, & Janssen, 2016; Rosenberg et al., 2013; Nakamura et al., 2001) and a MT ➔ MST/FST ➔ VIP pathway for 3-D motion processing. What neurocomputational differences between these pathways might explain the feature-dependent differences in cue integration? Simulations show that the strength of divisive normalization at the level of units combining left- and right-eye perspective cues can produce a range of dependencies between the two cues (Chang, Thompson, et al., 2020). Specifically, weaker divisive normalization would make the two perspective signals more independent. Thus, in the integration of the two eyes' perspective cues, divisive normalization may be weaker in the 3-D motion than the 3-D orientation pathway. This hypothesis can be tested using binocularly decorrelated random dot stimuli, which maintain each eyes' perspective cues. A key prediction is that monocular and binocular stimulation will produce similar responses from neurons processing 3-D orientation (Chang, Thompson, et al., 2020), but binocular facilitation (beyond a component attributable to interocular velocity differences) for neurons processing 3-D motion. Moreover, given that the degree of dependency between left- and right-eye perspective cues differed across monkeys, we anticipate that individual differences in the strength of divisive normalization within the motion pathway will covary with behavioral performance.

Consistent with "foveal advantages" for 2-D motion discrimination (Bower, Bian, & Andersen, 2012; Orban, Van Calenbergh, De Bruyn, & Maes, 1985), self-motion perception (Crowell & Banks, 1993; Warren & Kurtz, 1992), and time-to-contact estimation (Regan & Vincent, 1995), we found that 3-D motion sensitivity decreased with eccentricity. These findings may partially stem from eccentricity-dependent decreases in 2-D direction selectivity in visual cortex (Orban, Kennedy, & Bullier, 1986). However, sensitivity to 3-D motion across the visual field is also idiosyncratic (Thompson et al., 2019; Barendregt et al., 2014, 2016). Here, the stimulus parameters (e.g., dot size) were independent of the visual field location, and not surprisingly, sensitivity decreased as eccentricity increased. Future studies can compare the eccentricity dependence of 2-D and 3-D motion sensitivity on these parameters to test if the foveal advantage in 3-D motion sensitivity can be fully explained by lower level 2-D motion sensitivity. It would also be worthwhile to test if variability in 3-D motion sensitivity across the visual field has any relationship to the natural scene statistics of 3-D object motion.

Lastly, differences in 3-D orientation and motion processing likely constrain 3-D feature binding and limit performance when both features are behaviorally relevant. For example, to catch an object such as a frisbee, it is necessary to estimate its direction of motion and its orientation. We are unaware of a systematic analysis of the relative contributions of these factors to performance in

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

interception tasks, but it seems plausible that errors have more to do to with the grasp configuration (an orientation computation) than with the hand intercepting the object (a motion computation), especially because observers are generally accurate at estimating 3-D motion direction (Fulvio, Rosen, & Rokers, 2015). As such, it will be important to assess how performance on tasks requiring both 3-D orientation and 3-D motion information is constrained by the processing of each of these features.

Acknowledgments

This work was supported by National Science Foundation LUCID Training Program (1545481) and McPherson Eye Research Institute Graduate Student Support Initiative Awards to L. W. T. and the Alfred P. Sloan Foundation (FG-2016-6468), Whitehall Foundation (2016-08-18), Greater Milwaukee Foundation (Shaw Scientist Award), McPherson Eye Research Institute Expanding Our Vision 2020 Award, and National Institutes of Health Grant (EY029438) to A. R. Further support was provided by National Institutes of Health Grant P51OD011106 to the Wisconsin National Primate Research Center, University of Wisconsin-Madison.

Reprint requests should be sent to Ari Rosenberg, Department of Neuroscience, School of Medicine and Public Health, University of Wisconsin-Madison, 1111 Highland Ave., WIMR-II, Office 5505, Madison, WI 53705, or via e-mail: ari.rosenberg@wisc.edu.

Author Contributions

Lowell W. Thompson: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing—Original draft; Writing—Review & editing. Byounghoon Kim: Data curation; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Writing—Review & editing. Zikang Zhu: Data curation; Formal analysis; Investigation; Visualization; Writing—Review & editing. Bas Rokers: Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Supervision; Validation; Writing—Review & editing. Ari Rosenberg: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing—Original draft; Writing—Review & editing.

Funding Information

Ari Rosenberg, Alfred P. Sloan Foundation (https://dx.doi.org/10.13039/100000879), grant number: FG-2016-6468. Ari Rosenberg, Greater Milwaukee Foundation (https://dx.doi.org/10.13039/100007046), grant number: Shaw Scientist Award. Ari Rosenberg, Whitehall Foundation (https://dx.doi.org/10.13039/100001391), grant number: 2016-08-18. Ari Rosenberg, National Institutes of Health (https://dx.doi.org/10.13039/100000002), grant numbers: EY029438, P51OD011106. Ari Rosenberg, McPherson Eye Research Institute, grant numbers: Expanding Our Vision 2020 Award. Lowell W. Thompson, National Science Foundation (https://dx.doi.org/10.13039/100000001), grant number: LUCID Training Program 1545481. Lowell W. Thompson, McPherson Eye Research Institute (https://dx.doi.org/10.13039/100000001), grant number: Graduate Student Support Initiative Award.

Diversity in Citation Practices

A retrospective analysis of the citations in every article published in this journal from 2010 to 2020 has revealed a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .408, W(oman)/M = .335, M/W = .108, and W/W = .149, the comparable proportions for the articles that these authorship teams cited were M/M = .579, W/M = .243, M/W = .102, and W/W = .076 (Fulvio et al., *JoCN*, 33:1, pp. 3–7). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this article report its proportions of citations by gender category to be as follows: M/M = .744, W/M = .116, M/W = .070, and W/W = .070.

REFERENCES

Alizadeh, A.-M., Van Dromme, I., Verhoef, B.-E., & Janssen, P. (2018). Caudal intraparietal sulcus and three-dimensional vision: A combined functional magnetic resonance imaging and single-cell study. *Neuroimage*, *166*, 46–59. https://doi.org/10.1016/j.neuroimage.2017.10.045, PubMed: 29080712

Allen, B., Haun, A. M., Hanley, T., Green, C. S., & Rokers, B. (2015). Optimal combination of the binocular cues to 3-D motion. *Investigative Ophthalmology & Visual Science*, *56*, 7589–7596. https://doi.org/10.1167/iovs.15-17696, PubMed: 26618651

Barendregt, M., Dumoulin, S. O., & Rokers, B. (2014). Stereomotion scotomas occur after binocular combination. *Vision Research*, *105*, 92–99. https://doi.org/10.1016/j.visres.2014.09.008, PubMed: 25286416

Barendregt, M., Dumoulin, S. O., & Rokers, B. (2016). Impaired velocity processing reveals an agnosia for motion in depth. *Psychological Science*, *27*, 1474–1485. https://doi.org/10.1177/0956797616663488, PubMed: 27677897

Blake, R., & Cormack, R. H. (1979). On utrocular discrimination. *Perception & Psychophysics*, *26*, 53–68. https://doi.org/10.3758/BF03199861

Bower, J. D., Bian, Z., & Andersen, G. J. (2012). Effects of retinal eccentricity and acuity on global-motion processing. *Attention, Perception, & Psychophysics*, *74*, 942–949. https://doi.org/10.3758/s13414-012-0283-2, PubMed: 22382583

Brooks, K. R. (2002). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision*, *2*, 2. https://doi.org/10.1167/2.3.2, PubMed: 12678584

Burnham, K. P., & Anderson, D. R. (2010). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York: Springer.

Chang, T.-Y., Doudlah, R., Kim, B., Sunkara, A., Thompson, L. W., Lowe, M. E., et al. (2020). Functional links between sensory

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

representations, choice activity, and sensorimotor associations in parietal cortex. *eLife*, *9*, e57968. https://doi.org/10.7554/eLife.57968, PubMed: 33078705
Chang, T.-Y., Thompson, L., Doudlah, R., Kim, B., Sunkara, A., & Rosenberg, A. (2020). Optimized but not maximized cue integration for 3-D visual perception. *eNeuro*, *7*, ENEURO.0411-19.2019. https://doi.org/10.1523/ENEURO.0411-19.2019, PubMed: 31836597
Cooper, E. A., van Ginkel, M., & Rokers, B. (2016). Sensitivity and bias in the discrimination of two-dimensional and three-dimensional motion direction. *Journal of Vision*, *16*, 5. https://doi.org/10.1167/16.10.5, PubMed: 27537702
Cormack, L. K., Czuba, T. B., Knöll, J., & Huk, A. C. (2017). Binocular mechanisms of 3-D motion processing. *Annual Review of Vision Science*, *3*, 297–318. https://doi.org/10.1146/annurev-vision-102016-061259, PubMed: 28746813
Crowell, J. A., & Banks, M. S. (1993). Perceiving heading with different retinal regions and types of optic flow. *Perception & Psychophysics*, *53*, 325–337. https://doi.org/10.3758/BF03205187, PubMed: 8483696
Czuba, T. B., Huk, A. C., Cormack, L. K., & Kohn, A. (2014). Area MT encodes three-dimensional motion. *Journal of Neuroscience*, *34*, 15522–15533. https://doi.org/10.1523/JNEUROSCI.1081-14.2014, PubMed: 25411482
Dakin, C. J., & Rosenberg, A. (2018). Gravity estimation and verticality perception. *Handbook of Clinical Neurology*, *159*, 43–59. https://doi.org/10.1016/B978-0-444-63916-5.00003-3, PubMed: 30482332
Duffy, C. J. (1998). MST neurons respond to optic flow and translational movement. *Journal of Neurophysiology*, *80*, 1816–1827. https://doi.org/10.1152/jn.1998.80.4.1816, PubMed: 9772241
Eifuku, S., & Wurtz, R. H. (1998). Response to motion in extrastriate area MSTl: Center-surround interactions. *Journal of Neurophysiology*, *80*, 282–296. https://doi.org/10.1152/jn.1998.80.1.282, PubMed: 9658050
Elliott, S. L., Choi, S. S., Doble, N., Hardy, J. L., Evans, J. W., & Werner, J. S. (2009). Role of high-order aberrations in senescent changes in spatial vision. *Journal of Vision*, *9*, 24. https://doi.org/10.1167/9.2.24, PubMed: 19271934
Elmore, L. C., Rosenberg, A., DeAngelis, G. C., & Angelaki, D. E. (2019). Choice-related activity during visual slant discrimination in macaque CIP but not V3A. *eNeuro*, *6*, ENEURO.0248-18.2019. https://doi.org/10.1523/ENEURO.0248-18.2019, PubMed: 30923736
Fisher, N. I., Lewis, T., & Embleton, B. J. J. (1987). *Statistical analysis of spherical data*. Cambridge, United Kingdom: Cambridge University Press. https://doi.org/10.1017/CBO9780511623059
Fulvio, J. M., Ji, M., Thompson, L., Rosenberg, A., & Rokers, B. (2020). Cue-dependent effects of VR experience on motion-in-depth sensitivity. *PLoS One*, *15*, e0229929. https://doi.org/10.1371/journal.pone.0229929, PubMed: 32150569
Fulvio, J. M., Rosen, M. L., & Rokers, B. (2015). Sensory uncertainty leads to systematic misperception of the direction of motion in depth. *Attention, Perception, & Psychophysics*, *77*, 1685–1696. https://doi.org/10.3758/s13414-015-0881-x, PubMed: 25828462
Héjja-Brichard, Y., Rima, S., Rapha, E., Durand, J.-B., & Cottereau, B. R. (2020). Stereomotion processing in the nonhuman primate brain. *Cerebral Cortex*, *30*, 4528–4543. https://doi.org/10.1093/cercor/bhaa055, PubMed: 32227117
Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, *4*, 1. https://doi.org/10.1167/4.12.1, PubMed: 15669906
Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195084764.001.0001
Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). *Data analysis: A model comparison approach to regression, ANOVA, and beyond* (3rd ed.). New York: Routledge, Taylor & Francis Group. https://doi.org/10.4324/9781315744131
Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience*, *28*, 3017–3029. https://doi.org/10.1523/JNEUROSCI.4761-07.2008, PubMed: 18354005
Kim, B., Kenchappa, S. C., Sunkara, A., Chang, T.-Y., Thompson, L., Doudlah, R., et al. (2019). Real-time experimental control using network-based parallel processing. *eLife*, *8*, e40231. https://doi.org/10.7554/eLife.40231, PubMed: 30730290
Kleiner, M., Brainard, D. H., Pelli, D. G., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 1–16.
Knill, D. C. (1998). Surface orientation from texture: Ideal observers, generic observers and the information content of texture cues. *Vision Research*, *38*, 1655–1682. https://doi.org/10.1016/S0042-6989(97)00324-6, PubMed: 9747502
Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, *43*, 2539–2558. https://doi.org/10.1016/S0042-6989(03)00458-9, PubMed: 13129541
Komatsu, H., Kinoshita, M., & Murakami, I. (2002). Neural responses in the primary visual cortex of the monkey during perceptual filling-in at the blind spot. *Neuroscience Research*, *44*, 231–236. https://doi.org/10.1016/s0168-0102(02)00149-9, PubMed: 12413651
Lagae, L., Maes, H., Raiguel, S., Xiao, D. K., & Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: A comparison of areas MT and MST. *Journal of Neurophysiology*, *71*, 1597–1626. https://doi.org/10.1152/jn.1994.71.5.1597, PubMed: 8064337
Landy, M. S., Banks, M. S., & Knill, D. C. (2011). Ideal-observer models of cue integration. In J. Trommershäuser, K. P. Körding, & M. S. Landy (Eds.), *Sensory cue integration* (pp. 5–29). Oxford, United Kingdom: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195387247.003.0001
Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*, 1432–1438. https://doi.org/10.1038/nn1790, PubMed: 17057707
Maciokas, J. B., & Britten, K. H. (2010). Extrastriate area MST and parietal area VIP similarly represent forward headings. *Journal of Neurophysiology*, *104*, 239–247. https://doi.org/10.1152/jn.01083.2009, PubMed: 20427618
Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*, 1148–1167. https://doi.org/10.1152/jn.1983.49.5.1148, PubMed: 6864243
Nakamura, H., Kuroda, T., Wakita, M., Kusunoki, M., Kato, A., Mikami, A., et al. (2001). From three-dimensional space vision to prehensile hand movements: The lateral intraparietal area links the area V3A and the anterior intraparietal area in macaques. *Journal of Neuroscience*, *21*, 8174–8187. https://doi.org/10.1523/JNEUROSCI.21-20-08174.2001, PubMed: 11588190
Nakhla, N., Korkian, Y., Krause, M. R., & Pack, C. C. (2021). Neural selectivity for visual motion in macaque area V3A. *eNeuro*, *8*, ENEURO.0383-20.2020. https://doi.org/10.1523/ENEURO.0383-20.2020, PubMed: 33303620
Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, *1*, 155. https://doi.org/10.3389/fpsyg.2010.00155, PubMed: 21833221

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021

Nover, H., Anderson, C. H., & DeAngelis, G. C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *Journal of Neuroscience*, *25*, 10049–10060. https://doi.org/10.1523/JNEUROSCI.1661-05.2005, PubMed: 16251454

Orban, G. A., Kennedy, H., & Bullier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: Influence of eccentricity. *Journal of Neurophysiology*, *56*, 462–480. https://doi.org/10.1152/jn.1986.56.2.462, PubMed: 3760931

Orban, G. A., Van Calenbergh, F., De Bruyn, B., & Maes, H. (1985). Velocity discrimination in central and peripheral visual field. *Journal of the Optical Society of America A: Optics and Image Science*, *2*, 1836–1847. https://doi.org/10.1364/josaa.2.001836, PubMed: 4067694

Oruç, İ., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision Research*, *43*, 2451–2468. https://doi.org/10.1016/S0042-6989(03)00435-8, PubMed: 12972395

Poggio, G. F., & Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology*, *315*, 469–492. https://doi.org/10.1113/jphysiol.1981.sp013759, PubMed: 7310720

Priebe, N. J., Lisberger, S. G., & Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *Journal of Neuroscience*, *26*, 2941–2950. https://doi.org/10.1523/JNEUROSCI.3936-05.2006, PubMed: 16540571

Regan, D., & Vincent, A. (1995). Visual processing of looming and time to contact throughout the visual field. *Vision Research*, *35*, 1845–1857. https://doi.org/10.1016/0042-6989(94)00274-P, PubMed: 7660591

Rohde, M., van Dam, L. C. J., & Ernst, M. (2016). Statistically optimal multisensory cue integration: A practical tutorial. *Multisensory Research*, *29*, 279–317. https://doi.org/10.1163/22134808-00002510, PubMed: 29384605

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, *22*, 9475–9489. https://doi.org/10.1523/JNEUROSCI.22-21-09475.2002, PubMed: 12417672

Rokers, B., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2011). Seeing motion with two eyes in three dimensions. *Journal of Vision*, *11*, 10. https://doi.org/10.1167/11.2.10, PubMed: 21317358

Rokers, B., Fulvio, J. M., Pillow, J. W., & Cooper, E. A. (2018). Systematic misperceptions of 3-D motion explained by Bayesian inference. *Journal of Vision*, *18*, 23. https://doi.org/10.1167/18.3.23, PubMed: 29677339

Rosenberg, A., Cowan, N. J., & Angelaki, D. E. (2013). The visual representation of 3-D object orientation in parietal cortex. *Journal of Neuroscience*, *33*, 19352–19361. https://doi.org/10.1523/JNEUROSCI.3174-13.2013, PubMed: 24305830

Rosenberg, A., Wallisch, P., & Bradley, D. C. (2008). Responses to direction and transparent motion stimuli in area FST of the macaque. *Visual Neuroscience*, *25*, 187–195. https://doi.org/10.1017/S0952523808080528, PubMed: 18442441

Sabatini, S. P., & Solari, F. (2004). Emergence of motion-in-depth selectivity in the visual cortex through linear combination of binocular energy complex cells with different ocular dominance. *Neurocomputing*, *58–60*, 865–872. https://doi.org/10.1016/j.neucom.2004.01.139

Sanada, T. M., & DeAngelis, G. C. (2014). Neural representation of motion-in-depth in area MT. *Journal of Neuroscience*, *34*, 15508–15521. https://doi.org/10.1523/JNEUROSCI.1072-14.2014, PubMed: 25411481

Sasaki, R., Angelaki, D. E., & DeAngelis, G. C. (2019). Processing of object motion and self-motion in the lateral subdivision of the medial superior temporal area in macaques. *Journal of Neurophysiology*, *121*, 1207–1221. https://doi.org/10.1152/jn.00497.2018, PubMed: 30699042

Sasaki, R., Anzai, A., Angelaki, D. E., & DeAngelis, G. C. (2020). Flexible coding of object motion in multiple reference frames by parietal cortex neurons. *Nature Neuroscience*, *23*, 1004–1015. https://doi.org/10.1038/s41593-020-0656-0, PubMed: 32541964

Seilheimer, R. L., Rosenberg, A., & Angelaki, D. E. (2014). Models and processes of multisensory cue combination. *Current Opinion in Neurobiology*, *25*, 38–46. https://doi.org/10.1016/j.conb.2013.11.008, PubMed: 24709599

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*, 2565–2572. https://doi.org/10.1016/S0042-6989(00)00130-9, PubMed: 10958908

Smith, S. (1945). Utrocular, or 'which eye,' discrimination. *Journal of Experimental Psychology*, *35*, 1–14. https://doi.org/10.1037/h0059224

Stevens, K. A. (1981). The information content of texture gradients. *Biological Cybernetics*, *42*, 95–105. https://doi.org/10.1007/BF00336727, PubMed: 7326290

Sunkara, A., DeAngelis, G. C., & Angelaki, D. E. (2015). Role of visual and non-visual cues in constructing a rotation-invariant representation of heading in parietal cortex. *eLife*, *4*, e04693. https://doi.org/10.7554/eLife.04693, PubMed: 25693417

Sunkara, A., DeAngelis, G. C., & Angelaki, D. E. (2016). Joint representation of translational and rotational components of optic flow in parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *113*, 5077–5082. https://doi.org/10.1073/pnas.1604818113, PubMed: 27095846

Tanaka, K., Sugita, Y., Moriya, M., & Saito, H. (1993). Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *Journal of Neurophysiology*, *69*, 128–142. https://doi.org/10.1152/jn.1993.69.1.128, PubMed: 8433128

Thompson, L., Ji, M., Rokers, B., & Rosenberg, A. (2019). Contributions of binocular and monocular cues to motion-in-depth perception. *Journal of Vision*, *19*, 2. https://doi.org/10.1167/19.3.2, PubMed: 30836382

Van Dromme, I. C., Premereur, E., Verhoef, B.-E., Vanduffel, W., & Janssen, P. (2016). Posterior parietal cortex drives inferotemporal activations during three-dimensional object vision. *PLoS Biology*, *14*, e1002445. https://doi.org/10.1371/journal.pbio.1002445, PubMed: 27082854

Warren, W. H., & Kurtz, K. J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics*, *51*, 443–454. https://doi.org/10.3758/BF03211640, PubMed: 1594434

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*, 1293–1313. https://doi.org/10.3758/BF03194544, PubMed: 11800458

Woods, A. J. (2012). Crosstalk in stereoscopic displays: A review. *Journal of Electronic Imaging*, *21*, 040902. https://doi.org/10.1117/1.JEI.21.4.040902

Downloaded from http://direct.mit.edu/jocn/article-pdf/34/1/192/1976383/jocn_a_01781.pdf by University of Wisconsin, Parkside user on 22 December 2021