The initial torsional Ocular Following Response (tOFR) in humans: A response to the total motion energy in the stimulus?

B. M. Sheliga
Laboratory of Sensorimotor Research, National Eye Institute, Bethesda, MD, USA

E. J. FitzGibbon
Laboratory of Sensorimotor Research, National Eye Institute, Bethesda, MD, USA

F. A. Miles
Laboratory of Sensorimotor Research, National Eye Institute, Bethesda, MD, USA

We recorded the initial torsional Ocular Following Responses (tOFRs) elicited at short latency by visual images that occupied the frontal plane and rotated about the lines of sight. Using 1-D radial gratings, the local spatio-temporal characteristics of these tOFRs closely resembled those we previously reported for the hOFRs to horizontal motion with 1-D vertical gratings. When the 1-D radial grating was subdivided into a number of concentric annuli, each with the same radial thickness, tOFRs were less than predicted from the sum of the responses to the individual annuli: spatial normalization. However, the normalization was much weaker than that which we previously reported for the hOFRs. Further, when the number, thickness and contrast of these concentric annuli were varied systematically, the latency and magnitude of the tOFRs were well described by single monotonic functions when plotted against the product of the total area of the annuli and the square of their Michelson contrast ($A^*C^2$), consistent with the hypothesis that the onset and magnitude of the initial tOFR are determined by the total motion energy in the stimulus. When our previously published hOFR data were plotted against $A^*C^2$, a single monotonic function sufficed to describe the latency but not the magnitude.

Keywords: cycloversion, winner-take-all, normalization, visual motion


Introduction

This paper is concerned with the eye movements that are elicited when an observer experiences visual rotations about the naso-occipital (roll) axis. In our experiments, the observer faced a computer monitor that occupied the frontal plane and displayed a pattern that rotated in the plane of the screen and about its center. We concentrated on the special case in which the monitor was positioned so that the axis of this rotation passed midway between the two eyes and the observer’s gaze was directed at the center of the display. This meant that, although the visual rotation was defined with respect to the head, each eye saw rotation about its optic axis, and many studies have shown that this induces torsional tracking of the two eyes in the direction of the seen rotation, often termed cycloversion (Cheung & Howard, 1991; Cheung, Money, & Howard, 1995; Collewijn, van der Steen, Ferman, & Jansen, 1985; Farooq, Gottlob, Benskin, & Proudlock, 2008; Farooq, Proudlock, & Gottlob, 2004; Houben, Goumans, & van der Steen, 2006; Howard, Sun, & Shen, 1994; Howard & Templeton, 1964; Ibbotson, Price, Das, Hietanen, & Mustari, 2005; Kertesz & Jones, 1969; Lopez, Borel, Magnan, & Lacour, 2005; Morrow & Sharpe, 1993; Seidman, Leigh, & Thomas, 1992; Suzuki, Shinmei, Nara, & Ifukube, 2000; Thilo, Probst, Bronstein, Ito, & Greyst, 1999; van Rijn, van der Steen, & Collewijn, 1992, 1994a, 1994b; Wade, Swanston, Howard, Ono, & Shen, 1991; Washio, Suzuki, Sawa, & Ohtsuka, 2005; Zupan & Merfeld, 2003).

When such roll-axis visual rotation is prolonged, the pattern of torsional eye movements resembles classical optokinetic nystagmus, with periods of smooth tracking interrupted by resetting saccades (e.g., Cheung & Howard, 1991; Cheung et al., 1995; Collewijn et al., 1985; Farooq et al., 2004; Howard et al., 1994; Ibbotson et al., 2005). Horizontal and vertical optokinetic nystagmus (hOKN and vOKN)—for which the adequate stimuli are also generally defined with respect to the head and involve visual rotations about the yaw and pitch axes, respectively—show a gradual build-up over time and—after the visual stimulus is extinguished—an appreciable afternystagmus (OKAN), two features that have been attributed to a central...
velocity-storage mechanism (Cohen, Matsuo, & Raphan, 1977). Several studies have reported that the torsional optokinetic nystagmus (tOKN) elicited by roll-axis rotation does not show a gradual buildup and, at best, a very weak OKN, consistent with a very weak velocity-storage mechanism (e.g., Cheung & Howard, 1991; Cheung et al., 1995; Morrow & Sharpe, 1993), though a contrary view is offered by the study of Zupan and Merfeld (2003), which reported a reasonably robust tOKN.

All authors agree that this tOKN generally has a much lower steady-state gain than either hOKN or vOKN, maximum values of <0.1 being commonplace. A variety of rotating visual patterns have been used to elicit tOKN, including checkerboards (e.g., Collewijn et al., 1985), random dots (e.g., Cheung & Howard, 1991), and 1-D radial gratings (e.g., Farooq et al., 2004; Wade et al., 1991), and all yielded very similar steady-state gains. However, when the roll stimuli were 1-D horizontal or vertical grating patterns that oscillated about the optic axis, torsional ocular responses were much stronger with the horizontal gratings than with vertical ones (van Rijn et al., 1994a), an anisotropy that the authors suggested might reflect a special role for the horizon in everyday viewing.

The present report is the first to concentrate on the initial torsional eye movements elicited by brief roll-axis rotations and the emphasis is on the open-loop responses that occur during the period between one and two reaction times. There have been many previous studies of the initial ocular tracking responses elicited by other kinds of transient large-field visual stimuli, and these have identified three different reflexive eye movements that are thought to be involved in the visual stabilization of the moving observer’s gaze (for review of the older literature, see Miles, 1998). One of these reflexes, termed the Ocular Following Response (OFR), generates conjugate (version) eye movements in response to horizontal or vertical motion perpendicular to the line of sight—such as might result from head rotations about the yaw and pitch axes, respectively, and/or head translations in the coronal plane—and has been postulated to help prevent image motion in the plane of fixation. A second reflex, termed the Radial-Flow Vergence Response (RFVR), generates vergence eye movements in response to radial motion towards or away from the fovea—such as might result from fore-aft motion of the head—and has been postulated to help maintain binocular alignment on objects that lie ahead. A third reflex, termed the Disparity Vergence Response (DVR), generates vergence eye movements in response to binocular parallax (disparity)—such as might also result from fore-aft motion of the head—and has been assumed to work in parallel with the RFVR, helping to maintain binocular alignment on objects that lie ahead. All three reflexes have ultra-short latencies with machine-like response characteristics that closely reflect the detailed neural processing of the visual stimuli that elicit them, and there is strong evidence implicating the MT/MST area in their generation (Takemura, Murata, Kawano, & Miles, 2007). The torsional tracking responses in the present study are orthogonal to the OFRs (horizontal and vertical, hOFR and vOFR) studied previously and no combination of these cardinal responses can generate torsional responses, which also require a global decoding mechanism able to selectively sense visual rotations around the line of sight.

This report will document the initial torsional eye movements elicited by roll-axis rotations in a series of six main experiments that used a variety of visual stimuli specially designed to uncover the spatio-temporal characteristics and spatial summation properties. These visual stimuli were adapted from those used recently in studies of the hOFR, which used horizontal motions (Sheliga, Chen, FitzGibbon, & Miles, 2005; Sheliga, FitzGibbon, & Miles, 2008a; Sheliga, Kodaka, FitzGibbon, & Miles, 2006). In Experiment 1, which used random-dot patterns, the torsional eye movements displayed short latency and a skewed Gaussian dependence on log angular speed. In Experiment 2, which used 1-D radial gratings, the torsional eye movements showed a Gaussian dependence on log angular spatial frequency (“log-normal” distribution) and a sigmoidal dependence on contrast that saturated at relatively low levels (<20%). In Experiment 3, which used two overlapping 1-D radial gratings of different angular spatial frequency that rotated in opposite directions (the “3f stimulus”), the torsional eye movements showed a clear dependence on the motion of the two (Fourier) components rather than on the motion of the overall features, consistent with mediation by spatio-temporal filters sensitive to motion energy and also showed a highly nonlinear dependence on the relative contrast of the two gratings that resulted in Winner-Take-All (WTA) behavior when their contrasts differed more than an octave, consistent with mutual inhibition between the neuronal mechanisms mediating the responses to each of the two gratings. In Experiment 4, the roll-axis stimuli again consisted of two overlapping 1-D radial gratings of different angular spatial frequency similar to those in Experiment 3 except that the two gratings moved in the same direction (the “3f stimulus”). This dual grating again revealed a nonlinear dependence on relative contrast that resulted in WTA behavior when the two gratings differed in contrast by more than an octave. Thus, the clear implication was that the neural mechanisms sensitive to two different overlapping motions are negatively cross-coupled whether those motions are in the same or opposite direction. In Experiment 5, which used a single 1-D radial grating that was subdivided into a number of concentric annuli, each with the same radial thickness (0.5°–1.5°), the torsional eye movements were always less than predicted from the sum of the responses to the individual annuli (“less-than-linear-sum”), consistent with spatial normalization. In Experiment 6, which again used two 1-D radial gratings that rotated in opposite directions as in Experiment 3 (the “3f stimulus”) but now each was reduced to a single annulus of the same radial thickness...
(3°), WTA behavior was again evident but only when the annuli had the same radius (i.e., overlapped) and not when they differed in radius (i.e., were separated), indicating that the underlying nonlinear interactions were mostly local.

These various stimulus dependencies were, in all essentials, like those that we have previously reported for the hOFR and could even be described by the same mathematical functions, often with very similar parameter values (Sheliga et al., 2005, 2008a; Sheliga, FitzGibbon, & Miles, 2008b; Sheliga, Kodaka, et al., 2006). Accordingly, we refer to these initial torsional eye movements as tOFRs. However, the spatial normalization in Experiment 5 was much weaker than that which we had previously reported for the hOFR when horizontal motion was applied to a vertical grating subdivided into horizontal bands (Sheliga et al., 2008a) and this had a very interesting consequence: when we varied the thickness and contrast of the concentric annuli as well as their number (in Experiment 5), the latency and magnitude of most of the tOFR data were well described by single monotonic functions when plotted against the product of the Total Area of the annuli and the square of their Michelson contrast (“A*C^2”), as though the onset and magnitude of the tOFR were determined simply by the total motion energy in the stimulus. Reanalyzing our previously published hOFR data (Sheliga et al., 2008a) indicated that, when plotted against A*C^2, a single monotonic function sufficed to describe the latency but not the magnitude.

### Experiment 1: Random-dot patterns

We first describe the tOFRs elicited when broad-band random-dot patterns were rotated around the lines of sight. Major concerns were latency and speed dependence, and how these compared with those for the hOFRs elicited by horizontal translations. We also report a vertical anisotropy whereby masking off the lower half of the random-dot pattern reduced the initial amplitude of the tOFR more than masking off the upper half.

### Methods

Most of the techniques were very similar to those used previously in our laboratory (Sheliga et al., 2005; Sheliga, Chen, FitzGibbon, & Miles, 2006; Sheliga et al., 2008b; Sheliga, Kodaka, et al., 2006) and therefore will only be described in brief here. Experimental protocols were approved by the institutional review board concerned with the use of human subjects.

### Subjects

Three subjects participated: two were authors (FAM, BMS) and the third was a volunteer who was unaware of the purpose of the experiments (JKM). All had normal or corrected-to-normal acuity. Viewing was binocular for FAM and BMS and monocular for JKM (right eye viewing).

### Visual display and the random-dot stimuli

The subjects sat in a dark room with their heads positioned by means of adjustable rests for the forehead and chin and secured in place with a head band. For subjects FAM and BMS, stimuli were presented dichoptically using a Wheatstone mirror stereoscope but the two eyes always saw identical images (zero binocular disparity). Each eye viewed a computer monitor (ViewSonic G225F 21" CRT) through a 45° mirror, creating a vertical binocular surface straight ahead at 52.1 cm from the eye’s corneal vertex, which was also the optical distance to the images on the monitor screens. Each monitor was driven by an independent PC (Dell Precision 380) but the outputs of each computer’s video card (PC NVIDIA Quadro FX 5600) were frame-locked via NVIDIA Quadro G-Sync cards. This arrangement allowed the presentation of independent images simultaneously to each eye. Subject JKM viewed only the right monitor (left eye patched). Monitor screens were 400 mm wide × 300 mm high (subtense, 42° × 32°), with 1600 × 1200 pixels, and a vertical refresh rate of 100 Hz. The visual displays had a resolution of 36.6 pixels/° at the point directly ahead of each eye. The RGB signals from the video card provided the inputs to an attenuator whose output was connected to the “green” input of a video signal splitter (Black Box Corp., AC085A-R2); the three “green” video outputs of the splitter were then expanded to 2048 equally spaced levels by interpolation. The monitors were not turned off for the duration of the project and their luminance was checked for linearity at 2- or 3-week intervals (typically, r² = 0.99997). In the description that follows, we will refer only to the (single) binocular images.

The visual stimuli consisted of random dots (equal numbers of black and white dots, each circular with a diameter of ~0.5°) occupying a central circular area of the screen (diameter, 32°; dot-coverage, 50%; space-averaged luminance, 20.8 cd/m², as for the surround). The random-dot patterns were subjected to clockwise (CW) or counterclockwise (CCW) rotation around the screen center.
(“roll-axis rotation”). In Experiment 1A, concerned with speed dependence, the rotations could have one of 9 angular speeds randomly selected from a lookup table: 6.25, 12.5, 25, 50, 100, 200, 400, 800, and 1600°/s. In Experiment 1B, concerned with spatial anisotropies, the display always rotated at 135°/s and included trials in which the dots could fill the usual central circular area (32° in diameter), and other trials in which the dots in the lower or upper half of the screen were masked off. In Experiment 1C, the concern was to compare the tOFRs with hOFRs by determining the optimal speeds for, and minimum latencies of, the tOFRs elicited by roll-axis rotations (25, 50, 100, 200, 400, and 800°/s) and the hOFRs elicited by horizontal translations (5, 10, 20, 40, 80, and 160°/s). Note that the latter specified the speed of the pattern directly ahead of each eye, and retinal image speeds decreased roughly as a function the tangents of the angles of eccentricity with respect to these points in the display. On trials involving roll-axis rotations, the random dots occupied a central square area (32° on a side) whose boundaries remained fixed in location, i.e., did not move with the random dots.

Eye-movement recording

The torsional, horizontal, and vertical positions of the right eye were recorded with an electromagnetic induction technique (Robinson, 1963) using scleral search coils embedded in a silastin annulus (Skalar, Delft), as described by Collewijn, van der Mark, and Jansen (1975) and Collewijn et al. (1985). The annulus contained one coil in the frontal plane for recording horizontal and vertical eye position and a second (“folded figure eight”) coil in the sagittal plane for recording torsional eye position (Combination Coil, Skalar Co., Utrecht, The Netherlands). For calibration of the torsional coil, the annulus was mounted on a custom protractor that allowed it to be placed approximately at the location it would occupy during the experimental recordings and with approximately the same orientation. The recorded voltage output of the phase detector (CNC Engineering, Seattle) was a linear function of the torsional angle over the 20° range examined (typically, \( r^2 = 0.9996 \), and horizontal or vertical displacement by up to ±5° had only a minor impact on the torsional signal.

Procedures

All aspects of the experimental paradigms were controlled by three PCs, which communicated via Ethernet using the TCP/IP protocol. One of the PCs was running a Real-time EXPERimentation software package (REX) developed by Hays, Richmond, and Optican (1982) and provided the overall control of the experimental protocol as well as acquiring, displaying, and storing the eye-movement data. Two other PCs were running Matlab subroutines, utilizing the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997), and generated the visual stimuli upon receiving a start signal from the REX machine.

At the beginning of each recording session, the subject fixated targets located along the horizontal and vertical meridians to permit later calibration of the horizontal and vertical eye position signals. Following this initial calibration procedure, a random-dot pattern appeared at the beginning of each trial, together with a central target cross (2° high × 10° wide × 0.21° thick) that was either black or white on alternate trials and whose center the subject was instructed to fixate. After the subject’s right eye had been positioned within 2° of the center of the fixation cross and no saccades had been detected (using an eye velocity threshold of 18°/s) for a randomized period of 800–1100 ms, the fixation target disappeared and the random-dot pattern began to rotate CW or CCW around the screen center. This roll-axis rotation lasted for 200 ms, at which point the screen became a uniform gray (luminance, 20.8 cd/m²) marking the end of the trial. After an inter-trial interval of 500 ms, a new dot pattern appeared together with a central fixation target, commencing a new trial. The subjects were asked to refrain from blinking or shifting fixation except during the inter-trial intervals but were given no instructions relating to the roll stimuli. If no saccades were detected for the duration of the trial, then the data were stored on a hard disk; otherwise, the trial was aborted and subsequently repeated. Data were collected over several sessions until each condition had been repeated an adequate number of times to permit good resolution of the responses (through averaging); the actual numbers of trials will be given in the figure legends. In Experiment 1A, each block of trials had 18 randomly interleaved conditions (9 speeds and 2 directions of roll-axis rotation), and in Experiment 1B, each block of trials had 6 randomly interleaved conditions (1 “whole pattern,” 1 “upper-half pattern,” 1 “lower-half pattern,” and 2 directions of roll-axis rotation). In Experiment 1C, which compared tOFRs and hOFRs, each block of trials had 24 randomly interleaved conditions (6 roll-axis speeds, 6 horizontal speeds, and 2 directions).

Data analysis

The calibration data obtained with the torsional coil prior to the experiment were used to convert the voltage signals recorded during the experiment proper to torsional eye position. The horizontal and vertical eye-position data obtained during the calibration procedure at the start of each recording session were each fitted with second-order polynomials that were then used to linearize the horizontal and vertical eye position data recorded during the experiments proper. Trials with saccadic intrusions (that had
failed to reach the eye-velocity threshold of 18°/s used during the experiment) were deleted. All eye-position data were first smoothed with an acausal 6th-order Butterworth filter (3 dB at 30 Hz) and then mean temporal profiles time-locked to stimulus onset were computed from the data obtained for each of the stimulus conditions for each subject. Mean eye-velocity responses about yaw, pitch and roll axes were estimated at successive 1-ms intervals by computing the differences between the corresponding horizontal, vertical and torsional eye positions over 10-ms intervals. By convention, eye movements in the rightward, upward, and clockwise directions were positive. Because the tracking responses to roll-axis rotation could be very weak, the mean torsional position/velocity signals recorded with each CW motion stimulus were subtracted from the mean torsional eye position/velocity signals recorded with the CCW motion stimulus that had the same angular speed, yielding the “mean CW-CCW torsional eye position” and “mean CW-CCW torsional eye velocity” data. The initial torsional responses were quantified by measuring the changes in the mean CW-CCW torsional eye position signals over the 80-ms time periods commencing 70 ms after the onset of the roll-axis rotation (often referred to simply as, “response measures”). As the minimum latency of onset was ~83 ms, these response measures were restricted to the period prior to the closure of the visual feedback loop (i.e., twice the reaction time): initial open-loop responses. These response measures used a time-window defined with respect to the onset of the stimulus (“stimulus-locked measures”), in line with most previous studies of the initial eye movements elicited by visual stimuli at short-latency, and these will constitute the primary data reported for Experiment 1 (default measures). Although the dependent variable in all of our experiments is a motor response, our major interest is in the sensory processing that underlies this response and, for this, stimulus-locked measures have proved very useful. Such measures have the advantage that, unlike response-locked measures, they are not criterion dependent and allow the characterization of even the weakest responses elicited by suboptimal stimuli (useful for defining the full range over which a parameter exerts its influence). Of course, stimulus-locked response measures confound changes in latency and initial eye acceleration, which in practice are often negatively correlated. Although less robust, we will also report “response-locked measures,” based on the change in the mean CW-CCW torsional eye position signals over the 60-ms time periods commencing when mean CW-CCW torsional eye velocity first exceeded 0.15°/s, a rather conservative criterion for response onset that avoided spurious triggering by noise. It will be seen that stimulus- and response-locked measures generally show very similar dependencies on the stimulus parameters, which can generally be well characterized by simple functions with very few free variables (often only one or two). Importantly, it will also be seen that the dependence of stimulus-locked measures on these stimulus parameters is generally better described by these simple functions than is the dependence of the response-locked measures, i.e., the latter generally show more scatter, presumably in part because of the uncertainty in the estimated time of response onset. The response-locked measures will be reported in summary form only and always after the stimulus-locked measures. Unless otherwise indicated, all error bars are 1 standard deviation of the mean (SD), and the p-value for significance in all statistical tests was 0.05.

Results

Experiment 1A: Speed dependence

Figure 1 shows the mean torsional eye velocity profiles over time obtained from subject FAM in response to CW (A) and CCW (B) rotations with angular speeds ranging from 6.25 to 1600°/s in octave increments. It is evident that initial responses were usually in the direction of the seen rotation and minimum latencies were generally <85 ms. With CW rotations, the torsional eye velocity of this particular subject continued to increase throughout the recording period shown (up to 170 ms after stimulus onset) but with CCW rotations each profile tended to show saturation. The torsional responses to CW and CCW rotations showed a similar dependence on the angular speed insofar as the eye speed achieved within the time window shown was maximal when the angular speed was 200°/s and decreased progressively as the angular speed deviated from this level with both stimuli: note the traces in Figure 1 are shown in continuous line for stimuli speeds ≤200°/s and in dashed line for faster stimuli. Responses were always small and torsional eye speeds never reached 2°/s in the initial open-loop period.

The amplitude of the torsional responses was modest in all three subjects and generally showed only minor directional asymmetries. To improve the signal-to-noise ratio, we pooled the data obtained with CW and CCW stimuli by subtracting the responses to a CCW stimulus of a given speed from the responses to the CW stimulus of the same speed; see the traces in Figure 1C in the column labeled, “CW-CCW.” The mean CW-CCW torsional response measures (stimulus-locked) for these data are plotted in Figure 2A as a function of the angular speed (in °/s) on a logarithmic abscissa; see the open squares. Figure 2A also includes the data from the other two subjects, which are clearly very similar in form though smaller in amplitude: see the open circles (BMS) and diamonds (JKM). The data points for each subject have a negatively skewed Gaussian distribution when plotted on a log abscissa, showing a monotonic (roughly linear) rise to a peak (as speeds approach 100–150°/s) and a slightly steeper decline back towards zero (as speeds approach 1600°/s). For all three data sets, the dependence on log angular speed was well fitted (mean $r^2 = 0.988 \pm 0.014$) by

\[
\log v = \frac{-\pi}{4} \times \left\{ 1 + \frac{1}{\sqrt{1 + \left( \frac{\log \omega}{\log 150} \right)^2} \right\}
\]

where $v$ is the torsional eye velocity and $\omega$ is the log angular speed.
the following Expression consisting of a Gaussian function and its integral, the latter determining the skewness:

\[ A^* \exp \left( -\frac{(X-X_0)^2}{2*\sigma^2} \right) \left( 1 + \text{erf} \left( \alpha^* \frac{X-X_0}{\sqrt{2*\sigma}} \right) \right), \]  

(1)

where \( X_0 \) is the angular speed (in °/s) at which the Gaussian function has its maximum value, \( A \) (in degrees), \( \sigma \) is the Standard Deviation of the Gaussian function (in log units), and \( \alpha \) is a variable that determines the skewness. The continuous curves in Figure 2A are the least-squares best fits obtained with Expression 1, whose parameters are listed in Table 1A in the Supplementary material, together with \( E_{1\text{max}} \), the computed maximum value (in degrees) of Expression 1, and \( E_{1\text{speed}} \), the angular speed (in °/s) at which this value was achieved. The latter ranged from 136 to 196°/s with a mean of 173°/s.

Because of the low amplitude of many of the responses, estimates of the latency were attempted only with the mean CW-CCW torsional eye velocity profiles and the response onset was defined as the time when these profiles first exceeded 0.15°/s. This revealed that larger responses tended to have shorter latencies so that the dependence of latency on angular speed seen in Figure 2B is roughly the inverse of that seen for amplitude in Figure 2A. Expression 1 was fitted to the latency data with an additional term, \( \text{Offset} \), to allow nonzero asymptotes (mean \( r^2 = 0.979 \pm 0.022 \)) and the best-fit parameters are listed in Table 1B in the Supplementary material. The minimum latencies given by the troughs of these fits (\( E_{1\text{lat}} \)) and the angular speeds at which these occurred (\( E_{1\text{speed}} \)) are also listed. When so measured, minimum latencies were 84.9, 83.5, and 100.7 ms for the three subjects BSM, FAM, and JKM, respectively. Latencies could be 30–50 ms longer with less optimal stimuli, presumably in part because larger tOFRs had higher rates of acceleration and the criterion for response onset was a fixed velocity threshold.

The dependence of the tOFRs on angular speed based on response-locked measures was also well fit by Expression 1 (mean \( r^2 = 0.985 \pm 0.010 \)), and the values of the free parameters are listed in parentheses in Table 1A in the Supplementary material. Note that these fits were never quite as good as with the stimulus-locked measures, but the best-fit parameters were generally very similar for the two sets of measurements, i.e., the changes in latency had a relatively minor impact.
**Experiment 1B: A spatial anisotropy**

The torsional tracking responses to roll-axis visual rotations were much more sensitive to exclusion of the lower half of the pattern than to exclusion of the upper half, the average attenuation of the CW-CCW torsional response measures in the two cases being 41 ± 8% and 18 ± 3%, respectively, based on stimulus-locked measures, and 29 ± 9% and 7 ± 2%, respectively, based on response-locked measures.
measures. Again latency tended to vary inversely with response amplitude but the changes here were always small, on average increasing by 5 ± 1 ms with exclusion of the upper half and by 7 ± 2 ms with exclusion of the lower half.

Experiment 1C: Comparison of the tOFRs to roll-axis rotation with the hOFRs to horizontal translations

In a separate experiment, we interleaved trials with roll-axis rotations and horizontal translations to permit a direct comparison of the associated tOFRs and hOFRs, respectively, over a range of speeds distributed around the optimal (see Methods). In both cases, latency of onset (based on the usual 0.15°/s threshold criterion) showed a dependence on speed like that seen in Figure 2B and was generally well fitted by Expression 1 (after adding an offset term to allow nonzero asymptotes), with mean \(r^2\) values of 0.976 ± 0.013 for the tOFR data and 0.987 ± 0.018 for the hOFR data: see Table 2A in the Supplementary material, which lists the best-fit parameters and the minimum latencies (in ms) based on the troughs of these fits, \(E_{1}\),. On average, these estimates of the minimum latency were 17.7 ms shorter for the hOFRs than those for the tOFRs.

The dependence of response amplitude on speed, based on the mean CW-CCW torsional response measures for the tOFRs and on the equivalent mean rightward-minus-leftward measures for the hOFRs (all stimulus-locked), was similar to that in Figure 2A and always well fitted by Expression 1 with mean \(r^2\) values of 0.933 ± 0.115 for the tOFRs and 0.991 ± 0.005 for the hOFRs: the best-fit parameters are listed in Table 2B in the Supplementary material, together with \(E_{1}\)max, the computed maximum values of Expression 1 (means: 0.078° for tOFR and 0.178° for hOFR), and \(E_{1}\)speed, the (optimal) angular speeds at which these values were achieved (means: 188°/s for tOFR and 318°/s for hOFR). Thus, on average, based on stimulus-locked measures, maximum responses were 56% smaller and optimal speeds were 157°/s greater for the tOFRs than for the hOFRs. However, it was possible that the differences in amplitude here were all secondary to the differences in latency noted above. Response-based measures (values in parentheses in Table 2B in the Supplementary material) indicated that this was not the case: on average, maximum responses were 35% smaller (and optimal speeds were 111°/s greater) for the tOFRs than for the hOFRs.

Discussion of Experiment 1

The tOFR operates as a negative feedback tracking system and is assumed to provide a visual backup to the rotational vestibulo-ocular reflex (RVOR) during roll-axis rotations of the head. Of course, in everyday life such rotations usually occur about an axis located well below the eyes so that the latter undergo considerable translation. However, this translation does not disturb the retinal images of distant objects,4 which therefore undergo mostly pure rotation during normal roll rotations of the head. Thus, when viewing a distant visual scene, the potential exists for retinal images to undergo relatively pure rotations around the lines of sight, especially given that the torsional RVOR elicited by roll-axis head rotations has a gain substantially less than one (Migliaccio, Della Santina, Carey, Minor, & Zee, 2006). With near viewing, the situation during roll-axis head rotations is very complicated: again, there is only partial vestibular compensation, both for the rotation of the head (by the RVOR) and for the translation of the eyes (by the Translational Vestibulo-Ocular Reflex, TVOR) so that the associated retinal image motion will be expected to have both rotational and translational components, the latter depending on the distance of the object(s) from the plane of gaze stabilization (Schwarz & Miles, 1991). However, the eye rotations generated by the RVOR are around the roll axis and so result in vertical vergence, generally referred to as “skew deviation” (Bergamin & Straumann, 2001; Jauregui-Renaud, Faldon, Clarke, Bronstein, & Gresty, 1996, 1998; Jauregui-Renaud, Faldon, Gresty, & Bronstein, 2001; Migliaccio et al., 2006; Pansell, Schworm, & Ygge, 2003) and this is later partially corrected by “torsional” quick-phases about the roll axis (Jauregui-Renaud et al., 2001; Migliaccio et al., 2006). Thus, the residual visual motion—and likely involvement of the tOFR—with near viewing is not clear.

A direct comparison indicated that the tOFRs elicited by roll-axis rotations and the hOFRs elicited by horizontal translations were qualitatively very similar, their dependencies on log speed, for example, being well represented by negatively skewed Gaussian functions. However, there were quantitative differences and, compared to the hOFRs, the tOFRs had longer minimum latencies (on average, by 17.7 ms), smaller maximum amplitudes (on average, by 35%, based on response-locked measures), and peaked at higher stimulus speeds (on average, by 111°/s). Given that for both tOFRs and hOFRs larger responses generally had shorter latencies (perhaps in part because of our use of a fixed-velocity criterion for response onset), some of the differences in their minimum latencies might have been secondary to the differences in their response amplitude. However, even if one compares the minimum latencies of the tOFRs with the latencies of the hOFRs that had similar amplitudes the latter were still shorter, on average, by 13.3 ms. The differences in the maximal response amplitudes of the initial tOFRs and hOFRs are appreciably smaller than the differences in the maximum steady-state gains of the tOKN and hOKN: the latter is generally five to ten times the former (e.g., Cheung & Howard, 1991; Collewijn et al., 1985; Farooq et al., 2004).

The quantitative characterization of the magnitude of the tOFRs in the Results section was largely based on stimulus-locked measures but the response-locked
measures were qualitatively very similar in all essentials and differed only modestly in some quantitative details. Intriguingly, the skewed Gaussian functions used to describe the response dependency on log angular stimulus speed generally provided a slightly better representation when stimulus-locked measures were used.

Masking off the lower half of the motion display attenuated the tOFR much more than masking off the upper half, indicating that the visual inputs from the lower visual field are much more potent than those from the upper visual field. During normal viewing, the lower visual field must often be occupied largely by the ground plane, a rich source of feedback for a visual stabilization mechanism like the tOFR.

**Experiment 2: 1-dimensional radial gratings**

We now report the tOFRs elicited by visual rotation about the line of sight when the stimulus consisted of a 1-D radial grating with a sinusoidal luminance profile (Figure 3A). Major concerns were the dependence on angular spatial wavelength and contrast.

**Methods**

Many of the methods and procedures were identical to those used in Experiment 1, and only those that were different will be described here.

**Visual display and the grating stimuli**

Visual stimuli were presented on a single computer monitor (Silicon Graphics CPD G520K 19" CRT driven by a PC Radeon 9800 Pro video card) that was located straight ahead of the subject at 45.7 cm from the corneal vertex. The monitor screen was 385 mm wide and 241 mm high, with a resolution of 1920 × 1200 pixels and a vertical refresh rate of 100 Hz. The visual displays had a resolution of 40 pixels/° at the point directly ahead of each eye. Initially, a luminance look-up table with 64 equally spaced luminance levels ranging from 0 to 77.4 cd/m² was created from direct luminance measurements (IL1700 photometer; International Light Inc., Newburyport, MA) under software control. This table was then expanded to 2048 equally spaced levels by interpolation and subsequently checked for linearity (typically, r > 0.99997).

The visual stimuli consisted of 1-D radial gratings in which luminance modulated sinusoidally with angle: an example is shown in Figure 3A. The gratings occupied a circular area with a diameter of 29.3°, centered on the monitor screen, and had a mean luminance of 38.72 cd/m². The regions of the screen beyond the grating were luminance-matched gray. Roll-axis rotation was produced by substituting a new grating image every frame (i.e., every 10 ms) over a period of 200 ms (i.e., 20 images), each new image being identical to the preceding one except phase shifted CW or CCW by ¼ of the angular wavelength. In any given trial, the successive steps were all in the same direction (CW or CCW). The initial phase of a given grating stimulus was randomized from trial to trial at ¼-wavelength intervals. The central fixation target consisted of a white circular area (diameter, 0.25°) with a black dot (2 × 2 pixels) at the center.

In a first experiment, the independent variable was angular spatial wavelength, randomly sampled each trial from a lookup table (listed values, all a simple fraction of 360° to avoid discontinuities, were 2.8125°, 5.625°, 11.25°, 22.5°, 45°, 90°, and 180°), while the Michelson contrast was fixed at 32%. In order to avoid spatial aliasing problems near the centers of the radial gratings, the central region was masked off (luminance-matched gray). This central mask always had a diameter of 2.9°, which was sufficient to exclude areas where the local spatial frequency would have exceeded the Nyquist limit when using the highest angular spatial frequency (computed from the separation of the adjacent diagonal pixels, i.e., the worst case).

In a second experiment, the independent variable was the Michelson contrast, randomly sampled each trial from a lookup table (listed values: 1%, 2%, 4%, 8%, 16%, 32%, 64%, and 80%), while the angular wavelength was fixed throughout at 24° (BMS and FAM) or 15° (JKM) and the central mask had a diameter of 0.33° (BMS and FAM) or 0.53° (JKM).

**Results**

The initial tOFRs elicited when ¼-wavelength phase shifts were applied to the 1-D radial sine-wave gratings were always in the direction of those shifts, as though mediated by a sensing mechanism that gives greatest weight to the nearest-neighbor matches: see the traces in Figure 3B, which each show the mean CW-CCW torsional eye velocity responses over time elicited by a particular angular spatial wavelength (indicated in degrees by the numbers at the ends of the traces) for subject FAM.

**Dependence on angular spatial wavelength**

The initial tOFR showed a band-pass dependence on angular spatial wavelength that is evident in Figure 3C, which shows the mean CW-CCW torsional response measures (stimulus-locked) obtained from each of the three subjects plotted against the angular spatial wavelength on a log abscissa. These plots were well fit by Gaussian functions (“log-normal” distributions) with r² values ranging from 0.972 to 0.983: see the smooth curves
The two free parameters for these best-fit Gaussian functions—angular wavelength at the peak ($\lambda_o$) and standard deviation ($\sigma$)—are listed in Table 3A in the Supplementary material, together with the peak amplitude ($A_{\text{peak}}$) and the angular wavelengths at which the tuning curve was half its maximum (low-wavelength cutoff, $\lambda_{lo}$, and high-wavelength cutoff, $\lambda_{hi}$), derived as in Read and Cumming (2003). Although the absolute amplitude of the initial tOFR to a given stimulus differed substantially from one subject to another, the $\lambda_o$ and $\sigma$ parameters of their best-fit Gaussian functions were generally similar: $\lambda_o$ ranged from 15.5° to 26.0° (mean, 21.9°), and $\sigma$ ranged from 0.55 to 0.57 log units (mean, 0.56 log units).

Using the same criterion for response onset as was used in Experiment 1, measured onset latencies once again tended to be shorter for the larger responses and were well...
fit by Expression 1 with an additional term (Offset) to allow nonzero offsets: see Figure 3D, in which the latency data for all three subjects are each plotted as a function of angular wavelength (on a log abscissa) together with their associated best-fit skewed Gaussian functions (smooth curves). The best-fit parameters are listed in Table 3B in the Supplementary material, together with $E_{1 \text{lat}}$, the minimum latency given by these fits, which were 79.5, 80.3, and 99.0 ms for the three subjects BSM, FAM, and JKM, respectively, and $E_{1 \text{wave}}$, the wavelength at which the minimal value was achieved.

When these latency data were used to obtain response-locked measures of the tOFRs, dependence on angular wavelength was very similar to that seen with stimulus-locked measures, and the data for each subject were again well fit by a Gaussian function when plotted on a log abscissa (mean $r^2 = 0.947 \pm 0.036$), though these fits were never quite as good as for the stimulus-locked data, cf. Experiment 1. The values of the free parameters for these best-fit Gaussian functions were generally very similar to those for the stimulus-locked measures, though a notable exception was the $\lambda_n$ parameter for subject JKM, which was much higher with the response-locked measures (and more in line with the values of the other two subjects): see the values in parentheses in Table 3A in the Supplementary material. The $\sigma$ parameter tended to be slightly higher with the response-locked measures (mean $\pm SD$, 0.68 $\pm$ 0.06 log units).

dependence on contrast

The initial tOFRs showed gradual saturation as contrast increased and this is evident from the traces in Figure 4A, which each show the mean CW-CCW torsional eye velocity responses over time obtained from subject FAM with a particular contrast (indicated in % by the numbers at the ends of the traces): the traces are clearly close to maximal with a contrast of 16%. The quantitative details are apparent from the mean CW-CCW torsional response measures (stimulus-locked) plotted as a function of contrast for each of the three subjects in Figure 4B (note the logarithmic abscissa). Each plot was fitted with the following expression:

$$K \ast C^{-n} + L_0,$$

(3)

where $K$ is a coefficient, $C$ is contrast, $n$ is an exponent, $L_0$ is the minimum latency, and the plots in Figure 4C include these fits as smooth curves (mean $r^2 = 0.938 \pm 0.058$) whose best-fit parameters are listed in Table 4B in the Supplementary material.

When these latency measures were used to obtain response-locked measures of the tOFRs, dependence on contrast was very similar to that seen with stimulus-locked measures, and the data for each subject were again well fit by Expression 2 (mean $r^2 = 0.974 \pm 0.023$), though once more these fits were never quite as good as for the stimulus-locked data. The best-fit values of the free parameters ($n$, $c_{50}$), which are listed in parentheses in Table 4A in the Supplementary material, showed only minor differences from one subject to another and, compared with the values obtained with the stimulus-locked measures, $n$ and $c_{50}$ values were slightly lower (means, 1.4 $\pm$ 0.9% and 2.3 $\pm$ 0.8%, respectively).

Discussion of Experiment 2

Two features of our rotating radial grating stimulus are that its local spatial wavelength and local image speed (measured along the circular path of the motion) increase with (retinal) eccentricity. Thus, one might expect that the receptive fields of the local motion detectors that would be optimally responsive to such stimuli would increase in size with retinal eccentricity. Of course, there is considerable evidence indicating that visual receptive fields increase in size with retinal eccentricity (e.g., Smith, Singh, Williams, & Greenlee, 2001), so that our 1-D radial
Polar gratings might be expected to activate neurons over much broader regions of the visual field than either 1-D angular polar gratings or 1-D Cartesian gratings whose local spatial wavelength and speed (measured along the path of their motion) are uniform over their full extent. Despite this difference, the tOFRs elicited with 1-D radial gratings showed a clear Gaussian dependence on log wavelength, as previous studies have shown to be the case for the RFVRs elicited with 1-D angular gratings (Kodaka et al., 2007) and the hOFRs elicited with 1-D vertical gratings (Sheliga et al., 2005). However, the optimal angular wavelength for the tOFR (λ_o) averaged ~22°, whereas the optimal linear wavelengths for the RFVR and hOFR both averaged only 3–4°. Note that our radial gratings (angular wavelength, 22°) had a local wavelength (measured circumferentially) of 3–4° at an eccentricity of 7.8–10.4°.

Figure 4. The initial tOFRs obtained with 1-D radial sine-wave gratings: dependence on contrast. (A) Sample traces from subject FAM showing the mean CW-CCW torsional eye velocity over time; each trace is the mean response to 301–320 repetitions of the CW and the CCW stimuli; note that time on the abscissa starts 30 ms after the occurrence of the first step; contrast (in %) indicated by the numbers at the ends of the traces; upward deflections denote net eye rotations in the direction of the ¼-wavelength steps (forward direction). (B) Mean CW-CCW (stimulus-locked) response measures of the 3 subjects plotted as a function of contrast (%) on a logarithmic scale; continuous curves are the best fits obtained with Expression 2, which is based on the Naka–Rushton equation. (C) Mean latency measures of the 3 subjects plotted as a function of contrast (%) on a logarithmic scale; continuous curves are the best fits obtained with Expression 3. Circles, subject BMS (105–116 measures per condition in B and 106–113 measures per condition in C). Squares, subject FAM (301–320 measures in B and C). Diamonds, subject JKM (331–356 measures in B and C).
The mathematical function used to describe the dependence of the initial tOFR on contrast—Expression 2, based on the Naka–Rushton equation—was the same as that used with equal success to describe the contrast dependence of the initial hOFR (Sheliga et al., 2005) and the initial RFVR (Kodaka et al., 2007), and the best-fit parameters even had very similar values: \( n \) averaged 2.1 (tOFR), 1.8 (hOFR), and 1.6 (RFVR), while \( c_{50} \) averaged 4.7% (tOFR), 4.5% (hOFR), and 2.7% (RFVR). This is consistent with the notion that all three motion-based responses are mediated by magnocellular pathways (for review see Merigan & Maunsell, 1993).

As in Experiment 1, larger responses tended to have shorter latencies (cf. Albrecht, Geisler, Frazor, & Crane, 2002; Gawne, Kjaer, & Richmond, 1996; Reich, Mechler, & Victor, 2001; Sestokas & Lehmkuhle, 1986), but this had only a minor quantitative impact on the dependence of tOFR on the variables of interest—angular wavelength and contrast—which were very similar with stimulus- and response-locked measures. In addition, the simple functions used to describe these dependencies—a log-Gaussian and the Naka–Rushton equation—once more provided a slightly better fit when the stimulus-locked measures were used because the latter showed less scatter.

**Experiment 3: Two overlapping 1-dimensional radial gratings (“3f5f stimulus”)**

Recent studies indicate that the very earliest hOFRs are mediated by motion detectors that are sensitive to 1st-order motion energy, as in the well-known energy model of motion analysis (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Watson & Ahumada, 1985). Thus, hOFRs show clear reversal with “1st-order reverse-phi motion,” one of the hallmarks of an energy-based mechanism (Masson, Yang, & Miles, 2002), and are very sensitive to the Fourier composition of the luminance modulations in the motion stimulus (Sheliga et al., 2005).

One of the visual stimuli in this last study consisted of 1-D square-wave gratings lacking the fundamental—referred to as the missing fundamental (mf) stimulus—and motion was applied in discrete \( \frac{1}{4} \)-wavelength steps. The hOFRs associated with this apparent motion stimulus were always reversed, e.g., rightward steps resulted in leftward OFRs. The explanation advanced for this reversal was that the underlying motion detectors do not sense the motion of the raw images (or their features) but rather a spatially filtered version of the images, hence the strong dependence of the hOFR on the Fourier composition of the spatial stimulus. In the frequency domain, a pure square wave is composed entirely of the odd harmonics (1st, 3rd, 5th, 7th, etc.,) with progressively decreasing amplitudes such that the amplitude of the \( n \)th harmonic is proportional to \( 1/n \). Accordingly, the mf stimulus lacks the 1st harmonic and so is composed entirely of the higher odd harmonics, with the 3rd having the lowest spatial frequency and the largest amplitude. This means that when the mf stimulus shifts \( \frac{1}{4} \) of its (fundamental) wavelength, the largest Fourier component, the 3rd harmonic, shifts \( \frac{1}{4} \) of its wavelength in the same (forward) direction. However, a \( \frac{1}{4} \)-wavelength forward shift of a sine wave is exactly equivalent to a \( \frac{1}{4} \)-wavelength backward shift and, because the brain gives greatest weight to the nearest image matches (spatial aliasing), the hOFRs are in the backward direction. In fact, when \( \frac{1}{4} \)-wavelength steps are applied to the mf stimulus, all of the \( 4n−1 \) harmonics (where \( n \) is an integer), such as the 3rd, 7th, 11th, etc., will shift \( \frac{1}{4} \) of their wavelength in the backward direction whereas all of the \( 4n+1 \) harmonics, such as the 5th, 9th, 13th, etc., will shift \( \frac{1}{4} \) of their wavelength in the forward direction. The magnitude and contrast dependence of the initial hOFRs elicited by the mf stimulus generally approximated those of the hOFRs elicited when the same steps were applied to a pure sine wave whose spatial frequency and contrast matched those of the 3rd harmonic, consistent with the idea that the observed responses depended mostly on this single most prominent harmonic (Sheliga et al., 2005). Selectively altering the contrast of that 3rd harmonic of the mf stimulus indicated that its dominance resulted in part from nonlinear interactions between the neural mechanisms responding to the different harmonics (Sheliga, Kodaka, et al., 2006).

This last study included experiments in which the mf stimulus was reduced to just two competing harmonics, the 3rd and 5th, which shifted in opposite directions (termed “the 3f5f stimulus”) and showed that the initial hOFRs were solely determined by the relative contrast of those two harmonics: the motion of the overall features was irrelevant, consistent with mediation by spatio-temporal filters sensitive to 1st-order motion energy. Further, the dependence on relative contrast was highly nonlinear. Thus, the two harmonics were arranged to be of roughly equal efficacy when of equal contrast and presented singly, and when presented together with similar contrast both were effective (vector sum/averaging), but when the contrast of one was less than about \( \frac{1}{2} \) that of the other then the one with the higher contrast became dominant and the one with the lower contrast became ineffective: Winner-Take-All (WTA). This nonlinear interaction was attributed to mutual inhibition between the neuronal mechanisms mediating the responses to each of the two competing gratings. In Experiment 3, we undertook equivalent studies on the tOFR, using 1-D radial grating patterns whose angular luminance modulation was that of two superimposed sine waves with angular spatial frequencies in the ratio 3:5. This “3f5f stimulus” was subjected to apparent rotation consisting of successive \( \frac{1}{4} \)-wavelength steps so that its two component sine waves each underwent \( \frac{1}{4} \)-wavelength steps but in opposite directions. The major concern was
the dependence of the tOFR on the relative contrast of the two competing sine waves.

Methods

Many of the methods and procedures were identical to those used in Experiment 2, and only those that were different will be described here.

Visual stimuli

The visual images consisted of 1-D radial grating patterns that occupied a central circular area (diameter, 29.3°) on the screen facing the subject and could have one of three angular luminance profiles in any given trial: (1) a sum of two sine waves with angular spatial frequencies in the ratio, 3:5, creating a beat of spatial frequency, f (termed the “3f/5f stimulus”); (2) a pure sine wave with the same angular spatial frequency as the 3f component of the 3f/5f stimulus (the “3f stimulus”); and (3) a pure sine wave with the same angular spatial frequency as the 5f component of the 3f/5f stimulus (the “5f stimulus”). The successive angular phase shifts used to generate the roll-axis rotation were always of the same absolute amplitude, which was 1/4 of the fundamental wavelength of the 3f/5f stimulus, so that the 5f component (and the comparable 5f stimulus) underwent 1/4-wavelength forward steps whereas the 3f component (and the comparable 3f stimulus) underwent 1/4-wavelength backward steps. The angular spatial frequencies of the 3f and 5f stimuli/components were carefully selected so as to be of similar efficacy when of equal contrast, i.e., they elicited tOFRs of similar amplitude. Thus, for each subject, we selected the two sine waves that had (1) angular spatial frequencies in the ratio, 3:5; (2) roughly symmetrical locations on either side of the peak of the best-fit Gaussian describing the dependence of the tOFRs on (log) angular spatial frequency; and (3) angular wavelengths that were a simple fraction of 360° (to avoid discontinuities). The selected 3f and 5f sine waves had angular wavelengths of 30° and 18°, respectively, for subjects BMS and FAM, 20° and 12°, respectively, for subject JKM. The 3f and 5f components of the 3f/5f stimuli could have one of 13 Contrast Ratios randomly selected from a lookup table: 0.25, 0.33, 0.5, 0.5946, 0.7071, 0.8409, 1.0, 1.1892, 1.4142, 1.6818, 2.0, 3.0, and 4.0. The total contrast of these 3f/5f stimuli was fixed at 64% so that increases in the contrast of one component were balanced by decreases in the contrast of the other component. The entries in the lookup table, indicating the Michelson contrast (in %) of the 3f and 5f component pairs, respectively, were 13.3 and 53.1; 16.7 and 50.2; 22.6 and 45.1; 25.4 and 42.7; 28.3 and 40.1; 31.4 and 37.3; 34.5 and 34.5; 37.5 and 31.6; 40.6 and 28.7; 43.4 and 25.8; 46.2 and 23.1; 51.6 and 17.2; and 54.7 and 13.7. The contrasts of the pure 3f and 5f stimuli matched those of the corresponding components of the 3f/5f stimuli.

Procedures

Each block of trials had 78 randomly interleaved conditions: 13 Contrast Ratios, 3 stimulus types (3f/5f, 3f, 5f), and 2 directions of roll-axis rotation (CW, CCW).

Results

Over the range of contrasts used in this experiment (13–53%), the initial tOFRs elicited by the pure 5f and the pure 3f stimuli were mostly saturated and so showed little sensitivity to changes in the contrast level: see the sample mean CW-CCW torsional eye velocity profiles over time obtained from subject FAM shown in Figures 5A and 5B. The initial tOFRs elicited by the dual 3f/5f stimuli depended critically on the relative contrast of the 3f and 5f components. This is apparent from the mean CW-CCW torsional velocity profiles in Figure 5C, in which the Contrast Ratios, 3f/5f, are shown to the right of the traces. Thus, when the Contrast Ratio favored the 5f component by an octave or more, the response was strongly positive (denoting the forward direction) and very similar to that elicited by the 5f stimulus alone; when the Contrast Ratio favored the 3f component by a similar amount, the response was strongly negative (denoting the backward direction) and very similar to that elicited by the 3f stimulus alone. When the component contrasts were more similar, responses were intermediate, but showed a positive bias in favor of the 5f component when the two components were exactly equal in contrast.

The mean CW-CCW torsional response measures (stimulus-locked) based on the data of subject FAM in Figure 5 are plotted as a function of contrast in Figure 6A. The response measures for the pure 3f and 5f sine-wave data in Figure 6A (orange circles, green circles) are clearly saturated and show little sensitivity to changes in contrast over the range examined. The response measures for the 3f/5f data are plotted twice in Figure 6A (left-right mirror images): first, as a function of the contrast of the 5f component (black lines and open squares), to show how they merge with the data obtained with the pure 5f stimulus when the 5f component had high contrast; second, as a function of the contrast of the 3f component (gray lines and squares), to show how they merge with the data obtained with the pure 3f stimulus when the 3f component had high contrast. Thus, the data obtained with the 3f/5f stimuli show a sigmoidal dependence on contrast and deviate substantially from a simple linear prediction based on the vector sum of the responses to pure 3f and 5f stimuli of matching contrasts: in Figure 6A see the data labeled “Vector Sum,” which are plotted with respect to...
the contrast of the 5f component (black dots) and the contrast of the 3f component (gray dots).

To quantify the transition from dominance by one component to dominance by the other component more clearly, we computed the Response Ratio of Sheliga, Kodaka, et al. (2006) using the following expression:

$$R_{3f/5f} = \frac{R_{3f} - R_{5f}}{R_{5f} - R_{3f}},$$  \hspace{1cm} (4)

where $R_{3f/5f}$ is the mean response to the 3f/5f stimulus when the 3f and 5f components have particular contrast values, and $R_{3f}$ and $R_{5f}$ are the mean responses to pure 3f and 5f stimuli with contrasts matching those values. To the extent that the response to a given 3f/5f stimulus is determined exclusively by the 5f component (i.e., $R_{3f/5f} \approx R_{5f}$), the value of the numerator in Expression 4 will approach zero and the Response Ratio will therefore also approach zero. The Response Ratios of subject FAM, based on the 3f/5f response measures in Figure 6A, have been plotted in Figure 6B as a function of the Contrast Ratio, $3f/5f$, on a log abscissa (black filled circles). It is now clear that for Contrast Ratios less than $-0.5$, the 5f component was almost totally dominant and for Contrast Ratios greater than $-2$, the 3f component was almost totally dominant. Thus, when the Contrast Ratio was high or low only one component was effective (WTA) and the transition from one extreme to the other was rather abrupt. The Response Ratio data obtained from the other two subjects with the 3f/5f stimuli showed very similar nonlinear dependencies on the Contrast Ratio: see the red circles (subject BMS) and blue open squares (JKM) in Figure 6B.

To obtain a quantitative estimate of the abruptness of the transitions in Figure 6B, the data were fitted with cumulative Gaussian functions, forced through asymptotes at 0 and 1, using a least squares criterion: see the color-matched smooth curves in Figure 6B, the parameters of which are listed in Table 5 in the Supplementary material.
The $r^2$ values for these fits averaged 0.981 (range, 0.962–0.994), indicating that they provide a very adequate description of these data, and their standard deviations ($\sigma$) averaged 0.13 (range, 0.11–0.15 log units). We also wanted to obtain a quantitative estimate of how different the contrasts of the two components of the 3f/5f stimuli had to be for one of the components to effectively lose its influence. For this we used the cumulative Gaussian functions to identify a Transition Zone, which we defined as the range of Contrast Ratios over which the Response Ratio ranged from 0.05 to 0.95: see the “5%” and “95%” listings in Table 5 in the Supplementary material. On average, this Transition Zone extended from 0.72 to 1.87, indicating that a 1.6-fold difference in contrast generally sufficed for the sine wave with the lower contrast to effectively lose its influence on the tOFR.

At the center of the Transition Zone, where responses could be very weak, onset latencies (determined as in Experiment 1) with the 3f/5f stimuli could be elevated, but outside this region onset latencies showed only very minor changes (<10 ms), presumably in part because the contrast of the dominant component was always high. Consequently, when response-locked measures of the tOFRs were used, contrast dependencies with the pure 3f, pure 5f, and dual 3f/5f stimuli were little different from those with stimulus-locked measures. Thus, the plots of Response Ratio (based on response-locked measures) against Contrast Ratio were again well fit by the cumulative Gaussian function (mean $r^2 = 0.983$), though for 2/3 subjects these fits were not quite as good as for the stimulus-locked data: see the best-fit parameters listed in parentheses in Table 5 in the Supplementary material. The Transition Zone tended to be slightly narrower with the response-locked measures.

Discussion of Experiment 3

When two overlapping 1-D radial grating patterns were subject to roll-axis rotation in opposite directions—the 3f/5f stimulus—the resulting tOFRs showed a nonlinear dependence on the relative contrasts of those two gratings. On average, a difference in contrast of less than an octave was sufficient for the grating of higher contrast to completely dominate the response: WTA. These findings are analogous to those reported for three other oculomotor responses elicited at short latency by visual stimuli: the hOFR (Sheliga, Kodaka, et al., 2006), the DVR (Sheliga, FitzGibbon, & Miles, 2007), and the RFVR (Kodaka et al., 2007).
Taking our lead from those earlier studies, we attributed the WTA outcome in the present study to mutual inhibition between the neurons carrying the competing information. On average, a 1.6-fold difference in contrast was sufficient for the tOFR to show WTA behavior when confronted with the 3/5f stimulus and this is very close to the 1.8-, 2.2-, and 1.9-fold differences in contrast that these previous studies showed were sufficient to render the hOFR, the vDVR, and the RFVR, respectively, unresponsive to the component with the lower contrast. It is also significant that when the 3f component had the higher contrast, the tOFRs were in the opposite direction to the seen rotation of the whole pattern. These observations are all consistent with mediation by low-level motion detectors that (1) respond to a filtered version of the motion stimulus rather than its raw features and (2) are sensitive to the 1st-order motion energy in the stimulus (cf. Bostrom & Warzecha, 2009; Kodaka et al., 2007; Sheliga et al., 2005; Sheliga, FitzGibbon, et al., 2006).

The previous studies on the nonlinear characteristics of the other three oculomotor reflexes used a number of additional quantitative analyses and we have applied these to our present data with a remarkably similar outcome. We will now summarize the findings with these analyses and compare them with the previous ones but will describe the techniques in outline only: readers interested in more complete descriptions of the analytical methodology should consult the previous papers.

The first of these additional analyses addresses the fact that the analysis so far is based on responses averaged over many trials. With such data, the WTA outcome is manifest only when the Contrast Ratio is outside the Transition Zone. However, if the motion signals are very noisy then it is possible that a WTA situation also prevails inside the Transition Zone. For example, a mean Response Ratio of 0.5 might have resulted because torsional eye movements were effectively driven exclusively by the 5f component in half of the trials and exclusively by the 3f component in the other half of the trials. If this were the case, then we would expect the distribution of the tOFRs to a given 3/5f stimulus to be bimodal inside the Transition Zone and unimodal outside. In the previous studies, the response distributions were always unimodal and well fit by Gaussian functions with comparable SDs, even near the center of the Transition Zone when the competing sine waves were of similar contrast, and the response distributions with the pure 3f and 5f stimuli showed only minor overlap. We found the same in the present study: see the sample data in Figure 7A obtained from subject FAM when the 3f grating (contrast, 37.5%) rotated CCW and the 5f grating (contrast, 31.6%) rotated CW. The response distributions here with all three stimuli were well fit by a Gaussian function (smooth curves in Figure 7A), with r² values ranging from 0.88 to 0.97, and the SD was actually smaller with the dual (3/5f) stimuli than with the single (pure 3f, pure 5f) stimuli: see the orange, green, and gray histograms/curves, respectively, in Figure 7A. Very similar data were obtained from the other two subjects and, significantly, the standard deviation of the actual response distributions (rather than of the best-

![Figure 7](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932859/)
fit Gaussians) with the 3/f stimuli for which the Response Ratio was closest to 0.5 (the center of the Transition Zone) were never significantly larger—and in 3/12 cases were significantly smaller—than those of the response distributions with the pure 3f and pure 5f stimuli of matching contrast (Fischer test). Also, in 10/12 cases, the SDs of the 3/f distributions near the center of the Transition Zone were not significantly different from the SDs of the distributions for which the Response Ratios were closest to zero or unity (Fischer test). This all strongly suggests that the WTA situation does not operate inside the Transition Zone, and in order to confirm this we ran a simulation. For this, we first used the mean responses to the three stimuli and Expression 4 to estimate the Response Ratio and then simulated the response distribution predicted by the WTA model for the 3/f stimuli by summing the response distributions obtained with the pure 3f and 5f stimuli, weighted in accordance with this Response Ratio. It was clear from this that the simulated 3/f response distributions were indeed bimodal and extended well beyond the extremes of the actual 3/f response distributions, which were unimodal: see Figure 7B for an example, with the actual distribution in gray and its best-fit Gaussian function in black line (reproduced from Figure 7A) and the simulated distribution in blue. Overall, when the Response Ratios were closest to 0.5, the distributions of the “real” and the “simulated” responses to the 3/f stimuli for the data obtained from two out of three subjects were significantly different (p < 0.01 on the Kolmogorov–Smirnov two-sample test). For the third subject (JKM), the “simulated” distributions also tended to be broader than the “real” ones, but only by 0.2% and 16% for CW and CCW stimuli, respectively, and these differences were not statistically significant.6

These findings indicate that vector sum/averaging prevails near the center of the Transition Zone and WTA prevails outside this Zone, which is in line with the previous findings on the hOFR, DVR, and RFVR. Those previous studies were also able to fully account for the nonlinear dependence on relative contrast using a Contrast-Weighted Average model with just two free parameters. We tried this same approach on our present data by determining how well the 3/f data like those in Figure 6A (describing the dependence of the mean CW-CCW torsional eye position measures on the contrast of the 3f and 5f components) were fitted by the following Contrast-Weighted Average model, which has only two free parameters:

\[ R_{3/f} = \frac{(C_{3f})^{n_{3f}}}{(C_{3f})^{n_{3f}} + (C_{5f})^{n_{5f}}} R_{3f} + \frac{(C_{5f})^{n_{5f}}}{(C_{3f})^{n_{3f}} + (C_{5f})^{n_{5f}}} R_{5f}, \]

where \( R_{3/f} \) is the simulated tOFR to a given 3/f stimulus whose two components have contrasts of \( C_{3f} \) and \( C_{5f} \), respectively; \( R_{3f} \) and \( R_{5f} \) are the mean CW-CCW torsional response measures to pure 3f and 5f stimuli, respectively, with contrasts of \( C_{3f} \) and \( C_{5f} \); \( n_{3f} \) and \( n_{5f} \) are two free parameters that reflect the efficacies of the 3f and 5f components, respectively, of the given 3/f stimulus and thereby determine the abruptness of the transition. The least squares best-fit values of the \( n_{3f} \) and \( n_{5f} \) parameters, together with the \( r^2 \) values, for all of the 3/f data like those in Figure 6A, are listed for all three subjects in Table 6 in the Supplementary material. The \( r^2 \) values averaged 0.984 (range, 0.969–0.995), indicating that Equation 5 always provided a very good and complete description of the data. The exponents, which provide an estimate of the strengths of the postulated mutual inhibition between the neuronal mechanisms mediating the responses to each of the two sine waves, averaged 5.42 (\( n_{3f} \)) and 6.17 (\( n_{5f} \)). These values are slightly higher than those for the other three oculomotor reflex responses, which were, respectively: 5.43 and 5.20 for the hOFR (Sheliga, Kodaka, et al., 2006), 3.40 and 2.99 for the vDVR (Sheliga et al., 2007), and 4.09 and 4.71 for the RFVR (Kodaka et al., 2007). In summary, the Contrast-Weighted Average model, with only two free parameters, provided a very good description of our toOFR data and a quantitative estimate of the strength of the nonlinear interactions. The outcome was essentially the same when the response-locked measures were used for these analyses, though as already pointed out the transitions were much more abrupt with these measures: see the values listed in parentheses in Table 6 in the Supplementary material.

We also attempted to fit data like those in Figure 6A with a Response-Weighted-Average model in which CW-CCW torsional response measures were substituted for the contrast values in Equation 5. With \( r^2 \) values ranging from 0.003 to 0.25, this model never provided a good fit to the data, consistent with the idea that the nonlinear interactions occur at the sensory—rather than the motor—level where the competing motions are encoded. Once more this is in line with the findings in the studies of the other three oculomotor reflex responses and with their suggestions that the postulated mutual inhibition occurred between neurons carrying competing visual information. In the case of the two reflexes that, like the tOFR, relied on visual motion—the hOFR/vOFR and RFVR—the postulated mutual inhibition was likened to “motion opponency,” a competitive interaction for which there is substantial supporting evidence from psychophysical studies, functional magnetic resonance imaging, and single unit recordings in areas V1 and MT (Kodaka et al., 2007; Sheliga et al., 2007, 2008b; Sheliga, Kodaka, et al., 2006). In discussing the functional role of these competitive interactions, these studies also cited earlier suggestions that motion opponency could improve noise immunity, increase directional selectivity, and contribute to pattern selectivity. They also argued that, in favoring images of higher contrast, the interactions responsible for the WTA outcome would tend to favor objects in the
plane of fixation because their retinal images are better focused (due to accommodation) and so tend to have a higher contrast than those of objects in other depth planes. This apparent preference for the plane of fixation is reminiscent of the finding that the tOKN is compromised by horizontal disparity (Washio et al., 2005) but might seem to be redundant in a reflex supposedly concerned with pure rotational disturbances of the observer, which affect all retinal images equally, regardless of whether they are in the plane of fixation or not (Miles, 1998; Miles, Busettini, Masson, & Yang, 2004). However, as pointed out in the Discussion of Experiment 1, the eyes will generally undergo translation during normal roll-axis rotations of the head, disturbing the retinal images of objects in accordance with their distance from the plane of ocular stabilization (e.g., Schwarz & Miles, 1991), hence the need to ignore this reafferece insofar as it emanates from objects outside the plane of fixation. Another possibility is that mutual inhibition is a common feature in the early cortical processing of competing visual motions (Rust, 2004; Rust, Schwartz, Movshon, & Simoncelli, 2005); hence, any mechanism like the tOFR that utilizes visual motion signals will inherit this characteristic, useful or not.

In order for the postulated inhibition generated by the higher contrast component to totally suppress even the earliest tOFRs generated by the lower contrast component, the former must have the shorter latency. There is considerable evidence that higher contrast stimuli elicit activity in striate cortex (V1) at shorter latencies than do low contrast stimuli, and the same is often true of visually elicited eye movements (Gellman, Carl, & Miles, 1990; Kodaka et al., 2007). However, in the present experiment, contrasts were always relatively high and responses close to saturation, resulting in only minor differences in the latency. We examined this issue by comparing the latencies of the tOFRs to the highest contrast 3f stimuli with those to the lowest contrast 5f stimuli (and vice versa), as well as the latencies of the tOFRs to the second-highest contrast 3f stimuli with those to the second-lowest contrast 5f stimuli (and vice versa). These stimulus pairs corresponded to the components of the dual-grating stimuli that showed the most robust WTA responses and revealed only a slight tendency for the tOFRs elicited by the grating with the higher contrast to have the lower latency (on average, by 3 ± 9 ms).

### Methods

Most of the methods and procedures were identical to those used in Experiment 3, and only those that were different will generally be described here.

### Subjects

Two subjects participated: FAM and BMS.

### Visual stimuli

The visual images consisted of 1-D radial grating patterns that occupied a central circular area (diameter, 29.3°) on the screen facing the subject and could have one of three angular luminance profiles in any given trial: (1) a sum of two sine waves with angular spatial frequencies in the ratio, 3:7, creating a beat of spatial frequency, f (termed the “3f7f stimulus”); (2) a pure sine wave with the same angular spatial frequency as the 3f component of the 3f7f stimulus (the “3f stimulus”); and (3) a pure sine wave with the same angular spatial frequency as the 7f component of the 3f7f stimulus (the “7f stimulus”). The successive angular phase shifts used to generate the roll-axis rotation were always of the same absolute amplitude, which was ¼ of the fundamental wavelength of the 3f7f stimulus, so that both the 7f and the 3f components underwent ¼-wavelength backward steps. In Experiment 3, the competing motions were in opposite directions so that changes in their relative efficacy influenced both the magnitude and the direction of the tOFR, but in the present experiment the competing motions were in the same direction so that changes in their relative efficacy could influence only the magnitude of the tOFR. Thus, in order to distinguish between the contributions of the two motions are negatively cross-coupled and that this can result in WTA behavior. Those experiments used 1-D radial gratings composed of two sinusoids corresponding to the 3rd and 5th harmonics of the mf stimulus, which shift in opposite directions when the pattern is moved in ¼-wavelength steps. In the present experiment, we used 1-D radial gratings composed of two sinusoids corresponding to the 3rd and 7th harmonics of the mf stimulus, which shift in the same (backwards) direction—though at different speeds—when the pattern is moved in ¼-wavelength steps. This stimulus was directly analogous to the “3f7f stimuli” used in previous studies on the initial hOFR (Sheliga, Kodaka, et al., 2006) and vDVR (Sheliga et al., 2007), which again uncovered highly nonlinear dependencies on the relative contrast of the two harmonics similar to those seen with the 3f7f stimulus. Thus, in Experiment 4, our major concern was to determine the dependence of the tOFR on the relative contrast of the 3rd and 7th harmonics.
components in the present experiment, it was necessary to arrange for them to differ in efficacy as much as possible. However, it was essential that the less effective component was still able to generate robust responses, hence severely limiting the extent of any changes in magnitude that could occur as the response bias shifted from one component to the other (as it must when the Contrast Ratio was changed from very high to very low, or vice versa). Accordingly, the angular spatial wavelength of the 7f stimulus/component was selected to be as close as possible to the peak of the appropriate Gaussian curve describing the dependence on angular spatial wavelength in Figure 3C (while still a simple fraction of 360° to avoid discontinuities). This meant that, of necessity, the 3f stimulus/component was below the peak of this Gaussian but nonetheless was still very effective. The selected angular wavelengths were 25.7° (7f) and 60° (3f), corresponding to a 3f/7f pattern with a fundamental angular wavelength of 180°. The 3f and 7f components of the 3f/7f stimuli could have one of 13 Contrast Ratios, which were the same as for the 3f and 5f components in Experiment 3, and a total contrast that was always 64%. The entries in the lookup table, indicating the Michelson contrast (in %) of the 3f and 7f component pairs, respectively, of the 3f/7f stimuli were 12.8 and 51.2; 16.0 and 48.0; 21.3 and 42.7; 23.9 and 40.1; 26.5 and 47.5; 29.2 and 34.8; 32.0 and 32.0; 34.8 and 29.2; 37.5 and 26.5; 40.1 and 23.9; 42.7 and 21.3; 48.0 and 16.0; and 51.2 and 12.8. The contrasts of the pure 3f and 7f stimuli matched those of their respective components of the 3f/7f stimulus.

### Procedures

Each block of trials had 78 randomly interleaved conditions: 13 Contrast Ratios, 3 stimulus types (3f/7f, 3f, 7f), and 2 directions of roll-axis rotation (CW, CCW).

### Results

The initial tOFRs to the pure 7f and 3f stimuli showed relatively minor sensitivity to changes in contrast over the range employed and, as expected, the responses to the 7f stimuli were somewhat larger than those to the 3f stimuli: see the sample mean CW-CCW torsional eye velocity profiles over time obtained from subject FAM in Figure 8A (pure 7f stimuli) and Figure 8B (pure 3f stimuli) as well as the associated response measures for these data in Figure 9A (green circles, orange circles). Note that Figures 8 and 9 have the same general layout as Figures 5 and 6. The initial tOFRs elicited by the 3f/7f stimuli, in which the 3f and 7f components rotated always in the same direction, depended critically on the relative contrast of those components. This can be deduced from the mean CW-CCW torsional velocity profiles in Figure 8C, in which the Contrast Ratios, 3f/7f, are shown to the right of the traces. Thus, when the Contrast Ratio favored the 7f component by an octave or more, the response approximated that elicited by the 7f stimulus alone, and when the Contrast Ratio favored the 3f component by a similar margin, the response...
approximated that elicited by the 3f stimulus alone. The associated response measures for these 3f/7f data (stimulus-locked) are plotted twice in Figure 9A (cf. Figure 6A): first, as a function of the contrast of the 7f component (gray line/squares), when they merge with the data obtained with the pure 7f stimuli as the 7f component reaches the higher contrast levels; second, as a function of the contrast of the 3f component (black line/squares), when they merge with the data obtained with the pure 3f stimuli as the 3f component reaches the higher contrast levels. (Note that these data are plotted at very high resolution.; hence, the noisy appearance.) Of course, these 3f/7f data deviate substantially from the simple linear predictions based on the vector sum of the responses to pure 3f and 7f stimuli of matching contrasts, which all lie well below the abscissa and so are not visible in Figure 9A.

Again, we computed the Response Ratio of Sheliga, Kodaka, et al. (2006) using the following expression:

\[ \frac{R_{3f/7f} - R_{3f}}{R_{7f} - R_{3f}} \]

where \( R_{3f/7f} \) is the mean response to the 3f/7f stimulus when the 3f and 7f components have particular contrast values, and \( R_{3f} \) and \( R_{7f} \) are the mean responses to pure 3f and 7f stimuli with contrasts matching those values. The Response Ratios of subject FAM, based on the mean 3f/7f response measures in Figure 9A, are plotted in Figure 9B (black circles) as a function of the Contrast Ratio, 3f/7f, on a logarithmic scale. It is now clear that for Contrast Ratios less than ~0.5, the 7f component was almost totally dominant and for Contrast Ratios greater than ~2, the 3f component was almost totally dominant. Thus, when the Contrast Ratio was high or low only one component was effective (WTA) and the transition from one extreme to the other was fairly abrupt. The Response Ratio data obtained from the only other subject (BMS) showed very similar nonlinear dependence on the Contrast Ratio with similarly abrupt transitions between the zero and unity extremes: see the red circles in Figure 9B. The smooth color-matched curves in Figure 9B are the least-squares best-fit cumulative Gaussian functions, forced through asymptotes at 0 and 1. The \( r^2 \) values for these fits were 0.872 (FAM) and 0.864 (BMS), and their standard deviations (in log units) were 0.23 (FAM) and 0.25 (BMS). We again used the cumulative Gaussian functions to identify a Transition Zone, which we again defined as the range of Contrast Ratios over which the Response Ratio ranged from 0.95 to 0.05. This Transition Zone
extended from 0.52 to 2.88 for subject FAM and 0.42 to 2.74 for subject BMS. Based on the means of the 5% values and the reciprocals of the 95% values, on average, a 2.5-fold difference in contrast generally sufficed for the sine wave with the lower contrast to effectively lose its influence on the tOFR.

When response-locked measures were used, the cumulative Gaussian function fitted the plots of Response Ratio against Contrast Ratio only poorly: \( r^2 \) values were 0.708 (FAM) and 0.635 (BMS), and their standard deviations (in log units) were 0.39 (FAM) and 0.09 (BMS). This Transition Zone extended from 0.24 to 4.46 for subject FAM and 0.79 to 1.57 for subject BMS so that, on average, a 2.9-fold difference in contrast generally sufficed for the sine wave with the lower contrast to effectively lose its influence on the tOFR.

**Discussion of Experiment 4**

The data in Figure 9 indicate that when two superimposed 1-D radial sine waves differing only in spatial frequency and speed rotated in the same direction, the resulting tOFR depended critically on the relative contrasts of those two sine waves. This dependence was highly nonlinear, the tOFRs being essentially bimodal as responses tended to be determined by whichever of the two sine waves had the higher contrast: WTA. Using stimulus-locked measures, the \( SD \) of the best-fit cumulative Gaussians for the 3/7f data were about twice those for the same two subjects in Experiment 3 with the 3/5f stimuli. Presumably a major factor here is that the two components of the 3/5f stimulus differed in direction as well as in spatial frequency and speed, permitting the sensory mechanisms to distinguish between them more readily. Of course, another limiting factor is our ability to distinguish between two responses that differ only in latency. Indeed, when response-locked measures were used, the initial hOFR was maximal with just 3 strips (~10% coverage) and a further five-fold increase in the number of strips (to 15, giving ~50% coverage) produced almost no change, i.e., response amplitude had asymptoted. When the number of strips was then increased to 30 so as to fill all remaining gaps in the stimulus (100% coverage) then hOFRs actually decreased (an effect attributed to the inhibitory surrounds of the underlying motion detectors). When the experiment was repeated using a range of contrasts, the initial hOFR showed essentially the same pattern of dependence on the number of strips at any given contrast but, significantly, the higher the contrast, the higher the level at which the response asymptoted. This indicated that the asymptote in response amplitude was not due simply to the passive achievement of some intrinsic upper limit in the magnitude of the eye movement or the underlying motion signals (“ceiling effect”). Rather, this asymptote was seen as the result of an active process consistent with spatial normalization and was attributed to global divisive inhibition. We now report the results of equivalent studies on the tOFR using rotating 1-D radial sine-wave gratings that could completely fill a large circular area, as in Experiment 2, or occupy a number of equally spaced concentric annuli, each with the same radial thickness, angular wavelength, contrast and phase: see Figure 10A for a sample stimulus with 7 annuli. The major concern here was to determine the extent of any spatial normalization. In additional experiments, we varied the number, thickness, and contrast of the annuli and thereby uncovered evidence similar problem in the Transition Zone (Sheliga, Kodaka, et al., 2006).

**Experiment 5: Annular radial gratings and spatial summation properties**

In a recent study, we recorded the initial hOFRs when successive ¼-wavelength steps were applied to a 1-D vertical sine-wave grating and examined the spatial summation properties by varying only the vertical extent of the grating, i.e., by varying the extent of the stimulus only orthogonal to the axis of motion (Sheliga et al., 2008a). The grating could occupy the full monitor screen (45º wide, 30º high) or a number of horizontal strips, each 1º high and extending the full width of the display. These strips were always equally spaced vertically, and we examined the effect of increasing their number. Surprisingly, even a single (centered) strip (covering only 3.3% of the screen) elicited robust hOFRs, and the main effect of increasing the number of strips was to decrease the latency. Indeed, when response-locked measures were used, the initial hOFR was maximal with just 3 strips (~10% coverage) and a further five-fold increase in the number of strips (to 15, giving ~50% coverage) produced almost no change, i.e., response amplitude had asymptoted. When the number of strips was then increased to 30 so as to fill all remaining gaps in the stimulus (100% coverage) then hOFRs actually decreased (an effect attributed to the inhibitory surrounds of the underlying motion detectors).
suggesting that the tOFR was largely determined by the total motion energy in the stimulus.

**Methods**

The subjects, together with many of the methods and procedures, were identical to those used in Experiment 2, and only those aspects that were different will be described here.

**Grating stimuli**

All visual stimuli were derived from a centered 1-D radial grating with a circular outline (diameter, ~30°) whose luminance modulated sinusoidally with angle. On any given trial, this grating was subdivided into a number of equally spaced concentric annuli, all with the same angular wavelength (always 22.5°), radial thickness, Michelson contrast, phase, and mean luminance (always 38.7 cd/m²).

**Experiment 5A: Dependence on the number and contrast of the annuli**

In a first experiment, all annuli had the same radial thickness (20 pixels measured along the cardinal axes: nominal, 0.5°), and the independent stimulus variables—randomly sampled each trial from a lookup table—were the Michelson contrast (FAM: 3%, 6%, 12%, 32%; BMS: 2%, 4%, 8%, 32%; JKM: 6%, 12%, 32%) and the number of annuli. For the latter, the radial grating was subdivided into 30 equal-thickness concentric annuli, with numbered locations designated 1 to 30, such that Annulus 1 had inner and outer radii of 0° and 0.5°, Annulus 2 had inner and outer radii of 0.5° and 1°, and so forth, so that Annulus 30 had inner and outer radii of 14.5° and 15°. The total number of annuli displayed on any given trial (with their numbered locations given in parentheses) could be 1 (16), 2 (8, 24), 3 (8, 16, 24), 4 (4, 12, 20, 28), 7 (4, 8, 12, 16, etc.), 8 (2, 6, 10, 14, etc.), 15 (2, 4, 6, 8, etc.), or 29 (every annulus except the most central one): the cartoon in Figure 10A shows the stimulus with 7 annuli, and the diagram in Figure 10B indicates the locations of the annuli for all 8 combinations used. Note that the central 1° (occupied by Annulus 1) was always left blank to avoid spatial aliasing. All regions of the screen other than those occupied by the designated annuli were uniform gray (luminance, 38.7 cd/m²). Roll-axis rotation was produced by substituting a new grating image every frame (i.e., every 10 ms) over a period of 200 ms (i.e., 20 images), each new image being identical to the preceding one except rotated CW or CCW about its center by ¼ of the angular wavelength. In any given trial, the successive steps were all in the same direction (CW or CCW). The initial phase of the grating stimulus was randomized from trial to trial at ¼-wavelength intervals. With subjects FAM and BMS, each block of trials had 64 randomly interleaved conditions: 8 combinations of annuli with 4 contrasts and 2 directions of roll-axis rotation. With subject JKM, each block of trials had 48 randomly interleaved conditions: 8 combinations of annuli, 3 contrasts, and 2 directions of roll-axis rotation.

**Experiment 5B: Dependence on the thickness and contrast of the annuli**

In a second experiment, in any given trial there were either 7 or 8 annuli (whose locations were as indicated in Experiment 5A and Figure 10B) and all were of the same
radial thickness. However, this thickness varied from trial to trial and could be 20 pixels (measured along the cardinal axes, i.e., nominal 0.5°, as in Experiment 5A and Figure 10), 40 pixels or 60 pixels. Further, in a given trial, all annuli were of the same contrast but this contrast varied from trial to trial as follows: the 7 annuli could have one of two contrasts (6% or 32% for FAM; 4% or 32% for BMS; 12% or 32% for JKM) and the 8 annuli could have one of two different contrasts (3% or 12% for FAM; 2% or 8% for BMS; 6% or 25% for JKM). Each block of trials had 24 randomly interleaved conditions: 3 thicknesses, 2 contrasts, 2 stimulus combinations (7 or 8 annuli), and 2 directions (CW, CCW).

Data analysis

A major concern in this experiment was to document any evidence for spatial normalization. In our recent study of this phenomenon in the hOFR (Sheliga et al., 2008a), as pointed out in the introduction to this experiment, a five-fold change in the spatial extent of the stimulus affected mainly latency, and when the response measures were defined with respect to the onset of the response (“response-locked measures”), they often showed almost no change, consistent with divisive normalization. For this reason, in the present experiment we first provide response-locked measures, given by the change in the mean CW-CCW torsional eye position signals over the 60-ms time periods commencing when mean CW-CCW torsional eye velocity first exceeded 0.15°/s. Some stimulus-locked measures will be included at the end of the Results section because they often showed appreciably less scatter than the response-locked measures (as in the previous experiments).

Results

Experiment 5A: Dependence on the number of annuli

Increasing the area of the stimulus by increasing the number of annuli while keeping the contrast constant generally increased the magnitude and decreased the latency of the tOFR: see Figure 11, which plots the response-locked measures of the magnitude (Figures 11A–11C) and the latency (Figures 11D–11F) as a function of the area of coverage (expressed as a percentage of the maximum) at each of several contrasts and has separate graphs for each of the three subjects. The abscissas again have logarithmic scales and while a few of the plots seem to be linear others do not. As there are only two or three data points for a given contrast we did not attempt to fit regression lines.

Dependence on the product of the area and the square of the contrast

When the response measures in Figures 11A–11C and 12A–12C for a given subject were replotted as a function of the product of the Total Area of the Stimulus and the Square of its Contrast (referred to as “A*C^2” plots), all of the data points lay on a single monotonic curve: see Figures 13A–13C, which shows these plots for each of the three subjects (in log-linear graphs). The only consistent expectations to this were the responses obtained with the full pattern at the highest contrast (i.e., 29 annuli with 32% contrast), which were always outliers and located well below the rest of the data set. These outliers have been excluded from Figures 13A–13C (one data point per graph) on the grounds that they are subject to some special additional influences (cf. Sheliga et al., 2008a). The smooth curves in Figures 13A–13C are the least-squares best fits obtained when the following Expression was fitted to the plotted data:

$$K \left( 1 - e^{-\frac{\log X - X_0}{\lambda}} \right), \quad (7)$$

where $K$, $X_0$, and $\lambda$ were free parameters. This Expression provided a good fit to the data sets obtained from each of the three subjects (mean $r^2 = 0.877 \pm 0.025$): see Table 8A in the Supplementary material, which lists the best-fit parameters. We also did these fits using a range of different exponents for the contrast parameter (i.e., $A*C^n$). This
indicated that an exponent of unity (i.e., $A^*C$) gave substantially worse fits ($r^2 = 0.689 \pm 0.070$) and that the optimal exponents, which averaged $2.5 \pm 0.4$, resulted in only marginally better fits ($r^2 = 0.896 \pm 0.052$).

When the latency measures for these same responses were plotted as a function of $A^*C^2$ they too all lay on single monotonic curves: see Figures 13D–13F, which shows these data for each of the three subjects, again in log-linear plots. The smooth curves in Figures 13D–13F are the least-squares best fits obtained when the following Expression was fitted to the data:

$$K + \frac{1}{\log A - B} + L_0,$$

where $K$, $B$, and $L_0$ were free parameters. This Expression provided a good fit to the data sets obtained from each of the three subjects ($r^2 = 0.883 \pm 0.014$): see Table 3B in the Supplementary material, which lists the best-fit parameters. We also did these fits using different exponents for the contrast: again, an exponent of unity gave much worse fits ($r^2 = 0.697 \pm 0.047$) and the optimal exponents, which again averaged $2.5 \pm 0.2$, once more gave only slightly better fits ($r^2 = 0.901 \pm 0.016$). Note that the response-locked amplitude measures in Figures 13A–13C were inversely correlated with the corresponding latency measures in Figures 13D–13F ($r^2 = 0.774 \pm 0.090$).

All of the response measures shown so far for Experiment 5 were response-locked (see Methods) because the initial concern was to mirror the methodology that had been used in a previous study of spatial summation in the hOFR (Sheliga et al., 2008a). However, when stimulus-locked response measures were plotted against $A^*C^2$ and fitted with Expression 7, the fits were even better than for the response-locked measures, with a mean $r^2$ value of $0.940 \pm 0.013$: see Figures 13G–13I and Table 8C in the Supplementary material, which lists the best-fit parameters. Yet again, when the contrast had an exponent of unity the fits were substantially worse ($r^2 = 0.751 \pm 0.018$), and when the exponent was optimal for the subject.
Discussion of Experiment 5

Normalization

In Experiment 5 we manipulated the area of the roll-axis stimulus by changing only its radial extent, i.e., changing its dimensions only orthogonal to the circular path of motion. This was achieved by dividing the 1-D radial grating into concentric annuli and then varying their number and radial thickness. This revealed that increasing the area of the rotating stimulus by increasing the number or width of the annuli generally decreased the latency and increased the amplitude of the initial tOFR assessed with response-locked measures (Figures 11 and 12). The changes in amplitude often showed a linear dependence on the logarithm of the area (Figures 11A–11C) and were generally less than expected from a simple vector sum of the tOFRs to the component stimuli. This very gradual saturation is consistent with spatial normalization, which is also a prominent feature of the hOFR (Sheliga et al., 2008a). However, only at the highest contrast used (32%) did the tOFRs show any sign of the complete saturation (leveling off with increases in area) that was commonplace with hOFRs over a wide range of contrasts, i.e., the tOFR was subject to much weaker spatial normalization than the hOFR even though the visual stimuli used were equivalent in terms of their areal coverage and contrast.

The spatial normalization that we reported previously had rendered the hOFR almost totally insensitive to variations in the area of the motion stimulus over a fivefold range. This was seen as a very useful property for a visual tracking mechanism whose performance is probably optimized by an adaptive gain control mechanism (Miles & Kawano, 1986). We argued that any increase in the visual drive to the hOFR resulting from an increase in the physical extent of the stimulus would be equivalent to an increase in the forward-loop gain and hence could result in instability. Ideally, the initial visual drive to the OFR should reflect the velocity of the stimulus and be largely independent of its other physical characteristics. This is clearly not the case for the tOFR, and one
Figure 13. The initial tOFRs obtained with 1-D radial sine-wave gratings subdivided into annuli: dependence on the product of the area and the square of the contrast ($A^*C^2$). Mean response-locked CW-CCW response measures (top row, A–C) and latency estimates (middle row, D–F) from Figures 11 and 12 have been replotted as a function of $A^*C^2$ (on logarithmic abscissas), retaining the same colors and types of symbols except that those from Figure 12 are open symbols and those from Figure 11 are closed symbols; the amplitude measures were normalized with respect to those obtained with stimuli common to both recording sessions (annuli, 0.5°; contrast, 32%); the responses obtained with the full pattern at the highest contrast, i.e., 29 annuli with 32% contrast, have been omitted (see text). Bottom row, G–I: Mean stimulus-locked CW-CCW response measures for the same data sets have been plotted as a function of $A^*C^2$. The curves in A–C and G–I are the best fits obtained with Expression 7, and the curves in D–F are the best fits obtained with Expression 8. The data for each subject are plotted in a separate column: left column, FAM; middle column, BMS; right column, JKM.
possibility is that this tracking mechanism is not subject to adaptive gain control and has a gain that is well below the level at which instability becomes a problem. No studies have examined the adaptive capability of the tOFR and in the present study we found that, on average, the minimal change in area is roughly equivalent to a two-fold change in contrast. We suggest that this is all consistent with the idea that the latency and initial magnitude of the tOFR are determined by the total motion energy in the stimulus. Having already argued earlier (in Experiment 3) that the tOFR is mediated by local spatio-temporal filters sensitive to motion energy, we are now further suggesting that the signals that undergo normalization represent the total motion energy rather than the spatial extent or the contrast of the stimulus.

Initial tOFRs determined by the total motion energy in the stimulus?

The most intriguing observation in Experiment 5 was that, for each subject, the impact of changes in the number, thickness or contrast of the annuli on the magnitude and on the latency of the initial tOFRs was well captured by single monotonic functions over a broad range of the parameter space when the data were plotted with respect to $A^2C^2$ (Figure 13). Further, this was true for the magnitude whether stimulus- or response-locked measures were used, though the former always showed less scatter than the latter, perhaps in part, at least, because of uncertainty in the determination of the exact time of onset of the tOFR.

One implication of these findings is that when the spatial extent of the stimulus is increased orthogonal to the path of the motion, whether by increasing the number and/or thickness of the annuli, the critical factor is simply the total area of the stimulus. Further, changes in area have less impact than changes in contrast, e.g., a four-fold change in area is roughly equivalent to a two-fold change in contrast.

Comparison with the hOFR

The present experiments on the tOFR using concentric annuli evolved from analogous experiments that we had carried out previously on the hOFR using horizontal strips and which had led to the idea of spatial normalization by divisive inhibition (Sheliga et al., 2008a). When we now plotted those hOFR data against $A^2C^2$, it was evident that a single monotonic function sufficed to describe the latency but not the magnitude. The stimulus parameters used in that study of the hOFR were directly comparable with those in the present study of the tOFR, and in Figure 14 we show mean hOFR data for three subjects obtained with stimuli that had three different contrasts (8%; red circles; 16%; blue squares; 32%; black diamonds) and four different areas (1, 3, 7, and 15 strips, providing screen coverage of ~3%, ~10%, ~23%, and 50%, respectively). It is apparent that in Figure 14B, which plots the mean latency of the hOFRs (determined using the same criterion for response onset as in the present study) against $A^2C^2$, the data points were very well fitted by Expression 8 (smooth curve; $r^2 = 0.967$).
However, in Figure 14A, which plots the associated mean hOFR amplitude measures (normalized, response-locked) against $A^*C^2$, it is clear that responses have a much stronger dependency on $C^2$ than on $A$: the data points for a given contrast generally straddle the fitted function (Expression 7, smooth curve; $r^2 = 0.757$) because the dependence on $C^2$ (for any given number of strips) has a pronounced positive slope whereas the dependence on $A$ for any given contrast (when there are 3 or more bands) has either a smaller positive slope (8% contrast) or a slight negative slope (16%, 32% contrasts). The clear implication here is that the total motion energy determines the latency of onset of the hOFR but not its initial amplitude. However, the fact that the spatial normalization is so powerful for the hOFR, whose initial amplitude is independent of area over a five-fold range, precludes any possibility that its amplitude could directly reflect the motion energy in the stimulus.

Some years ago, Miles, Kawano, and Optican (1986) proposed a model in which the latency of the hOFR was determined by the time at which an integrated slip-velocity signal first exceeded a certain threshold level. A similar arrangement might also account for the tOFR latency data. For the tOFR (but not for the hOFR), the latency is inversely related to the initial amplitude of the response, which was ¼ of the fundamental wavelength of the stimulus, and so overlapped (“3f5f stimulus”), i.e., they were separated by a “gap” of $-3^\circ$; (2) one annulus had inner and outer radii of 4.5° and 7.5°, respectively, and the other of 7.5° and 10.5°, respectively, so the two annuli abutted one another, i.e., gap, 0°; (3) one annulus had inner and outer radii of 3.5° and 6.5°, respectively, and the other of 8.5° and 11.5°, respectively, i.e., gap, 2°; and (4) one annulus had inner and outer radii of 2.5° and 5.5°, respectively, and the other of 9.5° and 12.5°, respectively, i.e., gap, 4°. An example of these paired annular stimuli (with a gap of 2°) is shown in Figure 15A, and the diagram in Figure 15B illustrates the layout of the pairs of annuli with each of the 4 separations (one pair for each quadrant). In these illustrated examples, the inner annulus is always the 5f grating but the opposite arrangement was also used (making a total of 7 configurations). The successive angular phase shifts used to generate the roll-axis rotations were always of the same absolute amplitude, which was ¼ of the fundamental wavelength of the 3f5f stimulus, as in Experiment 3, so that the component sine waves advanced in ¼-wavelength steps in opposite directions. The 3f and 5f sine waves had angular wavelengths of 60° and 36°, respectively and could have one of 5 Contrast Ratios (0.33, 0.5, 1.0, 2.0, and 3.0) randomly selected from a lookup table. One concern was to avoid contrast normalization when the pair of annuli overlapped. For this reason the contrasts of the pairs were selected so that, when overlapping, the total contrast of these 3f5f 

**Methods**

The methods and procedures were an amalgamation of those in Experiment 3, which used the 3f5f stimulus, and those in Experiment 5, which used annular gratings.

**Subjects**

Two subjects participated: FAM and BMS.

**Grating stimuli**

The competing visual stimuli consisted of two 1-D radial sine-wave gratings with angular spatial frequencies in the ratio, 3:5, each restricted to a single annulus with a nominal radial thickness of 3° (120 pixels along the cardinal axes). The two annuli could have one of 4 radial separations (“gaps”) in any given trial and were always distributed symmetrically about a radius of 7.5°: (1) both annuli had inner and outer radii of 6° and 9°, respectively, and so overlapped (“3f5f stimulus”), i.e., they were separated by a “gap” of $-3^\circ$; (2) one annulus had inner and outer radii of 4.5° and 7.5°, respectively, and the other of 7.5° and 10.5°, respectively, so the two annuli abutted one another, i.e., gap, 0°; (3) one annulus had inner and outer radii of 3.5° and 6.5°, respectively, and the other of 8.5° and 11.5°, respectively, i.e., gap, 2°; and (4) one annulus had inner and outer radii of 2.5° and 5.5°, respectively, and the other of 9.5° and 12.5°, respectively, i.e., gap, 4°. An example of these paired annular stimuli (with a gap of 2°) is shown in Figure 15A, and the diagram in Figure 15B illustrates the layout of the pairs of annuli with each of the 4 separations (one pair for each quadrant). In these illustrated examples, the inner annulus is always the 5f grating but the opposite arrangement was also used (making a total of 7 configurations). The successive angular phase shifts used to generate the roll-axis rotations were always of the same absolute amplitude, which was ¼ of the fundamental wavelength of the 3f5f stimulus, as in Experiment 3, so that the component sine waves advanced in ¼-wavelength steps in opposite directions. The 3f and 5f sine waves had angular wavelengths of 60° and 36°, respectively and could have one of 5 Contrast Ratios (0.33, 0.5, 1.0, 2.0, and 3.0) randomly selected from a lookup table. One concern was to avoid contrast normalization when the pair of annuli overlapped. For this reason the contrasts of the pairs were selected so that, when overlapping, the total contrast of these 3f5f

**Experiment 6: Two annular radial gratings with competing motions (“3f5f stimuli”)**

In Experiment 3, we reported that the tOFRs elicited by two overlapping 1-D gratings with competing motions (3f5f stimuli) were subject to powerful nonlinear interactions that resulted in WTA behavior when the two gratings differed in contrast by more than an octave. In the present experiment, we used competing 3f and 5f gratings that each consisted of a single, relatively narrow, annulus and we examined the effect of physically separating them so that they no longer overlapped.
stimuli was fixed at 64%: for this, increases in the contrast of one component had to be balanced by decreases in the contrast of the other component. The entries in the lookup table, indicating the Michelson contrast (in %) of the 3f and 5f component pairs, respectively, whether overlapping or not, were 16.7 and 50.2; 22.6 and 45.1; 34.5 and 34.5; 46.2 and 23.1; and 51.6 and 17.2. This same selection of contrasts (and radii) was also used for control trials in which only one annulus was present, i.e., pure 3f and 5f stimuli.

**Procedures**

Each block of trials had 210 randomly interleaved conditions: 5 Contrast Ratios, 3 stimulus types (3f/5f, pure 3f, and pure 5f) for each of the 30 angles. The stimuli consisted of a single annulus (radial thickness, 3°)—rotate in opposite directions: dependence on their relative contrast (Contrast Ratio) and the effect of physical separation. (A) X-Y luminance plot of a sample stimulus configuration in which an inner 5f annulus is separated from an outer 3f annulus by a gap of 2°. (B) Diagram indicating the locations of the two competing annuli when separated by each of the four gaps used (−3°, 0°, 2°, 4°); the two annuli were always radially symmetrical with respect to a circle of radius 7.5° (dotted circle); note that when the gap between them was −3° the two annuli exactly overlapped, and when the gap was 0° the two annuli exactly abutted one another; the reverse configuration (inner 3f, outer 5f) was also used (not shown). (C, D) Response Ratio, computed using Expression 4 and mean CW-CCW (stimulus-locked) response measures, is plotted as a function of the Contrast Ratio, 3f/5f for subjects BMS (C) and FAM (D); preliminary analysis had indicated that the data were essentially the same whether the inner annulus was occupied by the 3f or the 5f grating, so the two sets of data were pooled; black filled circles, gap = −3°; blue open squares, gap = 0°; green closed diamonds, gap = 2°; red open circles, gap = 4°; continuous curves with matching colors are best-fit cumulative Gaussian functions; dotted lines with matching colors show the vector sum predictions. Each datum point based on 93–100 samples (BMS) and 93–100 samples (FAM).

Figure 15. The initial tOFRs obtained when two competing 1-D radial sine-wave gratings (3f/5f stimulus)—each consisting of a single annulus (radial thickness, 3°) —rotate in opposite directions: dependence on their relative contrast (Contrast Ratio) and the effect of physical separation. (A) X–Y luminance plot of a sample stimulus configuration in which an inner 5f annulus is separated from an outer 3f annulus by a gap of 2°. (B) Diagram indicating the locations of the two competing annuli when separated by each of the four gaps used (−3°, 0°, 2°, 4°); the two annuli were always radially symmetrical with respect to a circle of radius 7.5° (dotted circle); note that when the gap between them was −3° the two annuli exactly overlapped, and when the gap was 0° the two annuli exactly abutted one another; the reverse configuration (inner 3f, outer 5f) was also used (not shown). (C, D) Response Ratio, computed using Expression 4 and mean CW-CCW (stimulus-locked) response measures, is plotted as a function of the Contrast Ratio, 3f/5f for subjects BMS (C) and FAM (D); preliminary analysis had indicated that the data were essentially the same whether the inner annulus was occupied by the 3f or the 5f grating, so the two sets of data were pooled; black filled circles, gap = −3°; blue open squares, gap = 0°; green closed diamonds, gap = 2°; red open circles, gap = 4°; continuous curves with matching colors are best-fit cumulative Gaussian functions; dotted lines with matching colors show the vector sum predictions. Each datum point based on 93–100 samples (BMS) and 93–100 samples (FAM).
3f, pure 5f), 7 configurations, and 2 directions of roll-axis rotation.

Results

When the two annuli overlapped, the tOFR data were very similar to those in Experiment 3. Thus, when the contrast of one annulus exceeded that of the other by more than an octave, then the annulus with the lower contrast lost almost all of its influence: WTA. This is apparent from the plots of Response Ratio (computed using Expression 4) against Contrast Ratio, 3f/5f, in Figures 15C and 15D: see the black filled circles. The data in Figures 15C and 15D strongly resemble those in Figure 6B and were well fit by cumulative Gaussian functions (forcing the asymptotes through “0” and “1”) with SDs of 0.16 (FAM, $r^2 = 0.995$) and 0.14 (BMS, $r^2 = 0.998$): see the smooth black curves. Indeed, these SD values were very similar to those obtained with these same subjects in Experiment 3 using overlapping large-field 3f/5f stimuli: 0.15 and 0.12, respectively (from Table 5 in the Supplementary material). The mean Transition Zone (based on these best-fit cumulative Gaussian functions, computed as in Experiment 3 and averaged across the two subjects) extended from 0.70 to 2.22, indicating that, on average, a 1.8-fold difference in contrast sufficed for the stimulus with the lower contrast to almost totally lose its influence.

Separating the two annuli by giving them different radii had a dramatic impact and WTA behavior was no longer evident: see the plots in colored symbols in Figures 15C and 15D. These data were also well fit by cumulative Gaussian functions (mean $r^2 = 0.957$) with SDs that increased steadily as the gap between the annuli increased: when the gaps were 0°, 2°, and 4°, SDs were 0.77, 0.90, and 1.19, respectively, for FAM, and 0.96, 1.61, and 1.64, respectively, for BMS. Note that the effects of the gaps were essentially the same whether the inner annulus was occupied by the 3f or the 5f grating, so that in Figure 15 these two data sets have been pooled. Separating the two annuli never quite eliminated all interactions between them, and this is evident from the fact that the data never quite merge with the vector sum predictions (shown in Figures 15C and 15D as dotted lines with matching colors). However, the required difference in contrast before one annulus lost its influence was now substantial, e.g., on average, a 36-fold difference in contrast was required when the two annuli abutted one another (based on the Transition Zones computed from the best-fit cumulative Gaussian functions).

Discussion of Experiment 6

When the competing 3f and 5f annuli overlapped, the tOFR’s dependence on the relative contrast of the two annuli was highly nonlinear such that, on average, a 1.8-fold difference in contrast was sufficient for the annulus of lower contrast to lose most of its influence: WTA behavior. These data were very similar to those obtained in Experiment 3 with the large-field 3f/5f stimulus, which we postulated reflect mutual inhibition between the neural pathways mediating the opponent motions. The major new finding in the current experiment was that separating the competing annuli reduced these nonlinear interactions substantially, indicating that the postulated mutual inhibition would have to be mostly local. However, there was a tendency for the gratings with the higher contrast to continue to exert a slightly greater influence than expected from the vector sum even with the largest separation used (4°), clearly suggesting that there are some more global inhibitory connections involved, albeit very weak. There have been analogous experiments and findings on the hOFR using 3f/5f stimuli arranged in horizontal strips, though in this case a separation of only 1° was sufficient to eliminate all of the nonlinear interactions (Sheliga et al., 2008b).

General discussion: Neural mediation

Rotation-sensitive neurons

Neurons that are selectively sensitive to visual rotations in the frontal plane have been recorded in the dorsal region of MST (Duffy & Wurtz, 1991a, 1991b, 1995, 1997; Geesaman & Andersen, 1996; Graziano, Andersen, & Snowden, 1994; Heuer & Britten, 2004, 2007; Saito et al., 1986; Takahashi et al., 2007; Tanaka, Fukada, & Saito, 1989; Tanaka & Saito, 1989), the superior temporal polysensory area (Bruce, Desimone & Gross, 1981), and area 7a (Sakata, Shibutani, Ito, & Tsurugai, 1986) of monkeys as well as the dorsolateral pons of the cat (Mower, Gibson, & Glickstein, 1979). The recordings in MST are of particular interest because lesions in this area have implicated it in the genesis of the initial hOFR/vOFR and RFVR (Takeamura et al., 2007), which share many features with the tOFR. Unfortunately, most studies in MST largely ignored the earliest discharges elicited by rotation, which are the only ones relevant to the present study. In fact, the earlier studies of Duffy and Wurtz (1991a, 1991b, 1995) subdivided the responses of MST neurons into initial transient and later tonic components and all of their quantitative assessments were based on the later one, which commenced 400 ms after stimulus onset, i.e., long after the initial tOFRs in the present study were recorded. In a later study, Duffy and Wurtz (1997) found that their initial transient responses were mostly due to the change in luminance that accompanied motion onset in their experiments. Note that in the present study (and in all previous studies of the OFR and RFVR) the stimulus...
pattern always appeared some time before it was moved. In the case of the hOFR, Kawano and colleagues have shown that if the onset of motion coincides with the appearance of the pattern, the initial hOFR is suppressed (Kawano, Shidara, Watanabe, & Yamane, 1994): presumably, the motion detectors are activated by the luminance change (ON-response) without regard for their preferred directions of motion, momentarily freezing the eyes in place. Another problem with most MST single-unit studies that used rotational stimuli is that these were centered within the receptive fields of the neurons rather than on the fovea as in the present study. Exceptions to this were the studies of Mower et al. (1979), Sakata et al. (1986), and Takahashi et al. (2007). However, the study of Takahashi et al. reported only one neuron in MST (out of 127 tested) actually tuned to roll-axis rotational visual stimuli centered on the fovea, though many neurons were broadly tuned for large-field motions and nearly half modulated their activity with such stimuli.

Many studies have reported vertical anisotropies in the efficacy of visual inputs, the most common being a preference for the lower visual field, as in Experiment 1 in the present study on the tOFR (e.g., Amenedo, Pazo-Alvarez, & Cadaveira, 2007; Levine & McAnany, 2005; Pflugshaupt et al., 2009). This might be linked to the overrepresentation of the lower visual field—a rich source of optic flow signals—in the geniculo-striate projections and MT (Maunsell & Van Essen, 1983b; Van Essen, Maunsell, & Bixby, 1981; Van Essen, Newsome, & Maunsell, 1984).

### Local and global processing

A recent series of publications on two ocular reflexes that like the tOFR respond to visual motion at short latency—the hOFR/vOFR and RFVR—developed the hypothesis that both are mediated by MT/MST, where each acquires the global properties that define its sensitivity to optic flow, and share local spatio-temporal properties that are acquired at earlier shared levels of processing (Kodaka et al., 2007; Matsuura, Miura, Tabata, Kawano, & Miles, 2008; Miura et al., 2006; Sheliga et al., 2005, 2008a, 2008b; Sheliga, Kodaka, et al., 2006; Takemura et al., 2007). The shortcomings in the available neuronal data about roll-axis rotations around the fovea notwithstanding, we suggest that the tOFR is yet another reflex mediated by MT/MST. In this scheme, all three tracking systems utilize motion energy signals extracted by direction-selective complex cells in the striate cortex that function as local spatio-temporal filters (Adelson & Bergen, 1985; Emerson, Bergen, & Adelson, 1992; Heeger, 1992a; Watson & Ahumada, 1985). At least some of these complex cells are known to project directly to MT (Movshon & Newsome, 1996), a major source of inputs to MST (Maunsell & Van Essen, 1983a; Ungerleider & Desimone, 1986). Two studies that recorded single unit activity in the cortical motion pathway while manipulating the motion energy in the stimulus are relevant here. The first showed that altering the coherence of dynamic random-dot motion stimuli caused linear changes in the motion energy and resulted in mostly linear changes in the discharge rate of MT neurons (Britten, Shadlen, Newsome, & Movshon, 1993). A second more recent study on cells in MST that were selectively sensitive to radial motion or rotation (or combinations of the same) in the frontal plane also reported that responses showed linear dependence on coherence (Heuer & Britten, 2007). These studies provide strong evidence that neurons in MT and MST linearly encode the motion energy within their receptive fields.

### Motion opponency and normalization

The WTA behavior in Experiments 3 and 6, in which motions in opposite directions had mutually inhibitory effects, is often termed, “motion opponency,” and has substantial supporting evidence from psychophysical studies, functional magnetic resonance imaging, and single unit recordings in area MT and area V1. For review of this evidence see the recent studies of the hOFR, RFVR, and DVR that also used the 3/5f stimulus and reported WTA behavior very similar to that in the present study (Kodaka et al., 2007; Sheliga et al., 2007, 2008b; Sheliga, Kodaka, et al., 2006). Those studies postulated that the WTA behavior was the result of local mutual inhibition between direction-selective complex cells in striate cortex (Rust et al., 2005), though similar local inhibitory interactions in MT might also contribute (Qian & Andersen, 1994; Rust, 2004). We suggest that these same neural mechanisms might also mediate the WTA behavior of the tOFR. Interestingly, a recent study reported similar nonlinear behavior in MSTd neurons when the competing stimuli were of different modality—specifically, visual and vestibular self-motion cues (Morgan, DeAngelis, & Angelaki, 2008). For example, the responses of a neuron could be completely dominated by the motion of the random-dot visual stimulus when the latter’s coherence was high (i.e., the net motion energy in a given direction was high), but were completely dominated by the competing vestibular stimulus when the coherence of the visual motion stimulus was lowered (i.e., the net motion energy in any given direction was low). In that study, the nonlinear interaction was seen as favoring the more reliable source of information, and the same functional argument can be made for the nonlinear interactions that we have described in the tOFR and other ocular tracking reflexes.

Saturation at relatively low contrast (<20%) like that seen in Experiment 2 is also a characteristic of the hOFR/vOFR (Sheliga et al., 2005) and RFVR (Kodaka et al., 2007) and is regarded as a feature of the magnocellular pathway (Mergian & Maunsell, 1993). It is seen as
evidence for contrast gain control or contrast normalization and attributed to divisive inhibition, which is a common feature of cortical motion-selective neurons in striate cortex and MT (Britten & Heuer, 1999; Carandini & Heeger, 1994; Carandini, Heeger, & Movshon, 1997; Heeger, 1992b; Heuer & Britten, 2002; Simoncelli & Heeger, 1998). Of particular interest are recent studies on the population responses to moving plaids in cat striate cortex, which report summation when the two components are of similar contrast and WTA when they are different (cf. our findings in Experiment 3), and show that the divisive normalization model can account for the full range of behaviors (Busse, Katzner, Benucci, & Carandini, 2009).

The tOFR also showed gradual saturation with increases in the spatial extent of the stimulus (Experiment 5), a phenomenon noted in recent reports on the hOFR, in which it was referred to as spatial normalization and attributed to global divisive inhibition (Sheliga et al., 2008a, 2008b). However, this effect was much weaker for the tOFR than for the hOFR. One reason for this might be that divisive inhibition occurs at two (or more?) stages of cortical motion processing: for example, an early one (striate cortex and/or MT?), where it is relatively local and the neurons involved contribute to both the tOFR and the hOFR, and a later one (MST?), where it is more global and restricted to neurons specific to the hOFR.

**Total motion energy**

An interesting observation in Experiment 5 was that the latency and magnitude of the initial tOFRs were each well described by single monotonic functions when plotted against the product, $A^*C^2$, consistent with the hypothesis that these response parameters are determined by the total motion energy in the stimulus. Implicit in this hypothesis is that the signals that are normalized represent the total motion energy rather than just the spatial extent or the contrast of the stimulus. This is all consistent with the findings suggesting that single cells in the cortical motion pathway encode the motion energy within their receptive fields (Britten et al., 1993; Heuer & Britten, 2007). Models that treat direction selective neurons in striate cortex as motion energy filters are increasingly used to examine population responses, e.g., recent studies suggest that the maps of V1 activation obtained with intrinsic-signal optical imaging represent stimulus energy rather than isolated stimulus features such as orientation, direction, and speed (Basole, Kreft-Kerekes, White, & Fitzpatrick, 2006; Basole, White, & Fitzpatrick, 2003; Mante & Carandini, 2005).

When the hOFR data were plotted against $A^*C^2$, a single monotonic function sufficed to describe the latency but not the magnitude (Figure 14) and we pointed out earlier that the latter was inevitable, given that the hOFR was subject to such powerful spatial normalization whereby responses failed to grow with fivefold increases in stimulus area. Thus, the global organization of the tOFR allows some insights about the intrinsic information coding in early cortical motion pathways that are not always possible with the more commonly studied hOFR.

**Acknowledgments**

This research was supported by the Intramural Research Program of the National Eye Institute at the National Institutes of Health.

Commercial relationships: none.

Corresponding author: Boris M. Sheliga.

Email: bms@lsr.nei.nih.gov.

Address: Laboratory of Sensorimotor Research, National Eye Institute, Bldg. 49, Room 2A50, Bethesda, MD 20892, USA.

**Footnotes**

1In the 2-D polar (circular) coordinate system, each point is described by an angle and a radius. A 1-D grating in polar coordinates can have two forms: in a radial grating, luminance modulates with angle but not with radius, and in an angular grating, luminance modulates with radius but not with angle. The stimuli in the present study included 1-D radial gratings and, as in all other studies using these stimuli, the center of rotation was always positioned in the fovea so that the radius roughly equates with retinal eccentricity. Recently, 1-D angular gratings have been used to elicit so-called radial-flow vergence eye movements (Bussettin, Masson, & Miles, 1997; Kodaka et al., 2007; Yang, Fitzgibbon, & Miles, 1999). In the 2-D Cartesian (rectangular) coordinate system, each point is described by an abscissa (horizontal, X) and an ordinate (vertical, Y). A 1-D grating in Cartesian coordinates can have two cardinal forms: in a horizontal grating, luminance modulates with vertical but not with horizontal, and in a vertical grating, luminance modulates with horizontal but not with vertical.

2A setup with a single monitor was also available but the graphics cards used to drive the two (synchronized) monitors used in the Wheatstone stereoscope allowed a greater number of random-dot images to be stored and this was a limiting factor in the design of the experiment. We do not think that the use of separate monitors for the two eyes played a significant role in these experiments.

3The data of JKM showed little dependence on speed over the range studied and were omitted from this.

4The retinal image motion associated with translational disturbances is inversely proportional to the viewing distance (Busettini, Miles, Schwarz, & Carl, 1994).
Based on the mean 5% value and the reciprocal of the mean 95% value.

The SDs of the response distributions of subject JKM were—in absolute terms—higher than those of the other two subjects, i.e., the data were noisier, perhaps accounting for the lack of significance between the “simulated” and “real” distributions in this subject.

Note that the annuli present when there were 7 were all different from those present when there were 8 (Figure 10A), a deliberate attempt to reduce local adaptation effects.

Note that because the data from Experiments 5A and 5B were recorded in different sessions, the amplitude measures were normalized with respect to the data obtained with annuli that were 0.5° wide and had a contrast of 32% (stimuli common to both experiments).

References


