Nonretinotopic perception of orientation: Temporal integration of basic features operates in object-based coordinates

Andreas Wutz
Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy
Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA, USA

Jan Drewes
Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

David Melcher
Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

Early, feed-forward visual processing is organized in a retinotopic reference frame. In contrast, visual feature integration on longer time scales can involve object-based or spatiotopic coordinates. For example, in the Ternus-Pikler (T-P) apparent motion display, object identity is mapped across the object motion path. Here, we report evidence from three experiments supporting nonretinotopic feature integration even for the most paradigmatic example of retinotopically-defined features: orientation. We presented observers with a repeated series of T-P displays in which the perceived rotation of Gabor gratings indicates processing in either retinotopic or object-based coordinates. In Experiment 1, the frequency of perceived retinotopic rotations decreased exponentially for longer interstimulus intervals (ISIs) between T-P display frames, with object-based percepts dominating after about 150–250 ms. In a second experiment, we show that motion and rotation judgments depend on the perception of a moving object during the T-P display ISIs rather than only on temporal factors. In Experiment 3, we cued the observers' attentional state either toward a retinotopic or object motion-based reference frame and then tracked both the observers' eye position and the time course of the perceptual bias while viewing identical T-P display sequences. Overall, we report novel evidence for spatiotemporal integration of even basic visual features such as orientation in nonretinotopic coordinates, in order to support perceptual constancy across self- and object motion.

Introduction

One seminal finding in characterizing the visual system was the discovery of orientation-selective receptive fields in V1 neurons (Hubel & Wiesel, 1962). Primary visual cortex organization resembles a retinotopic map, in which nearby neurons respond to stimuli from adjacent retinal locations with specific feature tuning for visual properties such as orientation, color, motion direction, or shape. For functional magnetic resonance imaging, as well, an orderly arrangement of receptive fields based on retinal location underlies the methodology of retinotopic mapping to delineate different visual processing areas (for review, see Warnking et al., 2002). One could argue that vision science has been built upon the foundation of the retinotopic coordinate system, with orientation encoding being perhaps the most paradigmatic example.

Retinotopic coding might be particularly useful for the initial, feed-forward sweep of visual processing, in which distributed processing areas specialize for encoding of basic visual features such as orientation, motion, or color. Retinotopic organization may help to solve the feature-binding problem by organizing information in terms of spatiotemporal coincidence (Freiwald, 2007; Lin & He, 2009; Melcher, Papathomas, & Vidnyánszky, 2005). Moreover, an important aspect of foveal perception is its ability to support high-acuity processing. During the beginning of a fixation,

Citation: Wutz, A., Drewes, J., & Melcher, D. (2016). Nonretinotopic perception of orientation: Temporal integration of basic features operates in object-based coordinates. *Journal of Vision*, *16*(10):3, 1–15, doi:10.1167/16.10.3.

doi: 10.1167/16.10.3 Received January 20, 2016; published August 5, 2016 ISSN 1534-7362

such high resolution, aided by the retinotopic organization of receptive fields on the fovea, may underlie the ability to make fine discriminations and guide fine motor behavior.

At the same time, this local, retinotopic organization is not sufficient to explain all aspects of our stable perception of space or the binding of features to stable spatiotemporal objects. In fact, the map metaphor breaks down even in V1 with spatial discontinuities for visual quadrants, both in terms of upper and lower visual fields and, more dramatically, for left and right hemifields, in which separate cortical hemispheres are linked by the corpus callosum. Moreover, the firing rate of individual neurons in V1 depends not only on the retinal input but also on the configuration of the surrounding background (Lamme, 1995; Roelfsema, Lamme, & Spekreijse, 1998) providing evidence for contour integration and figure-ground segregation already in primary visual areas (Field, Hayes, & Hess, 1993). Also, basic features such as orientation and motion are processed at multiple levels, not just in V1. Thus, perception of space, even in retinal coordinates, requires putting together information across widely disparate brain areas.

More generally, however, the brain must construct a stable percept of space in a process that takes into account both self-motion and object motion (for review, see Herzog & Öğmen, 2014; Melcher, 2011; Melcher & Morrone, 2015; Öğmen & Herzog, 2010). In the case of self-motion, nonretinotopic effects have been found in studies of perception around the time of eye movements, both before and after the saccade. Single-cell neurophysiology investigations have reported a dramatic change in the spatial selectivity of neurons, called "remapping," in the parietal cortex (Duhamel, Colby, & Goldberg, 1992), frontal eye fields (Umeno & Goldberg, 1997), and in some areas implicated in visual processing (Nakamura & Colby, 2002). In other words, the receptive fields of many neurons cease to be retinotopic, in the traditional sense, around the time of saccades. Nonretinotopic effects have been reported in studies of human perception including orientation encoding (Cha & Chong, 2014; Melcher, 2007; Zimmermann, Morrone, & Burr, 2013; Zimmermann, Morrone, Fink, & Burr, 2013; Zirnsak, Gerhards, Kiani, Lappe, & Hamker, 2011), with the visual system combining presaccadic and postsaccadic orientation information in a nearly optimal way (Ganmor, Landy, & Simoncelli, 2015; Wolf & Schütz, 2015). More generally, there is a growing number of reports of nonretinotopic effects across saccades that involve matching information based on external spatial location (spatiotopic, object-based, or allocentric coordinates) for features including motion (Fracasso, Caramazza, & Melcher, 2010; Melcher & Fracasso, 2012; Melcher & Morrone, 2003; Ong, Hooshvar, Zhang, & Bisley, 2009), form (Demeyer, De Graef, Verfaillie, & Wagemans, 2011; Demeyer, De Graef, Wagemans, & Verfaillie, 2009, 2010; Fracasso et al., 2010; Gordon, Vollmer, & Frankl, 2008; Melcher, 2005; Van Eccelpoel, Germeys, De Graef, & Verfaillie, 2008), and color (Oostwoud Wijdenes, Marshall, & Bays, 2015; Tas, Moore, & Hollingworth, 2014; Wittenberg, Bremmer, & Wachtler, 2008).

It still remains a matter of debate, however, whether visual features are remapped, with some studies failing to find any nonretinotopic transfer of visual features across a saccade or series of saccades, in particular for the case of orientation/tilt (Knapen, Rolfs, Wexler, & Cavanagh, 2010; Mathôt & Theeuwes, 2013). This failure has been taken as evidence for an alternative hypothesis that only an attention index or pointer, devoid of any features, is remapped prior to the saccades while visual representations remain strictly retinotopic (Cavanagh, Hunt, Afraz, & Rolfs, 2010). Thus, understanding whether or not orientation can be integrated over time in nonretinotopic coordinates is vital to characterizing the mechanisms underlying visual stability.

Additional evidence for object-based and frame-based representations can be found in experiments using object motion displays. Examples include the perception of intact objects that are presented moving through a slit (Ağaoğlu, Herzog, & Öğmen, 2012; Parks, 1965; Zöllner, 1862), color fusion along the trajectory of motion (Nishida, Watanabe, Kuriki, & Tokimoto, 2007), and displacements of the object/frame surrounding features (Lin, 2013; Lin & He, 2012). One important set of findings comes from studies using the Ternus-Pikler (T-P) apparent motion display (Figure 1A), which is the focus of the current study. When the two frames of the display are separated by a sufficiently long interstimulus interval (ISI; typically >100 ms), the three disks are seen as moving back and forth as a group and features are combined across the object motion path rather than based on retinal position (Boi, Öğmen, & Herzog, 2011; Boi, Öğmen, Krummenacher, Otto, & Herzog, 2009; Otto, Ögmen, & Herzog, 2006, 2009; Pooresmaeili, Cicchini, Morrone, & Burr, 2012). These studies have reported nonretinotopic effects for form and motion but, at least in the case of tilt adaptation, not for orientation (Boi, Öğmen, et al., 2011). Thus, it remains an open question whether orientation, a paradigmatic example of retinotopic processing, can be represented in a nonretinotopic coordinate system due to object-based grouping.

The fact that the visual system encodes information, at least initially, in retinotopic coordinates raises the question of how the nonretinotopic effects described above, which include cases of both object motion and self-motion, might arise at all. One potential mecha-

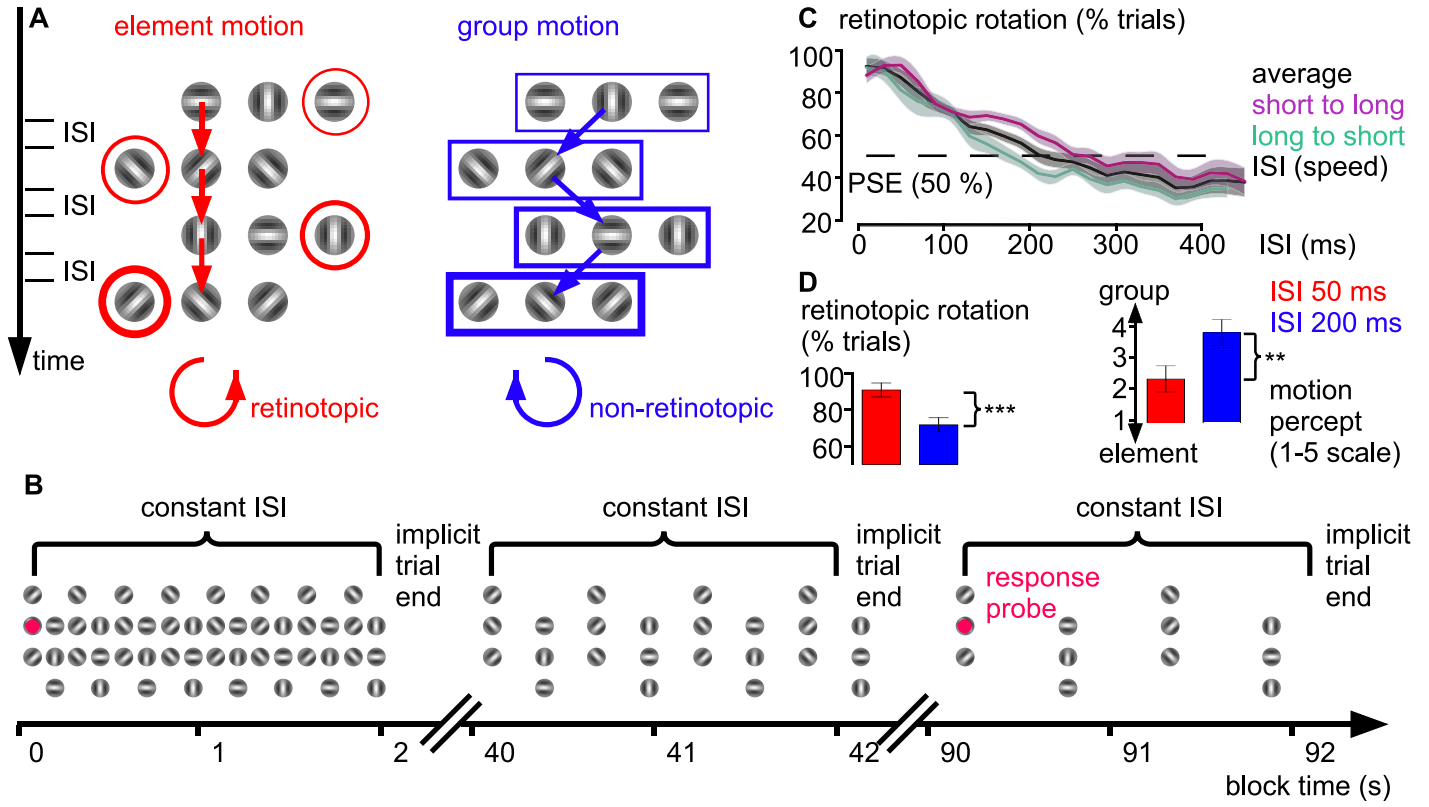

Figure 1. Methods and results for Experiments 1 and 2a. (A) Illustration of the Ternus-Pikler (T-P) sequence used in all experiments. Element motion (red) biases perception toward a static object-percept rotating counterclockwise and group motion toward a dynamic object with clockwise rotation. (B) Illustration of the implicit trial structure in Experiment 1 with fixed ISI for ≈2 s and continuously increasing ISI over the block. A red dot probes the subject's response. (C) Perceived rotation as a function of ISI between T-P displays in Experiment 1. The dashed line denotes subjective equality (PSE) between retinotopic and nonretinotopic rotations. (D) Perceived rotation (left panel) and perceived motion (right panel) for fixed ISIs (50 ms, 200 ms) between T-P displays in Experiment 2a. Shaded areas and error bars reflect ±1 *SEM* (Morey, 2008; see also Franz & Loftus, 2012). Asterisks indicate the significance level for 50 versus 200 ms ISI (with $*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

nism for nonretinotopic perception of visual features is the formation and maintenance of "object files" or "object pointers" (Irwin, 1996; Kahneman & Treisman, 1984; Kahneman, Treisman, & Gibbs, 1992; Melcher & Colby, 2008; Tas, Moore, & Hollingworth, 2012; Wutz & Melcher, 2013). Object-based representations have been theorized to link feature properties to spatiotemporally matched entities ("objects") in order to solve the correspondence problem across object motion or saccades (Herzog & Öğmen, 2014; Öğmen & Herzog, 2010) and to provide a link between perception and cognition (Kahneman et al., 1992).

How might object files support nonretinotopic perception? Starting with each new eye fixation and/or stimulus onset, visual features are initially encoded in retinotopic coordinates. The visual system then accumulates retinal position information into stable, spatial representations and individuates a limited set of object-files in each scene. This supports feature binding in an object-based spatiotemporal reference frame (Xu & Chun, 2009) and avoids motion smear in the retinal image due to object motion or gaze shifts (Öğmen & Herzog, 2010). Recent evidence reported interactions between the capacity of selected or tracked objects and the speed of sensory change, as shown for integration masking (Wutz, Caramazza, & Melcher, 2012) or object motion (Holcombe & Chen, 2013). We suggest that the visual system temporally buffers object individuation and feature integration within a narrow time window of 100 to 150 ms of feedforward input processing (for review, see Wutz & Melcher, 2014). The output of this temporal integration window can then provide sufficient temporal resolution for the individuation of a limited set of object files for subsequent feedback processing of more high-level object properties in nonretinotopic coordinates (Drewes, Zhu, Wutz, & Melcher, 2015; Wutz & Melcher, 2013). Consistent with this idea, studies of nonretinotopic perceptual effects across saccades have reported that such representations take 200 to 300 ms to reach their full expression (Mathôt & Theeuwes, 2011; Zimmermann,

Morrone, & Burr, 2013; Zimmermann, Morrone, Fink, et al., 2013).

In summary, although orientation processing is based on retinotopic coordinates initially, our subjective perception of objects reflects additional processes that are not strictly retinotopic. One potential mechanism for freeing perception from retinotopy may be an object-based reference frame. Here we test whether high-level object perception in a T-P display can bias the temporal integration of basic visual features, such as orientation, toward a nonretinotopic reference frame. In line with object-based theories of visual stability, we predict that when spatial and temporal parameters allow for object matching, nonretinotopic, object-based feature integration will occur, even in case of the most paradigmatic example of a retinotopic visual feature: orientation.

General materials and methods

Participants

All procedures were approved by the ethics committee of the University of Trento and adhere to the Declaration of Helsinki. Participants had normal or corrected-to-normal vision, gave written informed consent before the experimental session, and received a monetary compensation. Ten subjects (seven female, mean 21.9 years ± 2.9 *SD*) took part in Experiment 1. Thirteen subjects (12 female, mean 22.5 years ± 3 *SD*) participated in Experiment 2a, of which a subset of seven participants was also run in in Experiment 2b. Ten subjects (nine female, mean 23.5 years ± 0.9 *SD*) were recruited for Experiment 3.

Apparatus

Visual stimuli were generated using MATLAB 8.2 (MathWorks, Natick, MA) and Psychophysics Toolbox Version 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). In each experiment, subjects viewed the stimuli on a 19-inch CRT monitor running at 100 Hz (ViewSonic, Walnut, CA; 1,024- × 768-pixel resolution in Experiments 1 and 2; 1,280-× 800-pixel resolution in Experiment 3), situated in a dimly lit room. A chin rest controlled for proper head position throughout the session and set the viewing distance at 50 cm in Experiments 1 and 2. In Experiment 3, the viewing distance was set at 54 cm, and an Eyelink 1000 Desktop Mount eye tracker (SR Research, Ontario, Canada) recorded the left eye position at 1000 Hz.

Experimental stimuli

In all experiments, participants viewed a variant of the T-P apparent motion display (Boi et al., 2009; Pikler, 1917; Ternus, 1926). Each stimulus frame contained three Gabor patches (each 0.75° visual angle [v.a.] in Experiment 1 or 1° v.a. in Experiments 2 and 3) presented centered on a uniform gray background for 100 ms. Each Gabor grating was odd symmetric, had a spatial frequency of three cycles per degree (wavelength: 0.25° v.a. in Experiment 1, 0.33° v.a. in Experiments 2 and 3) and a Michelson contrast of 1 on the monitor's luminance range. The standard deviation of the Gaussian envelope was one sixth of the grating size (0.13° v.a. in Experiment 1, 0.17° v.a. in Experiments 2 and 3). The space between adjacent grating positions was 0.125° v.a. in Experiment 1 (total extent of all three gratings 2.5° v.a.) and 1° v.a. in Experiments 2 and 3 (total extent 5° v.a.). After the presentation of a uniform gray blank display for a systematically controlled ISI, the three Gabor gratings were redrawn but horizontally shifted by one grating position. Thereafter, the direction of this horizontal position shift alternated with each next stimulus frame (T-P display). As shown in Figure 1A, each Gabor grating was assigned a specific orientation (0°, 45°, 90°, 135°), such that the reported Gabor grating rotation over time (clockwise, counterclockwise) signaled the combination of Gabor grating orientations in either retinotopic or nonretinotopic coordinates across stimulus frames.

Experimental design and results

Experiment 1

In the first study, we tested the impact of temporal factors on the perceived rotation of the Gabor gratings by continuously varying the ISI between stimulus frames. Each block started with a red central fixation cross on a gray background (0.75° v.a., 1-s duration), followed by four continuous sequences of Gabor grating displays of accelerating or decelerating speed (twice in alternation) with ISIs ranging from 10 ms to 450 ms (in steps of 20 ms). The fixation cross was not shown during grating presentation, in order to avoid disrupting the percept of object motion at this spatial location during the ISI. Instead, participants were instructed to fixate the screen center and covertly attend to the gratings. Each ISI was presented for the number of full rotation cycles (4 × stimulus display + 3 × ISI) closest to 2 s in duration (on average 2.3 ± 0.4 s; i.e., four cycles for 50-ms ISI ≈ 2.2 s; two cycles for 200 ms ISI ≈ 2 s) imposing an implicit trial structure of four ISIs per block (Figure 1B). Participants were

instructed to press a key corresponding to clockwise or counterclockwise based on their perceived rotation of one of the two centrally located Gabor gratings on the screen according to three specific time points during the block: (a) at the start of each block, (b) every time they perceived a change in rotation, and (c) upon every occurrence of a small red dot (0.2° v.a.), presented superimposed on one of the two central Gabor gratings at a pseudo-random position in the block. The probes appeared always at the implicit trial start on approximately every third trial. The response rate on the probes was on average 87.7% (±11 *SD*). The starting speed (decelerating or accelerating) and the Gabor grating rotation were counterbalanced and randomized over the blocks. The starting Gabor grating orientation was random on each block. Each of the 12 blocks lasted about 3.5 min.

The perceived Gabor grating rotation for each ISI was then calculated in terms of retinotopic or nonretinotopic coordinates for each trial, based on the direction of rotation reported (retinotopic rotation, shown in red in Figure 1A). On average, retinotopic rotation responses decreased exponentially with increasing ISI (exponential fit to data: $R^2 = 0.97$; linear fit to log-transformed data: $R^2 = 0.94$), crossing the 75% threshold at 89.8 ms ($P_{0.75}$, 95% confidence interval [CI] = [66.7–116.7 ms]) and the point of subjective equality (PSE; 50%) at 248.1 ms ($P_{0.5}$, CI = [212–290.6 ms]; Figure 1C). When split into trials containing only decelerating or accelerating sequences, an exponential decay provided an equally strong account for the data (both $R^2 = 0.96$), but the 50% threshold estimates for the two psychometric curves were shifted (long to short ISI: $P_{0.5} = 204.9$ ms, CI = [168.1–249.4 ms]; short to long ISI: $P_{0.5} = 292.2$ ms, CI = [250.4–341.8 ms]). Thus, the perceived Gabor grating rotation depended not only on the ISI but also in part on the history of sequence speeds (rate-dependent hysteresis), probably reflecting a tenacious, lingering bias of the dominant element or group motion percept on real-time feature integration. For the single-subject exponential fits (median $R^2 = 0.79 \pm 0.31$ interquartile range [IQR]), the median 50% threshold between perceived retinotopic and nonretinotopic Gabor rotation was at 256.6 ms (±170.1 ms IQR over subjects). In sum, these results demonstrate a strong bias toward retinotopic processing with short ISIs (<100 ms), but on longer time scales (200–300 ms), the rotation percept was either ambiguous or predominantly nonretinotopic for individual observers.

Experiment 2

In the second study, we used T-P displays with a fixed ISI repeated for multiple iterations and a single response for both the perceived rotation and motion on each trial. In Experiment 2a, we aimed to confirm the impact of temporal factors on the perceived Gabor grating rotation at two different fixed ISIs (50 ms and 200 ms) and across a wider spatial distance between adjacent grating locations (1° v.a. vs. 0.125° v.a. in Experiment 1). Each trial started with the presentation of a red, central fixation cross (0.75° v.a.) followed by a uniform gray blank screen for each 500 ms. Then multiple iterations of T-P displays with a fixed ISI were presented. The fixation cross was not shown during grating presentation to avoid disrupting the percept of object motion at this spatial location during the ISI. Instead, participants were instructed to fixate the screen center and covertly attend to the gratings. The first 12 displays contained three black annuli (1° v.a. in size, 0.1° v.a. line width) that served as placeholders before the presentation of 20 displays with Gabor gratings surrounded by the annuli (five full rotation cycles). After the presentation of a gray blank screen (500 ms), participants were required to first report the perceived Gabor grating rotation (by pressing a corresponding key) and then judge the perceived object motion by pressing a number key on a scale from 1 to 5 (1 = element motion; 5 = group motion). Participants were instructed to report their dominant rotation and motion percept during the entire display sequence. The ISI between displays (50 or 200 ms) and the Gabor grating rotation were counterbalanced and randomized in each block. The starting Gabor grating orientation was random on each trial. Each of the four blocks contained 20 trials and lasted about 4.3 min.

The temporal factors of the T-P sequence (50 ms vs. 200 ms ISI) had a strong impact on both measures of retinotopic versus nonretinotopic processing (Gabor grating rotation, object motion). In accordance with previous results (Boi et al., 2009), participants predominantly reported an element motion percept with a 50-ms ISI compared with stronger group motion with 200 ms ISI ($t[12] = -3.5$, $p < 0.004$). Replicating the findings of Experiment 1, a short ISI between displays also yielded significantly more retinotopic rotation responses compared with a long ISI (50 ms vs. 200 ms, $t[12] = 5$, $p < 0.001$; Figure 1D). In the majority of trials, however, participants perceived a Gabor grating rotation in retinotopic coordinates for both short and long ISIs (50-ms ISI, 87.5% ± 18% *SD*, *t* test versus PSE (50%), $t[12] = 7.4$, $p < 0.001$; 200-ms ISI: 70% ± 22% *SD*, $t[12] = 3.2$, $p < 0.008$), consistent with the idea that the temporal factors alone may not be sufficient for a rotation percept in nonretinotopic coordinates (for review, see Petersik & Rice, 2006).

We hypothesized that the perception of object motion might be a determining factor for nonretinotopic processing, in addition to the temporal parameters of the T-P display. Indeed, in Experiment 2a,

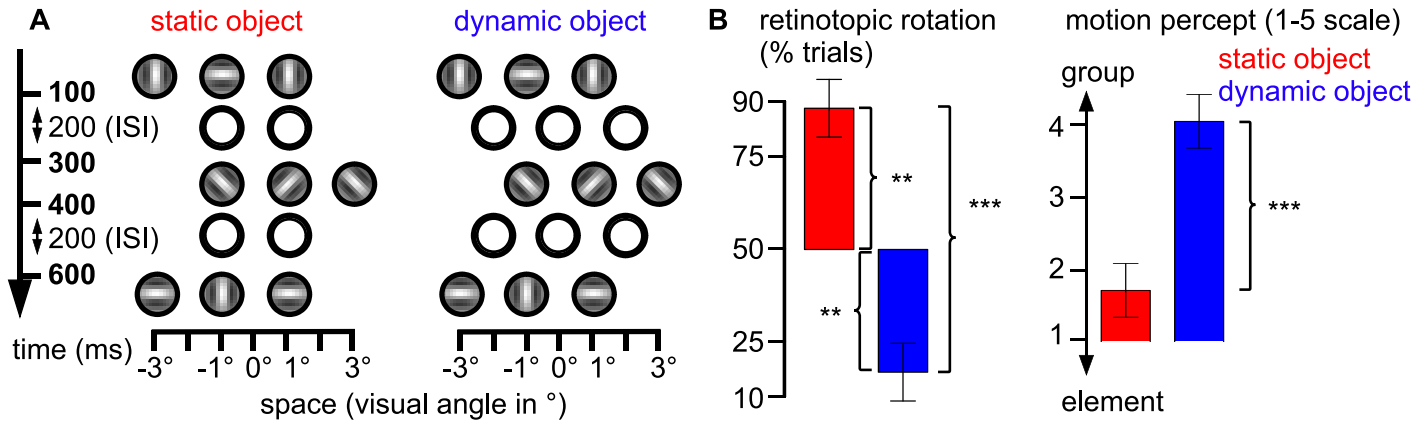

Figure 2. Methods and results for Experiment 2b. (A) Illustration of the modified T-P sequence used in Experiment 2b. The insertion of annuli during the fixed ISI (200 ms) is hypothesized to bias perception toward either a static (red) or dynamic object (blue). (B) Perceived rotation (left panel) and perceived motion (right panel) for static and dynamic object cueing in Experiment 2b. Error bars reflect ±1 *SEM* (Morey, 2008; see also Franz & Loftus, 2012). Asterisks indicate the significance level for static/dynamic trials against the PSE (50%) and against each other (with *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

participants reported a stronger percept of element motion in retinotopic rotation trials ($M = 3.1 \pm 0.7$ *SD*) and more group motion in nonretinotopic rotation trials ($M = 3.6 \pm 0.7$ *SD*; $t[12] = -2.2$, $p < 0.05$). Likewise, when splitting the data at the median motion response (average median 3.2 ± 0.9 *SD*), we observed the expected pattern of more retinotopic processing with element motion (81.4% ± 24% *SD* for below median) and more nonretinotopic processing with group motion (76.3% ± 19% *SD* for above median; $t[11] = 1.2$, $p =$ n.s.). Experiment 2b was designed to explicitly test the relation between motion perception and nonretinotopic processing by cueing either static spatial object positions (element motion) or apparent group motion while keeping the ISI between Gabor grating stimulus frames fixed at 200 ms (Figure 2A). Either two annuli (1° v.a. in size, 0.1° v.a. line width) at the two central grating locations (cueing a static object) or three annuli at intermediate grating locations (cueing a dynamic object) were presented during the ISI between stimulus frames. The annuli were briefly flashed for 40 ms, preceded and followed by a gray blank display 80 ms in duration for each, preserving the total ISI duration between grating displays (modified ISI). The T-P display sequences with different object cues were run in separate blocks, starting with 12 display iterations (stimulus presentation + modified ISI) containing only annuli and then 20 display iterations with Gabor gratings and annuli (five full rotation cycles). All other experimental procedures were identical to Experiment 2a. Each of the four blocks (two each for static/dynamic cue; block order counterbalanced over subjects) contained 20 trials and lasted about 4.3 min.

We found strong effects of static versus dynamic object cueing on both measures of nonretinotopic processing. For the motion percept, the static cue resulted in significantly stronger element motion compared with stronger group motion for the dynamic cue (static versus dynamic, $t[6] = -6.3$, $p < 0.001$). Critically, the frequency of retinotopic rotation reports differed strongly between object cues (static versus dynamic, $t[6] = 9.2$, $p < 0.001$), and was significantly above the PSE (50%) in the static case (85.4% ± 20% *SD*, $t[6] = 4.8$, $p < 0.003$) and significantly below the PSE in the dynamic case (19.3% ± 19% *SD*, $t[6] = -4.3$, $p < 0.005$; Figure 2B). Splitting the data at the median motion response (average median $M = 3.1 \pm 0.9$ *SD*) or between (non-) retinotopic processing showed that the two measures were highly interdependent (below median: 79.8% ± 20% *SD* retinotopic rotations; above median: 28.7% ± 20 *SD*, $t[6] = 5$, $p < 0.003$; retinotopic trials: 2 ± 0.6 *SD* average motion response; non-retinotopic trials: 3.9 ± 0.8 *SD*, $t[6] = -5.4$, $p < 0.002$). These findings suggest that an object-based frame of reference, defining either a static object or an object in motion, determines the perceived Gabor grating rotation in retinotopic or nonretinotopic coordinates.

Experiment 3

In the final experiment, the T-P display sequence was designed to measure the nature and time course of this object-based processing bias with physically identical grating displays. On each trial, we first induced either a static or dynamic object-based reference frame and then tested processing at several, successive intervals from this inducing phase. Concurrently, we recorded the participants' eye position during the entire T-P display sequence (inducer + test) to precisely track the spatial location and timing of the retinal input during

perception in retinotopic or nonretinotopic coordinate frames. The eye position data were important for two main reasons. First, eye position provided an additional, independent measure of perceived motion indicated by some small eye movements (Boman & Hotson, 1988, 1989; Schor, Lakshminarayanan, & Narayan, 1984). Second, we wanted to ensure that participants were following instructions and mainly kept their eyes near the fixation point, even if they did make small deviations in eye position. If they entirely displaced their gaze to the position of the object on each frame, then object-based coordinates would be identical to retinal coordinates. Given the long value of ISI, it is theoretically possible that previous studies of nonretinotopic perception with the T-P display also included trials in which participants moved their eyes along with the object motion, although this potential confound has been excluded in studies of slit viewing (Rieger, Grüschow, Heinze, & Fendrich, 2007).

A five-point calibration was performed at the start of each block. The trial started with a drift check, followed by the presentation of a central, red fixation cross (0.75°) and a uniform gray blank screen for each 500 ms. Then the T-P sequence began with 12 display iterations containing only placeholder annuli (1° in size, 0.1° line width) and designed to cue the participants' reference frame either toward a static or dynamic object motion path (modified ISI, see Experiment 2b). The fixation cross was not shown during grating presentation to avoid disrupting the percept of object motion at this spatial location during the ISI. Instead, participants were instructed to fixate the screen center and covertly attend to the gratings. After four display iterations with annuli and Gabor gratings together (one full rotation cycle), 16 iterations with only Gabor gratings were presented (four full rotation cycles). In separate blocks, different reference frame cues (static/dynamic inducer) were shown, but the presentation of the Gabor gratings with a fixed 200-ms ISI was identical across blocks (Figure 3A). On each trial, the participants' perceived Gabor grating rotation was probed with a small red dot (0.2°) presented superimposed on one of the two central Gabor gratings during one of the five rotation cycle following the inducers (always at the start of the cycle from 3.6 s in steps of 1.2 s in Figure 3A). Participants were instructed to press a corresponding key (clockwise, counterclockwise) based on their perceived rotation after probe onset. At the end of each trial, participants judged their perceived object motion by pressing a number key on a scale from 1 to 5 (1 = element motion; 5 = group motion). Participants were instructed to report their motion percept at the time of the probe presentation. The probe timing (first to fifth rotation cycle) and the Gabor grating rotation were counterbalanced and randomized in each block. The starting Gabor grating orientation was random on each trial. Each of the 10 blocks (five each for static/dynamic inducer; block order counterbalanced over subjects) contained 20 trials and lasted about 5 min.

As shown in Figure 3B, the horizontal eye position did vary as a function of the display sequence, in particular for the dynamic object-motion inducer condition, at a 1.67-Hz rhythm (600-ms cycle period, 2× stimulus presentation + ISI). As can be noted from the figure, there were more deviations in gaze position for the group motion displays, as predicted, but the amplitude of the gaze variations was much smaller than the displacement of the images (ca. 25% of a pure sinus with 1° of visual angle). Thus, although perception of group motion did drive the oculomotor system, subjects maintained their gaze position to be near the fixation point rather than moving their eyes the entire 2° of visual angle of the element displacement between Gabor grating centers.

For both static and dynamic inducer blocks, we quantified the horizontal amplitude, measuring its spatial magnitude, and the intertrial phase concentration of the eye trace, measuring its timing precision, by its 1.67-Hz frequency characteristics in three successive trial segments (inducer phase: 0–4.8 s; two test phases: 4.8–7.2 s and 7.2–9.6 s). Before applying the Fourier transform, the horizontal eye trace of each trial segment was smoothed with a 10-ms average sliding window, detrended by subtracting the linear fit and tapered with a symmetric Hanning window. To estimate the spatial extent of the horizontal eye trace in each time segment, we computed the absolute value of the Fourier spectra and averaged over trials in the frequency domain. Thus, amplitude values represent the spatial extent of the eye movements independent of the exact timing on each trial. These data were used for the statistical tests reported below (see also the colored dot inset in Figure 3C). Similar results were found when first averaging over trials in the time domain before applying the Fourier transform. These data are shown in Figure 3C because of the reduced 1/f noise in the time-averaged amplitude spectra. The amplitude spectra were normalized relative to the expected values of a pure 1.67-Hz sinusoidal function covering 1° of visual angle and are thus reported as amplitude percentage (i.e., abs[$\text{Fourier}_{\text{trial i}}$ at 1.67 Hz]/abs[$\text{Fourier}_{\text{sinus 1°}}$ at 1.67 Hz].

To estimate the timing precision, we quantified the amount of phase coherence across trials by the resultant length of the complex Fourier vectors along the unit circle (intertrial coherence [ITC]; Makeig, Debener, Onton, & Delorme, 2004; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). ITC can take values between 0 (random phase angle distribution across trials) and 1 (perfect synchronization). All data partitions were equated in trial number before the

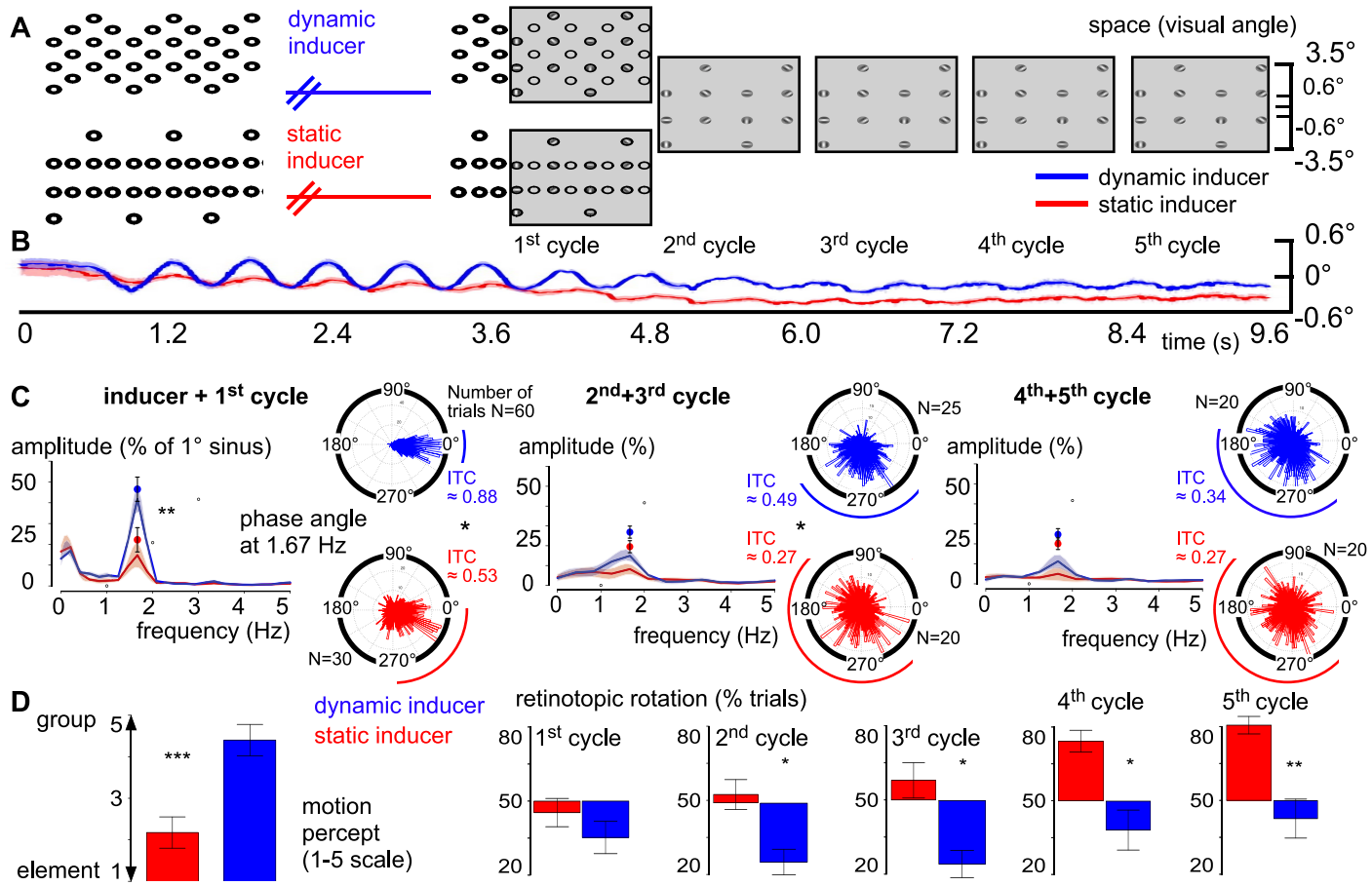

Figure 3. Methods and results for Experiment 3. (A) Illustration of the trial sequence in Experiment 3. In separate blocks, either static (red) or dynamic inducing sequences (blue) were presented (0–3.6 s), followed by one full rotation cycle with annuli + gratings (3.6–4.8 s). Then the same grating sequences were shown for four full rotation cycles (4.8–9.6 s). (B) Averaged horizontal eye traces as a function of trial time for each inducer block (static/dynamic). (C) Fourier statistics of the horizontal eye traces for consecutive trial segments (left panels: 0–4.8 s; middle panels: 4.8–7.2 s; right panels: 7.2–9.6 s). Left plots: Eye movement amplitude in percentage of a pure 1° sinus as a function of frequency based on each subject's averaged time course per inducer block. The inset red/blue dots show eye movement amplitude averaged over frequency spectra (instead of time series) as reported in the main text. Right plots: Phase angles for all trials and all subjects at 1.67 Hz (right plots) per inducer block. (D) Behavioral performance. Left panel: Perceived motion response per inducer block (static/dynamic). Right panel: Perceived rotation response per inducer block and trial segment (left to right plots: 3.6–9.6 s for 1.2 s each, one rotation cycle). Shaded areas and error bars reflect ±1 *SEM* (Morey, 2008; see also Franz & Loftus, 2012). Asterisks indicate the significance level for static versus dynamic inducer blocks (with $*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

phase analysis. Eye movement (amplitude and ITC at 1.67 Hz) and behavioral measures (rotation and motion responses) were compared between static and dynamic inducer blocks and across different time intervals from the inducer sequence (Block × Time design; three trial segments for eye movements, five rotation cycles for behavior).

For both eye movement measures, a two-way within-subjects analysis of variance (ANOVA) revealed significant main effects of the factors Inducer Block and Trial Segment, as well as a significant interaction between these main factors (Table 1). The strongest eye movement amplitude was found during the dynamic inducer sequences at trial start (0–4.8 s), covering on

			Amplitude		ITC	
Eye movements	*df*1	df2	*F*	*p*	*F*	*p*
Inducer	1	9	18	0.003	13	0.006
Time	2	18	7.1	0.006	43.8	0.001
Inducer × Time	2	18	8.4	0.003	6.1	0.01

Table 1. Two-way within-subjects ANOVA for the eye movement data in Experiment 3. *Notes*: Degrees of freedom (*df*), *F* values, and corresponding *p* values for the main factors (inducer block: static vs. dynamic; time: first to third trial segment) and the interaction for the eye movement data (amplitude and intertrial coherence [ITC]).

average 47% of a pure sinus with 1° visual angle (±14% *SD*) and differing significantly from static inducer sequences (22% ± 14% *SD*, $t[9] = -4.2$, $p < 0.008$; all reported t tests Bonferroni-corrected for multiple comparisons). Later in the trial, the eye movement amplitude decreased and showed no significant differences between inducer blocks (4.8–7.2 s, static 20% ± 12% *SD*, dynamic 27% ± 11% *SD*, $t[9] = -2.4$, $p < 0.12$; 7.2–9.6 s: static 20% ± 12% *SD*, dynamic 25% ± 16% *SD*, $t[9] = -1.6$, $p < 0.43$). In both static and dynamic inducer blocks, the horizontal eye trace stabilized between the two central Gabor gratings (±0.5° relative to the screen center) during the test intervals (4.8–9.6 s; Figure 3B and C). A similar pattern was found in terms of timing precision measured with ITC. At trial start (0–4.8 s), dynamic sequences induced a more consistent eye movement pattern over trials compared with static sequences (static ITC = 0.53 ± 0.3 SD, dynamic ITC = 0.88 ± 0.12 SD), $t(9) = -3.7$, $p < 0.016$. This timing precision decreased over the trial and differed—if at all—only marginally between inducer blocks during the later test intervals (4.8–7.2 s, static ITC 0.27 ± 0.17 *SD*, dynamic ITC 0.49 ± 0.15 *SD*, $t[9] = -3.1$, $p < 0.041$; 7.2–9.6 s, static 0.27 ± 0.15 *SD*, dynamic 0.34 ± 0.15 *SD*, $t[9] = -1.4$, $p < 0.58$). Moreover, the eye trace phase angle tracked the stimulation sequence only at trial start (0–4.8 s, average phase angle difference: dynamic −10.3° ± 13.5° *SD*, static −37.3° ± 58.9° *SD*), but eye trace and stimulation shifted out of phase during the test intervals later in the trial (4.8–7.2 s: static −141.9° ± 62.4° *SD*, dynamic −101.3° ± 40.3° *SD*; 7.2–9.6 s: static −134.7° ± 62.6° *SD*, dynamic −134.1° ± 43.5° *SD*; Figure 3C). In sum, these findings suggest that both eye movement amplitude and precision were largely independent from the stimulation sequence and relatively stable during the test intervals for both static and dynamic inducer blocks.

For the reported motion percept (element-group motion), a two-way within-subjects ANOVA revealed a significant main effect of inducer block (Table 2). Although viewing identical T-P display sequences, participants perceived element motion more often in static inducer blocks compared with stronger group motion in dynamic inducer blocks, and this was found for all of the test intervals (post hoc t tests static vs. dynamic for first to fifth rotation cycle: all $t[9] < -5.4$, all $p < 0.002$). For the perceived Gabor grating rotation, the two-way ANOVA showed significant main and also interaction effects (Table 2). From the second to the fifth test interval (rotation cycle), we found a significant bias in the reported grating rotation between static and dynamic inducer blocks (first cycle, $t[9] = 1.3$, $p =$ n.s.; second cycle, $t[9] = 3.4$, $p < 0.041$; third cycle, $t[9] = 3.5$, $p < 0.037$; fourth cycle, $t[9] = 4.2$, $p < 0.013$; fifth cycle, $t[9] = 5.4$, $p < 0.003$). In dynamic inducer blocks, participants perceived the grating rotation predominantly in nonretinotopic coordinates immediately after the inducer sequence (post hoc t test vs. PSE [50%]: first cycle, $t[9] = -3.3$, $p < 0.05$; second cycle, $t[9] = -5.5$, $p < 0.003$; third cycle, $t[9] = -5.6$, $p < 0.002$; fourth cycle, $t[9] = -1.5$, $p =$ n.s.; fifth cycle, $t[9] = -0.8$, $p =$ n.s.). For static inducer blocks, in contrast, retinotopic rotation responses increased over the trial and were significantly above the PSE for the last two rotation cycles (first cycle, $t[9] = -0.6$, $p =$ n.s.; second cycle, $t[9] = 0.5$, $p =$ n.s.; third cycle, $t[9] = 0.8$, $p =$ n.s.; fourth cycle, $t[9] = 3.7$, $p < 0.027$; fifth cycle, $t[9] = 5.8$, $p < 0.002$; Figure 3D). Overall, motion perception and (non-) retinotopic processing were highly interdependent, replicating the findings from Experiment 2b: below median, 60.2% ± 23% *SD* retinotopic rotations and above median (36.7% ± 11% *SD*, $t[9] = 2.8$, $p < 0.02$; retinotopic trials, 3.1 ± 0.9 *SD* average motion response; nonretinotopic trials, 4 ± 0.5 *SD*, $t[9] = -2.9$, $p < 0.017$). In sum, we found strong differences between static and dynamic inducer blocks for both behavioral measures (the perceived object motion and the Gabor grating rotation) that result from a shift from nonretinotopic/object-based to retinotopic coordinate frames with increasing time from the inducing sequence.

			Motion		Rotation	
Behavior	*df*1	*df*2	*F*	*p*	*F*	*p*
Inducer	1	9	34.9	0.001	18.4	0.002
Time	4	36	0.6	0.65	7.1	0.001
Inducer × Time	4	36	1.1	0.38	5	0.003

Table 2. Two-way within-subjects ANOVA for the behavioral data in Experiment 3. *Notes*. Degrees of freedom (*df*), *F* values, and corresponding *p* values for the main factors (inducer block: static vs. dynamic; time: first to fifth rotation cycle) and the interaction for the behavioral data (motion and rotation percept).

The reported eye movement data analysis suggests that the observed effects of static versus dynamic object cueing on nonretinotopic processing were largely independent from the participant's eye position. Nevertheless, we wanted to exclude the possibility that fast gaze shifts between grating locations via saccadic eye movements might have influenced the results. To this end, we detected saccades in the horizontal eye traces of each participant (saccade velocity threshold: 35° visual angle/s) and reanalyzed the data excluding possibly contaminated trials. Static and dynamic inducer blocks were comparable both in the number of saccades during the 9.6-s-long trial sequence (static: 11.3 ± 5.3 *SD*, dynamic: 12.5 ± 4.9 *SD*, $t[9] = -1.5$, $p =$ n.s.) and in terms of saccade amplitude (median absolute gaze position difference before and after saccade: static, 0.57° ± 0.1° *SD*; dynamic, 0.61° ± 0.08°

SD, $t[9] = -1.9$, $p =$ n.s.). Saccades could be problematic for the interpretation of the results if they covered the space between adjacent gratings (1° visual angle between grating edges, 2° visual angle between grating centers) and/or if their timing coincided with grating presentation (100-ms stimulus duration, 200-ms ISI). Trials were excluded when the therein-contained saccades occurred during the Gabor grating presentation period (4.8–9.6 s), had a spatial extent between 0.75° to 3° visual angle, and were followed by a second saccade in the opposite direction within 200 to 400 ms. This led to the rejection of on average 19.7 trials (±20.4 *SD*) for static blocks and 26.8 trials (±21.6 *SD*) for dynamic blocks. Between static and dynamic blocks, we were able to confirm the main results both for perceived object motion (static: 2.3 ± 1.2 *SD* average motion response; dynamic: 4.7 ± 0.4 *SD*; $t[9] = -5.8$, $p < 0.001$) and for (non-) retinotopic feature integration (static: 64.5% ± 19% *SD* retinotopic rotations; dynamic: 35.2% ± 13% *SD*; $t[9] = 4.6$, $p < 0.002$), even when excluding those saccade-contaminated trials.

Discussion

The main findings are that participants reported nonretinotopic, object-based integration of orientation in a T-P display, and this nonretinotopic percept was influenced by both object-based cues and temporal factors (ISI and time from inducer/cue). We found strong evidence for nonretinotopic processing when a sufficiently long ISI allowed for the perception of a moving object between stimulus frames (Experiments 1 and 2a). Moreover, both perceived object motion and the integration of orientation were influenced by the presentation of cues, either between the frames of the T-P sequence (Experiment 2b) or with inducers at the beginning of the trial (Experiment 3).

In the third experiment, we also measured eye position to outline the relation between the observers' gaze in retinal coordinates and their perceived object motion and/or temporal integration of the oriented Gabor gratings. Dynamic inducers, which increased the percept of object motion and nonretinotopic integration, also increased the amplitude and timing precision of eye movements compared with static inducers. However, the magnitude of these eye movements was small compared with the stimulus displacements, and both the magnitude and precision of these eye movements tended to decay over time and to shift out of sync with the T-P display. Thus, the amplitude and precision of eye movements were largely independent from the stimulus sequence and relatively stable during the test intervals for both types of inducers, during which we still observed reliable differences in both perceived motion and feature integration. Many previous studies using the T-P display did not measure eye position, raising the question of whether eye movements that exactly tracked the stimulus shifts might have been a confound in those studies. The findings reported here are not consistent with an account based on successive sampling of grating locations through gaze shifts and subsequent feature integration over the retinal samples. Instead, the pattern of results suggests that for T-P displays, such as with slit viewing (Rieger et al., 2007), perceptual experience involves nonretinotopic representations.

Previous work has shown a relationship between stimulus orientation and perception in the T-P display. The orientation of the Gabor stimuli can bias perception toward either the group or element interpretation of motion (Alais & Lorenceau, 2002; Scott-Samuel & Hess, 2001; Wallace & Scott-Samuel, 2007). For example, horizontally oriented Gabors would tend to favor an interpretation of horizontal motion. Conversely, the current results show how the interpretation of motion, in terms of group or element motion, can bias the integration of orientation information.

Beyond retinotopy

Our results indicate that temporal integration of even basic visual features such as orientation can be biased toward nonretinotopic processing, which may support perceptual constancy and object-based representations. These results confirm previous findings for other features with the T-P display, as well as a variety of other paradigms using moving objects as stimuli (Ağaoğlu et al., 2012; Nishida et al., 2007; Otto et al., 2006; Parks, 1965; Zöllner, 1862). However, a previous study using the T-P display and the tilt aftereffect rather than temporal integration for rotation judgments did not find object-based effects (Boi, Öğmen, & Herzog, 2011).

Likewise, perceptual learning has been demonstrated in nonretinotopic coordinates for orientation (Otto, Öğmen, & Herzog, 2010) and motion direction discrimination (Zhang & Li, 2010). Nonretinotopic effects for orientation judgments have been reported across saccades (Cha & Chong, 2014; Melcher, 2005, 2007; Zimmermann, Morrone, & Burr, 2013; Zimmermann, Morrone, Fink, et al., 2013), consistent with an optimal integration of presaccadic and postsaccadic orientation information for the same spatial location (Ganmor et al., 2015; Wolf & Schütz, 2015). In terms of nonretinotopic effects for orientation and shifts in gaze, the tilt aftereffect has been reported to compress toward the saccade target in the perisaccadic time period (Melcher, 2007; Zirnsak et al., 2011) and show spatiotopic effects over a period of hundreds of

milliseconds (Cha & Chong, 2014; Melcher, 2005; Zimmermann, Morrone, & Burr, 2013; Zimmermann, Morrone, Fink et al., 2013). The current study provides further evidence that orientation information can be temporally integrated in nonretinotopic coordinates and furthermore emphasizes the importance of both temporal factors and the object-based interpretation of the two stimulus presentations.

Spatiotemporal object perception in a phenomenal reference frame

The current findings are consistent with the idea that dynamic scene perception, such as in the case of the T-P display, is based on the construction of spatiotemporal reference frames that cannot be reduced to just spatial or temporal factors (Ağaoğlu et al., 2012; Gepshtein & Kubovy, 2000). Moreover, previous studies have shown that object features can also influence this grouping process (Feldman & Tremoulet, 2006). The current findings are consistent with the idea that spatial, temporal, and object/Gestalt grouping all interact in determining our subjective perceptual interpretation of the stimulus.

In the current study, both perception of group motion and the coordinate system underlying temporal integration of orientation depended on ISI duration but also on the nature of the inducers. In the final experiment, we found that the influence of the dynamic inducer decayed over time, whereas retinotopic processing built up with increasing time from the inducing sequence. In both cases, the temporal factors might be related to the way objects are individuated and tracked. Object formation is not instantaneous, but instead, the binding of object representations to specific spatial locations evolves over time (for review, see Wutz & Melcher, 2014). To account for sensory changes through object or self-motion during this feedforward computation, the visual system adopts temporal integration windows of limited duration (100–150 ms). With longer time windows, the goal of achieving object stability balances with the needs for sensitivity to new sensory samples (Wutz et al., 2012). It is interesting to note that the longer time periods implicated here are on the same order as fixation durations (Rayner, 2009; Walshe & Nuthmann, 2014). Probably it is not coincidental then that feature integration over relatively long time scales, which allows for the stable perception of objects, was biased toward nonretinotopic processing.

In addition to feedforward processing, increasing evidence suggests important contributions to sensory encoding by feedback (reentrant) projections from higher visual areas (Ro, Breitmeyer, Burton, Singhal, & Lane, 2003; for review, see Lamme & Roelfsema, 2000). Indeed, our perception of even local features reflects also the influence of these top-down signals altering the bottom-up information such that we perceive a cognitive interpretation rather than the actual sensory input. One possible solution to the problem of how to carve the continuous flow of information from the retina to the cortex into coherent percepts, despite the presence of feedback loops, might involve the temporal organization of sensory sampling and consolidation in distinct time windows. In particular, brain oscillations could provide natural time frames for sensory integration with feedforward and feedback signals sent through distinct frequency channels from and to visual cortex (Bastos et al., 2015; Wutz, Muschter, van Koningsbruggen, Weisz, & Melcher, 2016; Wutz, Weisz, Braun, & Melcher, 2014). Reentrant processes may then operate on the entire integration buffer to consolidate meaningful and coherent object percepts despite changes in retinal position. Because of this temporal organization, low-level feature integration over longer time scales may be biased toward nonretinotopic processing, in order to directly support the perceived constancy of moving object representations.

Conclusion

Consistent with recent evidence (Boi, Vergeer, Ogmen, & Herzog, 2011), our findings suggest that subjective perceptual experience does not always reflect a hard-coded, low-level processing architecture but instead depends on the construction of a "phenomenal reference frame" (Rock & Brosgole, 1964; Rock & Ebenholtz, 1962). Perceptual experience reflects a spatiotemporal construct, organized around stable objects and dynamic (four-dimensional) scenes. This construct goes beyond retinotopy to support active, real-world vision.

Keywords: Ternus-Pikler display, orientation, retinotopic, object-based, eye movements

Acknowledgments

This research was supported by a European Research Council Grant "Construction of Perceptual Space-Time" (CoPeST; agreement No. 313658).

Commercial relationships: none.
Corresponding author: Andreas Wutz.
Email: andreas.wutz-1@unitn.it.
Address: Palazzo Fedrigotti, Corso Bettini 31, 38068 Rovereto (TN), Italy.

References

Ağaoğlu, M. N., Herzog, M. H., & Öğmen, H. (2012). Non-retinotopic feature processing in the absence of retinotopic spatial layout and the construction of perceptual space from motion. *Vision Research*, *71*, 10–17.

Alais, D., & Lorenceau, J. (2002). Perceptual grouping in the Ternus display: Evidence for an 'association field' in apparent motion. *Vision Research*, *42*, 1005–1016.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., … Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*, 390–401.

Boi, M., Öğmen, H., & Herzog, M. H. (2011). Motion and tilt aftereffects occur largely in retinal, not in object, coordinates in the Ternus–Pikler display. *Journal of Vision*, *11*(3):7, 1–11, doi:10.1167/11.3.7. [PubMed] [Article]

Boi, M., Öğmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino-vs. non-retinotopic processing. *Journal of Vision*, *9*(13):5, 1–11, doi:10.1167/9.13.5. [PubMed] [Article]

Boi, M., Vergeer, M., Ogmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology*, *21*, 1732–1737.

Boman, D. K., & Hotson, J. R. (1988). Stimulus conditions that enhance anticipatory slow eye movements. *Vision Research*, *28*, 1157–1165.

Boman, D. K., & Hotson, J. R. (1989). Motion perception prominence alters anticipatory slow eye movements. *Experimental Brain Research*, *74*, 555–562.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, *14*, 147–153.

Cha, O., & Chong, S. C. (2014). The background is remapped across saccades. *Experimental Brain Research*, *232*, 609–618.

Demeyer, M., De Graef, P., Verfaillie, K., & Wagemans, J. (2010). Perceptual grouping of contour elements survives saccades. *Journal of Vision*, *10*(7): 515, doi:10.1167/10.7.515. [Abstract]

Demeyer, M., De Graef, P., Verfaillie, K., & Wagemans, J. (2011). Perceptual grouping of object contours survives saccades. *PLoS One*, *6*, e21257.

Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2009). Transsaccadic identification of highly similar artificial shapes. *Journal of Vision*, *9*(4):28, 1–14, doi:10.1167/9.4.28. [PubMed] [Article]

Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2010). Parametric integration of visual form across saccades. *Vision Research*, *50*, 1225–1234.

Drewes, J., Zhu, W., Wutz, A., & Melcher, D. (2015). Dense sampling reveals behavioral oscillations in rapid visual categorization. *Scientific Reports*, *5*, 16290.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*(5040), 90–92.

Feldman, J., & Tremoulet, P. D. (2006). Individuation of visual objects over time. *Cognition*, *99*, 131–165.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*, 173–193.

Fracasso, A., Caramazza, A., & Melcher, D. (2010). Continuous perception of motion and shape across saccadic eye movements. *Journal of Vision*, *10*(13): 14, 1–17, doi:10.1167/10.13.14. [PubMed] [Article]

Franz, V. H., & Loftus, G. R. (2012). Standard errors and confidence intervals in within-subjects designs: Generalizing Loftus and Masson (1994) and avoiding the biases of alternative accounts. *Psychonomic Bulletin & Review*, *19*, 395–404.

Freiwald, W. A. (2007). Attention to objects made of features. *Cortex*, *14*, 744–751.

Ganmor, E., Landy, M. S., & Simoncelli, E. P. (2015). Near-optimal integration of orientation information across saccades. *Journal of Vision*, *15*(16):8, 1–12, doi:10.1167/15.16.8. [PubMed] [Article]

Gordon, R. D., Vollmer, S. D., & Frankl, M. L. (2008). Object continuity and the transsaccadic representation of form. *Perception & Psychophysics*, *70*, 667–679.

Gepshtein, S., & Kubovy, M. (2000). The emergence of visual objects in space–time. *Proceedings of the National Academy of Sciences, USA*, *97*, 8186–8191.

Herzog, M. H., & Öğmen, H. (2014). Apparent motion and reference frames. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization*. Oxford, UK: Oxford University Press.

Holcombe, A. O., & Chen, W. Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to < 3 Hz when tracking three. *Journal of Vision*, *13*(1):12, 1–19, doi:10.1167/13.1.12. [PubMed] [Article]

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106.

Irwin, D. E. (1996). Integrating across saccadic eye movements. *Current Directions in Psychological Science*, *5*, 94–100.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. *Varieties of Attention*, *1*, 29–61.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36,* ECVP Abstract Supplement.

Knapen, T., Rolfs, M., Wexler, M., & Cavanagh, P. (2010). The reference frame of the tilt aftereffect. *Journal of Vision*, *10*(1):8, 1–13, doi:10.1167/10.1.8. [PubMed] [Article]

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605–1615.

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579.

Lin, Z. (2013). Object-centered representations support flexible exogenous visual attention across translation and reflection. *Cognition*, *129*, 221–231.

Lin, Z., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87*, 195–211.

Lin, Z., & He, S. (2012). Automatic frame-centered object representation and integration revealed by iconic memory, visual priming, and backward masking. *Journal of Vision*, *12*(11):24, 1–18, doi:10.1167/12.11.24. [PubMed] [Article]

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*, 204–210.

Mathôt, S., & Theeuwes, J. (2011). Visual attention and stability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 516–527.

Mathôt, S., & Theeuwes, J. (2013). A reinvestigation of the reference frame of the tilt-adaptation aftereffect. *Scientific Reports*, *3*, 1152.

Melcher, D. (2005). Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Current Biology*, *15*, 1745–1748.

Melcher, D. (2007). Predictive remapping of visual features precedes saccadic eye movements. *Nature Neuroscience*, *10*, 903–907.

Melcher, D. (2011). Visual stability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 468–475.

Melcher, D., & Colby, C. L. (2008). Trans-saccadic perception. *Trends in Cognitive Sciences*, *12*, 466–473.

Melcher, D., & Fracasso, A. (2012). Remapping of the line motion illusion across eye movements. *Experimental Brain Research*, *218*, 503–514.

Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience*, *6*, 877–881.

Melcher, D., & Morrone, M. C. (2015). Nonretinotopic visual processing in the brain. *Visual Neuroscience*, *32*, E017.

Melcher, D., Papathomas, T. V., & Vidnyánszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron*, *46*, 723–729.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61–64.

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences, USA*, *99*, 4026–4031.

Nishida, S. Y., Watanabe, J., Kuriki, I., & Tokimoto, T. (2007). Human visual system integrates color signals along a motion trajectory. *Current Biology*, *17*, 366–372.

Öğmen, H., & Herzog, M. H. (2010). The geometry of visual perception: Retinotopic and nonretinotopic representations in the human visual system. *Proceedings of the IEEE*, *98*, 479–492.

Ong, W. S., Hooshvar, N., Zhang, M., & Bisley, J. W. (2009). Psychophysical evidence for spatiotopic processing in area MT in a short-term memory for motion task. *Journal of Neurophysiology*, *102*, 2435–2440.

Oostwoud Wijdenes, L., Marshall, L., & Bays, P. (2015). Visual updating across saccades by working memory integration. *Journal of Vision*, *15*(12):785, doi:10.1167/15.12.785. [Abstract]

Otto, T. U., Ögmen, H., & Herzog, M. H. (2006). The flight path of the phoenix—The visible trace of invisible elements in human vision. *Journal of Vision*, *6*(10):7, 1079–1086, doi:10.1167/6.10.7. [PubMed] [Article]

Otto, T. U., Ögmen, H., & Herzog, M. H. (2009).

Feature integration across space, time, and orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1670–1686.

Otto, T. U., Öğmen, H., & Herzog, M. H. (2010). Perceptual learning in a nonretinotopic frame of reference. *Psychological Science*, *21*, 1058–1063.

Parks, T. E. (1965). Post-retinal visual storage. *American Journal of Psychology*, *78*, 145–147.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Petersik, J. T., & Rice, C. M. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the Ternus effect. *Perception*, *35*, 807–821.

Pikler, J. (1917). *Sinnesphysiologische Untersuchungen*. Leipzig, Germany: Barth.

Pooresmaeili, A., Cicchini, G. M., Morrone, M. C., & Burr, D. (2012). "Non-retinotopic processing" in Ternus motion displays modeled by spatiotemporal filters. *Journal of Vision*, *12*(1):10, 1–15, doi:10.1167/12.1.10. [PubMed] [Article]

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Rieger, J. W., Grüschow, M., Heinze, H. J., & Fendrich, R. (2007). The appearance of figures seen through a narrow aperture under free viewing conditions: Effects of spontaneous eye motions. *Journal of Vision*, *7*(6):10, 1–13, doi:10.1167/7.6.10. [PubMed] [Article]

Ro, T., Breitmeyer, B., Burton, P., Singhal, N. S., & Lane, D. (2003). Feedback contributions to visual awareness in human occipital cortex. *Current Biology*, *13*, 1038–1041.

Rock, I., & Brosgole, L. (1964). Grouping based on phenomenal proximity. *Journal of Experimental Psychology*, *67*, 531–538.

Rock, I., & Ebenholtz, S. (1962). Stroboscopic movement based on change of phenomenal rather than retinal location. *American Journal of Psychology*, *75*, 193–207.

Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*, 376–381.

Schor, C. M., Lakshminarayanan, V., & Narayan, V. (1984). Optokinetic and vection responses to apparent motion in man. *Vision Research*, *24*, 1181–1187.

Scott-Samuel, N. E., & Hess, R. F. (2002). Orientation sensitivity in human visual motion processing. *Vision Research*, *42*, 613–620.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, *16*, 4240–4249.

Tas, A. C., Moore, C. M., & Hollingworth, A. (2012). An object-mediated updating account of insensitivity to transsaccadic change. *Journal of Vision*, *12*(11):18, 1–13, doi:10.1167/12.11.18. [PubMed] [Article]

Tas, A. C., Moore, C. M., & Hollingworth, A. (2014). The representation of the saccade target object depends on visual stability. *Visual Cognition*, *22*, 1042–1046.

Ternus, J. (1926). Experimentelle untersuchungen über phänomenale Identität. *Psychologische Forschung*, *7*, 81–136.

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology*, *78*, 1373–1383.

Van Eccelpoel, C., Germeys, F., De Graef, P., & Verfaillie, K. (2008). Coding of identity-diagnostic information in transsaccadic object perception. *Journal of Vision*, *8*(14):29, 1–16, doi:10.1167/8.14.29. [PubMed] [Article]

Wallace, J. M., & Scott-Samuel, N.E. (2007). Spatial versus temporal grouping in a modified Ternus display. *Vision Research*, *47*, 2353–2366.

Walshe, R. C., & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision Research*, *100*, 38–46.

Warnking, J., Dojat, M., Guérin-Dugué, A., Delon-Martin, C., Olympieff, S., Richard, N., … Segebarth, C. (2002). fMRI retinotopic mapping—Step by step. *NeuroImage*, *17*, 1665–1683.

Wittenberg, M., Bremmer, F., & Wachtler, T. (2008). Perceptual evidence for saccadic updating of color stimuli. *Journal of Vision*, *8*(14):9, 1–9, doi:10.1167/8.14.9. [PubMed] [Article]

Wolf, C., & Schütz, A. C. (2015). Trans-saccadic integration of peripheral and foveal feature information is close to optimal. *Journal of Vision*, *15*(16):1, 1–18, doi:10.1167/15.16.1. [PubMed] [Article]

Wutz, A., Caramazza, A., & Melcher, D. (2012). Rapid enumeration within a fraction of a single glance: The role of visible persistence in object individuation capacity. *Visual Cognition*, *20*, 717–732.

Wutz, A., & Melcher, D. (2013). Temporal buffering

and visual capacity: The time course of object formation underlies capacity limits in visual cognition. *Attention, Perception, & Psychophysics*, *75*, 921–933.

Wutz, A., & Melcher, D. (2014). The temporal window of individuation limits visual capacity. *Frontiers in Psychology*, *5*, 952.

Wutz, A., Muschter, E., van Koningsbruggen, M. G., Weisz, N., & Melcher, D. (2016). Temporal integration windows in neural processing and perception aligned to saccadic eye movements. *Current Biology, 26,* 1659–1668.

Wutz, A., Weisz, N., Braun, C., & Melcher, D. (2014). Temporal windows in visual processing:"prestimulus brain state" and "poststimulus phase reset" segregate visual transients on different temporal scales. *Journal of Neuroscience*, *34*, 1554–1565.

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*, 167–174.

Zhang, E., & Li, W. (2010). Perceptual learning beyond retinotopic reference frame. *Proceedings of the National Academy of Sciences*, *USA*, *107*, 15969–15974.

Zimmermann, E., Morrone, M. C., & Burr, D. C. (2013). Spatial position information accumulates steadily over time. *Journal of Neuroscience*, *33*, 18396–18401.

Zimmermann, E., Morrone, M. C., Fink, G. R., & Burr, D. (2013). Spatiotopic neural representations develop slowly across saccades. *Current Biology*, *23*, R193–R194.

Zirnsak, M., Gerhards, R. G., Kiani, R., Lappe, M., & Hamker, F. H. (2011). Anticipatory saccade target processing and the presaccadic transfer of visual features. *Journal of Neuroscience*, *31*, 17887–17891.

Zöllner, F. (1862). Über eine neue Art anorthoskopischer Zerrbilder. *Annalen der Physik und Chemie*, *117*, 477–484.