

Extrastriate Visual Areas Integrate Form Features over Space and Time to Construct Representations of Stationary and Rigidly Rotating Objects

J. Daniel McCarthy¹, Peter J. Kohler², Peter U. Tse³, and Gideon Paul Caplovitz⁴

Abstract

■ When an object moves behind a bush, for example, its visible fragments are revealed at different times and locations across the visual field. Nonetheless, a whole moving object is perceived. Unlike traditional modal and amodal completion mechanisms known to support spatial form integration when all parts of a stimulus are simultaneously visible, relatively little is known about the neural substrates of the spatiotemporal form integration (STFI) processes involved in generating coherent object representations from a succession visible fragments. We use fMRI to identify brain regions involved in two mechanisms supporting the representation of stationary and rigidly rotating objects whose form features are shown in succession: STFI and position updating. STFI allows past and present form cues to be integrated over space and time into a coherent object even when the object is not visible in any given frame. STFI

can occur whether or not the object is moving. Position updating allows us to perceive a moving object, whether rigidly rotating or translating, even when its form features are revealed at different times and locations in space. Our results suggest that STFI is mediated by visual regions beyond V1 and V2. Moreover, although widespread cortical activation has been observed for other motion percepts derived solely from form-based analyses [Tse, P. U. Neural correlates of transformational apparent motion. *Neuroimage*, 31, 766–773, 2006; Krekelberg, B., Vatakis, A., & Kourtzi, Z. Implied motion from form in the human visual cortex. *Journal of Neurophysiology*, 94, 4373–4386, 2005], increased responses for the position updating that lead to rigidly rotating object representations were only observed in visual areas KO and possibly hMT+, indicating that this is a distinct and highly specialized type of processing. ■

INTRODUCTION

We rarely experience objects in their entirety. Surfaces of objects are generally only partially visible, whether because of occlusion by other objects, self-occlusion, or shadows, shading, and other lighting conditions. Under imperfect real-world circumstances, our experience of coherent object representations relies on neural mechanisms that integrate form information across the visual scene. However, because we are mobile beings and objects around us are often in motion among other objects, the locations and times at which surface fragments become visible to us is typically in flux. Behavioral research has demonstrated that perceptual representations of such objects, whether they are stationary (Kojo, Liinasuo, & Rovamo, 1993), translating (Kellman & Shipley, 1991), or rigidly rotating (McCarthy, Strother, & Caplovitz, under review), can be formed on the basis of the integration of local form information across space and time. Despite the importance of such spatiotemporal integration processes, relatively little is known about the neural circuitry that supports them. Here, we use fMRI to identify brain regions involved in the representation of stationary and

rigidly rotating objects whose form features are revealed over time. Specifically, we identify neural correlates of two interrelated yet distinct mechanisms that underlie spatiotemporal form processing in the context of stationary and rigidly rotating objects: spatiotemporal form integration (STFI) and position updating.

To generate a perceptual representation of a stationary object whose local form features might appear disparately and in succession, the visual system cannot simply rely on spatial integration mechanisms that operate on simultaneously visible form information. Instead, the visual system must maintain local form representations that persist when the items are no longer visible, so that they can be integrated with subsequently visible sources of information—STFI (McCarthy et al., under review; Kojo et al., 1993). Although the mechanisms underlying spatial form integration, including processes such as modal and amodal completion, are relatively well understood (Pastor-Bernier, Tremblay, & Cisek, 2012; Murray, Wylie, et al., 2002; Mendola, Dale, Fischl, Liu, & Tootell, 1999; Pasupathy & Connor, 1999; De Weerd, Desimone, & Ungerleider, 1996; Merigan, 1996; Peterhans & von der Heydt, 1989; von der Heydt & Peterhans, 1989; Peterhans, Von der Heydt, & Baumgartner, 1986; von der Heydt, Peterhans, & Baumgartner, 1984), many questions about

¹Brown University, ²Stanford University, ³Dartmouth College, ⁴University of Nevada

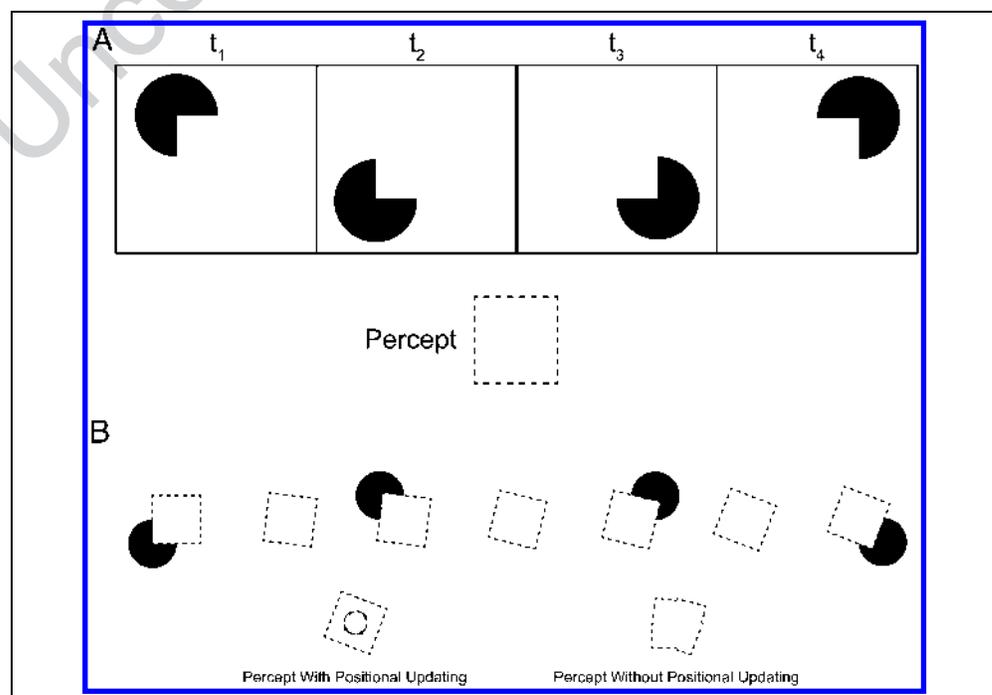
the neural processes that support STFI remain unanswered. For example, research on illusory contours—the paradigmatic case of spatial form integration—suggests that visual regions as early as V2 (and perhaps V1; see Grosf, Shapley, & Hawken, 1993), along with later visual areas including V4v and the lateral occipital complex (LOC) contribute to illusory contour formation. In addition, the right fusiform gyrus, lingual gyrus, posterior parietal cortex, and OFC have all been implicated in illusory contour perception, though there is less consistency across studies regarding the involvement of these higher-order areas (for reviews, see Murray & Herrmann, 2013; Seghier & Vuilleumier, 2006). It is unclear, however, if mechanisms within these regions are capable of spatially integrating disparate local form cues over time. The first goal of the current study is to address this question and identify the neural substrates of STFI.

Although STFI can give rise to a representation of a stationary object whose form features are revealed over time, this process alone is insufficient if the object is in motion. In the case of a rotating object, STFI alone will lead to a stationary deformed figure (Figure 1B, bottom right) whose integrated shape is inconsistent with the object's veridical shape (McCarthy et al., under review; Kellman & Shipley, 1991). To represent the true shape of the object, the positions of previously visible form information need to be updated and matched with currently visible form cues, such that the resulting percept will be of a rigidly moving object—position updating¹ (Palmer, Kellman, & Shipley, 2006). Here we focus on the special case of motion of a rigidly rotating object. Specifically, if the angular displacement is small enough that features can be matched across successive inducers, this change is interpreted as motion

such that a rigidly rotating square is perceived (McCarthy et al., under review). Indeed, there is a wealth of behavioral evidence demonstrating that shapes of translating and rigidly rotating objects can be perceived (McCarthy et al., under review; Agaoglu, Herzog, & Ogmen, 2012; Otto, Ogmen, & Herzog, 2009; Palmer et al., 2006; Kellman & Shipley, 1991; Helmholtz, 1867/1925). In some of these instances, the analysis of local motion information likely plays an important role in generating such percepts (Palmer & Kellman, 2014; Agaoglu et al., 2012; Otto et al., 2009; Palmer et al., 2006; Kellman & Shipley, 1991; Kellman & Cohen, 1984; Helmholtz, 1867/1925). However, at least for the case illustrated in Figure 1B, such percepts can only be generated by strictly form-based analyses that rely on position updating. This is because there is an absence of rigid rotation-correlated motion information in the stimulus.

In general, relatively little is known about how such form-generated motion percepts are instantiated in the brain. Research on transformational apparent motion (TAM), a phenomenon in which motion percepts are also generated solely from form-based analyses, implicates important roles for relatively high levels of visual processing: V3A, V4v, LOC, and hMT+ as well as the ventral subregions of early visual areas V1v, V2v, and V3v (Tse, 2006).² A similar observation was drawn from research on dynamic Glass patterns (Krekelberg, Vatakis, & Kourtzi, 2005; Glass, 1969) that suggests that motion constructed on the basis of local form cues is represented throughout visual cortex, including V1, V2, V3, V3A, V3B/KO, V4V, hMT+, and the LOC. It remains unknown whether position updating has a similar widespread representation in visual cortex or is instead mediated by specialized regions. Thus, the second goal of this article is to identify neural correlates

Figure 1. Objects whose form features are revealed over time. (A) Spatiotemporal integration of form: Sequentially presented inducers are spatiotemporally integrated to generate the percept of an illusory square. (B) The role of position updating: Inducers are presented sequentially, and the occluding square is rotated between each successive inducer. If the position of the features defined by the inducers are updated and spatiotemporally integrated, they will lead to the percept of a rigidly rotating rectangle (left). Without position updating, a misaligned figure would be perceived (right).



of the position updating process that allows us to perceive a rigidly rotating object whose disparate local form features are revealed in succession.

We propose two alternative hypotheses for each of the two primary goals of this article. For STFI, (1) one possibility is that the well-characterized neural networks supporting spatial integration are also capable of maintaining and integrating form information across time. According to this hypothesis, we would expect higher-level ventral stream areas such as V4v and the LOC as well as early visual areas such as V2 (and possibly V1) to exhibit differential responses to surfaces whose form features are revealed over time in comparison to control stimuli containing no figure. However, (2) the persistent nature of the form information that gets integrated in STFI is not easily addressed by the response properties of neurons in early visual cortex (Peterhans & von der Heydt, 1989; von der Heydt & Peterhans, 1989; Peterhans et al., 1986; von der Heydt et al., 1984; Hubel & Wiesel, 1968), so it may instead be the case that such spatiotemporal integration is mediated exclusively by higher-level visual areas beyond V1 and V2.

With respect to position updating, (1) one hypothesis is that it is instantiated within the same networks that support STFI—that is, both mechanisms are part of the same process. This would be the case, for example, if stationary figures represented a lower bound of position updating in which no updating was necessary. If so, we would expect the same areas involved in STFI to be involved in position updating as well. Ultimately, however, position updating leads to a percept of a rigidly rotating object. Accordingly, (2) it may instead be the case that areas of visual cortex specialized to process motion information are also involved in representing object motion derived from the position updating and spatiotemporal integration of form features revealed over time. For instance, it is well established that hMT+ is a central hub of motion perception. Although it receives its primary input from motion selective neurons in V1 and is known to integrate low-level motion energy (Rust, Mante, Simoncelli, & Movshon, 2006; Pack & Born, 2001), hMT+ has also been shown to play a role in supporting a wide range of motion percepts including those that arise from form-based analyses without the integration of local motion signals (Day & Palomares, 2014; Tse, 2006; Krekelberg et al., 2005). Therefore, it may be that hMT+, and perhaps other areas not generally involved in STFI, are involved in representing rigidly rotating object percepts supported by position updating.

We applied fMRI to localize brain activity correlated with percepts of stationary and rigidly rotating objects whose form features are revealed over time. Specifically, we sought to identify regions of visual cortex that support STFI as well as regions involved in position updating. We employed a standard block design to contrast the difference in BOLD signal activation between (A) conditions in which sequentially presented inducers are rotated inward

to generate the percept of an occluding square surface compared to when they are rotated outward so that no illusory square is present as well as (B) conditions in which the occluding square remains stationary throughout the trial compared to when it is rigidly rotating. The current results address a fundamental gap in the literature concerning the neural representation of surfaces and objects whose form features are revealed over time. This is an essential step toward a more complete understanding of surface formation and motion perception under natural viewing conditions.

METHODS

Participants

Eleven (two female) observers participated in the experiment. All observers were right-handed and reported normal or corrected-to-normal vision. Each observer gave written informed consent before participating in the experiment according to the guidelines of the Institutional Review Board and Department of Psychological and Brain Sciences at Dartmouth College and was paid \$20 per 1-hr scanning session.

Apparatus and Display

The stimulus computer was a 2.6-GHz MacBook Pro with an NVIDIA GeForce 8600M GT graphics processor (256 MB of DDR3 SDRAM). Stimuli were generated and presented using the Psychophysics Toolbox (Brainard, 1997) for MATLAB (Mathworks, Inc., Natick, MA). Stimuli were projected from a Panasonic DT-4000U DLP projector (60-Hz refresh rate) onto a frosted Plexiglas screen outside the bore of the magnet and viewed by way of a tangent mirror inside the magnet that permitted a maximum of $22.7^\circ \times 17^\circ$ visual area.

MRI Apparatus and Scanning Procedures

Continuous whole-brain BOLD signals were acquired at the Dartmouth Brain Imaging Center on a Phillips 3T Achieva Intera scanner using a 32-channel head coil. Functional images were obtained using T2* fast field echo, echo-planar functional images sensitive to BOLD contrast (repetition time [TR] = 2 sec, echo time = 35 msec, 32 axial slices, 3.0 mm^2 , matrix size = 80×80 , 3.5 mm thickness, interleaved slice acquisition, 0.5 mm gap, field of view = 240×240 , flip angle = 90° ; sense factor of 2). Dummy scans were collected for a minimum of 8 sec at the beginning of every scan to ensure baseline measures did not include transient activity from the initiation of the scans. High-resolution scans (T1-weighted 3-D MPRAGE sequence) were collected at the end of each scanning session and were used for anatomical reconstruction of the functional data.

fMRI Preprocessing

Functional fMRI data were analyzed using the AFNI software package (Cox, 1996), FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999), and MATLAB (MathWorks, Inc., Natick, MA). Functional images were motion-corrected to the image acquired closest in time to the anatomical scan. Runs in which head movement exceeded voxel size (3 mm) were excluded from the analysis. For each run, each voxel was temporally normalized so that the sum-of-squares was equal to one using AFNI's 3dDetrend command. We performed two group level analyses of the data: (1) a whole-brain analysis in which the data from each participant were smoothed with a 6-mm Gaussian kernel and spatially normalized into the standardized Talairach space (Talairach & Tournoux, 1988) and (2) a ROI-based analysis, where data were analyzed in the native space of each participant, without any smoothing or spatial normalization, and compared across participants based on functionally defined ROIs.

Defining ROIs and Retinotopic Mapping

ROIs were defined using localizer scans collected in separate scanning sessions. Functionally defined regions of visual cortex were independently identified for each of the 11 observers who participated in the main experiment. For each participant, the cortical mesh used to define the ROIs was coregistered to the experimental data using AFNI (Cox, 1996) and SUMA (Saad, Reynolds, Argall, Japee, & Cox, 2004). Functional mapping was conducted using procedures described previously (Caplovitz & Tse, 2010; Slotnick & Yantis, 2003; Sereno et al., 1995) and is only described briefly here.

Polar angle representation in visual cortex was measured using two wedges of an 8-Hz flickering black and white polar checkerboard grating bilaterally opposite like a bowtie to enhance the signal-to-noise ratio (Slotnick & Yantis, 2003; Sereno et al., 1995). Each wedge subtended 22.5° of 360° and occupied a given location in the visual field for one TR (2 sec) before moving to the adjacent location in the clockwise or counterclockwise direction (direction was alternated across runs). Cortical representation of eccentricity was measured using expanding 8-Hz flickering concentric rings that each spanned ~1° of visual angle in ring width. For every TR, a given ring was replaced by its outward neighbor, with the exception that the outermost ring was replaced by the innermost ring at the end of a cycle. This process was repeated until the end of the run. For each participant, five runs of each direction were collected for the wedge stimulus and three runs were collected for the concentric rings. Retinotopic areas (V1d, V1v, V2d, V2v, V3d, V3v, V4v, and V3AB) were defined as masks based on standard criteria (Sereno et al., 1995): contralateral quadrant representation for V1d, V1v, V2d, V2v, V3d, and V3v and contralateral hemifield representation for V4v and V3AB (Tootell et al.,

1997). Masks for visual areas V1, V2, and V3 were created by merging the dorsally and ventrally defined regions of each respective area.

Individual hMT+ and masks were made following procedures outlined by Tyler, Likova, Kontsevich, and Wade (2006) and detailed in Caplovitz and Tse (2010). hMT+ masks were identified by contrasting BOLD signal activity generated by stimulus blocks containing 64 randomly positioned expanding and contracting white dots (radius = 0.45°) with blocks in which the 64 dots remained stationary. In blocks containing motion, the dots switched between expansion and contraction every 1.5 sec. Three runs of the localizer task were collected for 7 of the 11 participants. For the localizer task, masks were identified for both hemispheres as an isolated cluster of voxels in which greater BOLD signal was observed for the motion condition compared to the stationary condition ($p < 10^{-8}$, uncorrected).

The kinetic occipital (KO) area has been identified from fMRI studies as a cortical region located between retinotopic area V3AB and hMT+ that is highly sensitive to kinetic or motion-defined borders (Van Oostende, Sunaert, Van Hecke, Marchal, & Orban, 1997) and depth structure (Tyler et al., 2006). Individual KO masks were defined using a standard kinetic contour localizer containing contours defined by motion (Van Oostende et al., 1997), detailed in Caplovitz and Tse (2010). BOLD signal activation in the kinetic condition was contrasted to a null condition containing a full field of random luminance-defined noise moving coherently back and forth. These two conditions were alternately presented for blocks of 20 sec. Observers completed a minimum of eight runs containing four blocks of each condition. Masks were identified as a cluster of voxels in which greater BOLD signal was observed for the kinetic condition compared to the uniform condition ($p < .0001$, uncorrected). The KO localizer was completed by 4 of the 11 participants.

The construction of individual LOC masks followed standard procedures outlined in Kourtzi and Kanwisher (2000) and detailed in Caplovitz and Tse (2010). Object images (7° × 7°) were displayed on a white background with a black grid superimposed. The same images were scrambled within the same grid to generate control images. Both conditions contained a central fixation point at all times. The center of the images was updated every TR to a random position within 1° of fixation to prevent perceptual fading. Stimuli were presented in 20-sec blocks and separated by blank periods containing only the fixation point. Participants each completed three runs containing four stimulus blocks (two scrambled, two unscrambled). LOC masks were defined for each hemisphere as a cluster of voxels in which the BOLD response was greater for unscrambled objects compared to scrambled objects ($p < .05$, uncorrected). Distinct, bilateral activations corresponding to the LOC were identified in all participants.

Finally, the boundaries of individual ROIs are defined one at a time, and voxels that lie along a given boundary can overlap with other visual areas. This can result in incorrectly assigning shared voxels to both areas on either side of the boundary. Because there is no objective way of assigning such voxels to one visual area or the other, the intersection of any two masks was removed from both masks. Potential overlapping voxels in the following visual areas were removed from the masks: V1d–V2d, V1v–V2v, V2d–V3d, V2v–V3v, V3v–V4v, V4v–LOC, V3d–KO, V3d–V3A, V3A–KO, KO–LOC, KO–hMT+, LOC–hMT+. Although this resulted in discarding voxels that were localized in retinotopic and other localized ROIs, thereby lowering statistical power, it is a conservative step because it further ensures that BOLD signal responses measured in a given ROI arose solely from that one area.

Experimental Stimuli

The stimuli for the main experiment consisted of “pac-men” style inducers with a diameter of 5.67° visual angle that were presented sequentially in random order and centered on locations consistent with four corners of a square. At the beginning of each stimulus block, the inducers were rotated either inward or 180° outward (Figure 2, Demonstration Videos 1–4).³ In the inward conditions, the corners of the square were located $5.67^\circ \times 5.67^\circ$ visual angle ($H \times V$) from the center of the display, generating the percept of an illusory square. In the outward conditions, the inducers were rotated

180° according to their center of gravity $\left[\frac{2r \left(\sin \frac{\pi}{2} \right)}{\frac{3\pi}{2}} \right]$ so

that no illusory square was present and the centroids of the inward and outward inducers were in the same location in both conditions ($6.30^\circ \times 6.30^\circ$ from the center of the display). This was done so that the empty space between the inducers was roughly equivalent in the inward and outward conditions to ensure roughly equivalent foveal stimulation. In all conditions, each inducer was presented for 250 msec with a 0-msec ISI, resulting in each cycle of four inducers lasting a total of 1000 msec. Within a given cycle, the order of the four possible locations at which each inducer could be presented was randomized.

Inducers were presented on a gray background and switched polarity between black and white after 125 msec during each presentation to minimize the formation of afterimages. In the stationary conditions, the occluding square surface maintained a constant angle throughout each block. In the motion conditions, the angle of the illusory square rotated back and forth, oscillating between -20° and 20° from vertical at an angular velocity of 12° per second. This was accomplished by updating the angle of the illusory square by 3° between successive inducers (as illustrated in Figure 1B). This rotational velocity was chosen based on psychophysical experiments demonstrating that robust percepts of a rigidly rotating square can be generated at this speed. As can be observed in Demonstration Video 5, increasing the angular velocity destroys the percept of a rigidly rotating square and instead gives rise to the percept of a stationary, deformed object. Importantly, this demonstration illustrates that the perceived shapes are not “inferred” from single inducers but are rather constructed through STFI and position updating.

In any given block, the inducers were rotated inward or outward and the illusory square either maintained a constant angle or oscillated throughout the block. Thus, in the inward conditions, the successively presented inducers generated the percept of a stationary or moving illusory square. In the outward conditions, no illusory square was perceived, and the stimulus appeared to be four outwardly facing pac-men whose “mouths” either remained constant or changed throughout the block.

Stimulus Protocol

In every run, each of the four stimulus conditions was pseudorandomly presented twice in 16-sec blocks (8 TRs) separated by blank periods of 12 sec (6 TRs) containing only the fixation spot. Each run began and ended with a blank period. Eight runs were collected for each participant. A central fixation square was presented throughout every run of the experiment and briefly changed color (to red or green) 78 times per run. Participants were instructed to covertly attend the stimulus while maintaining central fixation on the square and respond to color changes by pressing a button corresponding to

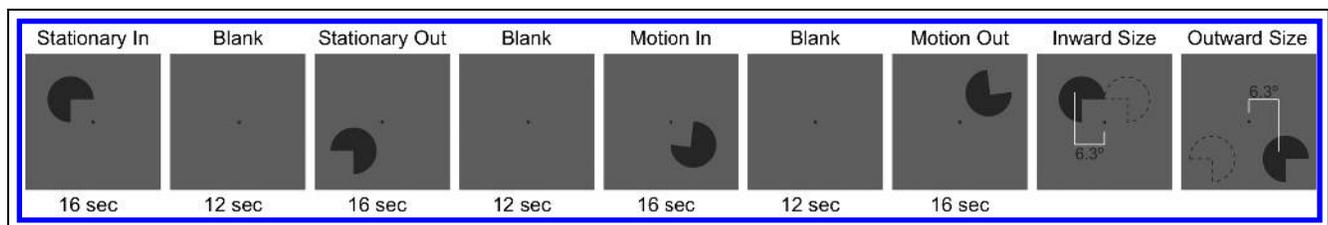


Figure 2. fMRI block design paradigm. Blocks of sequentially presented inducers were shown. The inducer orientations were either rotated inward, consistent with a global figure, or rotated outward according to their center of gravity (illustrated by the point at which end of the white line intersects the inducers) so that no illusory figure was present. In the static conditions, the inducer orientation remained constant throughout the block, and in the motion condition, the inducer orientation changed between each successive presentation.

each color. This assured that participants were alert, fixating, and maintaining a relatively constant level of attention across the different experimental conditions. The percentage of fixation color changes that were responded to within 1.5 sec of the change on each run was computed. Responses that took longer than this cutoff time were counted as misses. Runs with lower than 75% accuracy were rejected from further analysis.

Analyses

General linear model (GLM) analyses were performed on each voxel to obtain beta weights (coefficients) for boxcar predictors convolved with a canonical BOLD response to the four experimental stimulus blocks for each participant. The GLM included nine regressors of noninterest: three run-wise baseline parameters corresponding to constant signal, linear drift, and second-degree polynomials as well as six regressors derived from the motion registration parameters.

We performed two separate GLM analyses. (1) In the ROI-based GLM, no spatial smoothing or normalization was performed, and a threshold was applied so that only the most strongly activated voxels (i.e., $F > 6$) within each predefined region were included in the analysis. The average beta weight for each condition across each ROI was then computed for each participant. After verifying that there were no statistically significant activation differences between hemispheres, beta weights within left- and right-hemisphere ROIs were combined, thereby treating the small hemispheric differences in the number of voxels for a given area as a random variable. (2) For the whole-brain GLM, the data were spatially smoothed and normalized as described above before performing the GLM, allowing group level analysis within individual voxels.

For each ROI (in the ROI-based GLM) and voxel (in the whole-brain GLM), a group level 2×2 repeated-measures ANOVA with factors of Configuration (figure vs. non-figure) and Inducer Set (stationary vs. rotating) was then applied to the beta weights to identify regions in which differences in BOLD responses to each of the experimental conditions could be observed. Of particular interest were clusters of voxels and ROIs that exhibited a main effect of Configuration or an interaction between Configuration and Inducer Set. For the whole-brain GLM, the analysis was restricted to clusters exceeding a minimum size of 32 contiguous voxels at a threshold of $p < .005$ uncorrected. This cluster size was chosen based on Monte Carlo simulations using AFNI's 3dClustSim command that, given the level of spatial smoothing, indicated clusters exceeding this size would be unlikely to occur due to chance ($p < .01$). As illustrated by the hypothetical data in Figure 3A, a main effect of Configuration would suggest that neurons within that brain region support the integration of persistent form representations into a unified figure. In other words, STFI can be taken to involve ROIs that show a main effect

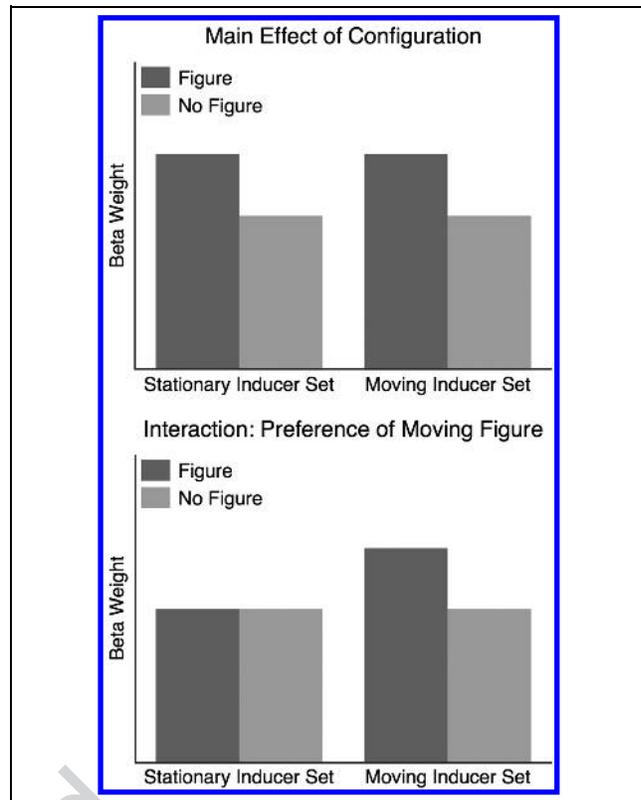


Figure 3. Hypothetical results. (A) A main effect of configuration would suggest that a given area supports STFI. (B) An interaction with a preferential response to the moving figure would suggest that a given area supports position updating.

of Configuration. An interaction, specifically one in which a preferential response was observed in the rotating figure condition, would suggest a role in the position updating of persistent form information that supports the rigid rotation percept (Figure 3B).

RESULTS

Fixation

On average, participants detected 95.62% ($SEM = 0.95\%$) of fixation changes. No participants performed below the 75% cutoff threshold. Because of technical complications with the scanner, two runs were rejected for one participant. In addition, due to excessive head movement, one and three runs from two additional participants, respectively, were rejected from further analysis.

ROI Analysis

Figure 4 illustrates the average beta weights for each condition in each of the ROIs. Of particular interest to our goal of identifying neural correlates of the integration of persistent local form information, we observed a significant main effect of Configuration in several ROIs: V3 ($F(1, 10) = 57.85, p < .001, \eta^2 = 0.85$), V3AB ($F(1, 10) = 17.40, p = .002, \eta^2 = 0.64$), V4v ($F(1, 10) = 22.60, p = .001,$

$\eta^2 = 0.69$), LOC ($F(1, 10) = 33.07, p < .001, \eta^2 = 0.77$), and KO ($F(1, 3) = 13.34, p = .035, \eta^2 = 0.82$). In each of these areas, the BOLD response was greater when the inducers formed a figure than when they did not, regardless of whether the figure was perceived as stationary or moving. There was no significant main effect of figure in V1, V2, or hMT+. This lack of V1/V2 involvement is consistent with the idea that STFI is instantiated in higher-order visual areas commonly associated with processing global form

than those supporting low-level spatial integration processes and suggests an interesting functional distinction between V3 and earlier visual areas.

Additionally, as expected from the whole-brain analysis, a main effect of Inducer Set (stationary or moving) emerged in V2 ($F(1, 10) = 5.15, p = .047, \eta^2 = 0.34$), V3 ($F(1, 10) = 17.49, p = .002, \eta^2 = 0.64$), V3AB ($F(1, 10) = 20.54, p = .001, \eta^2 = 0.67$), V4v ($F(1, 10) = 37.54, p < .001, \eta^2 = 0.79$), LOC ($F(1, 10) = 29.89, p < .001,$

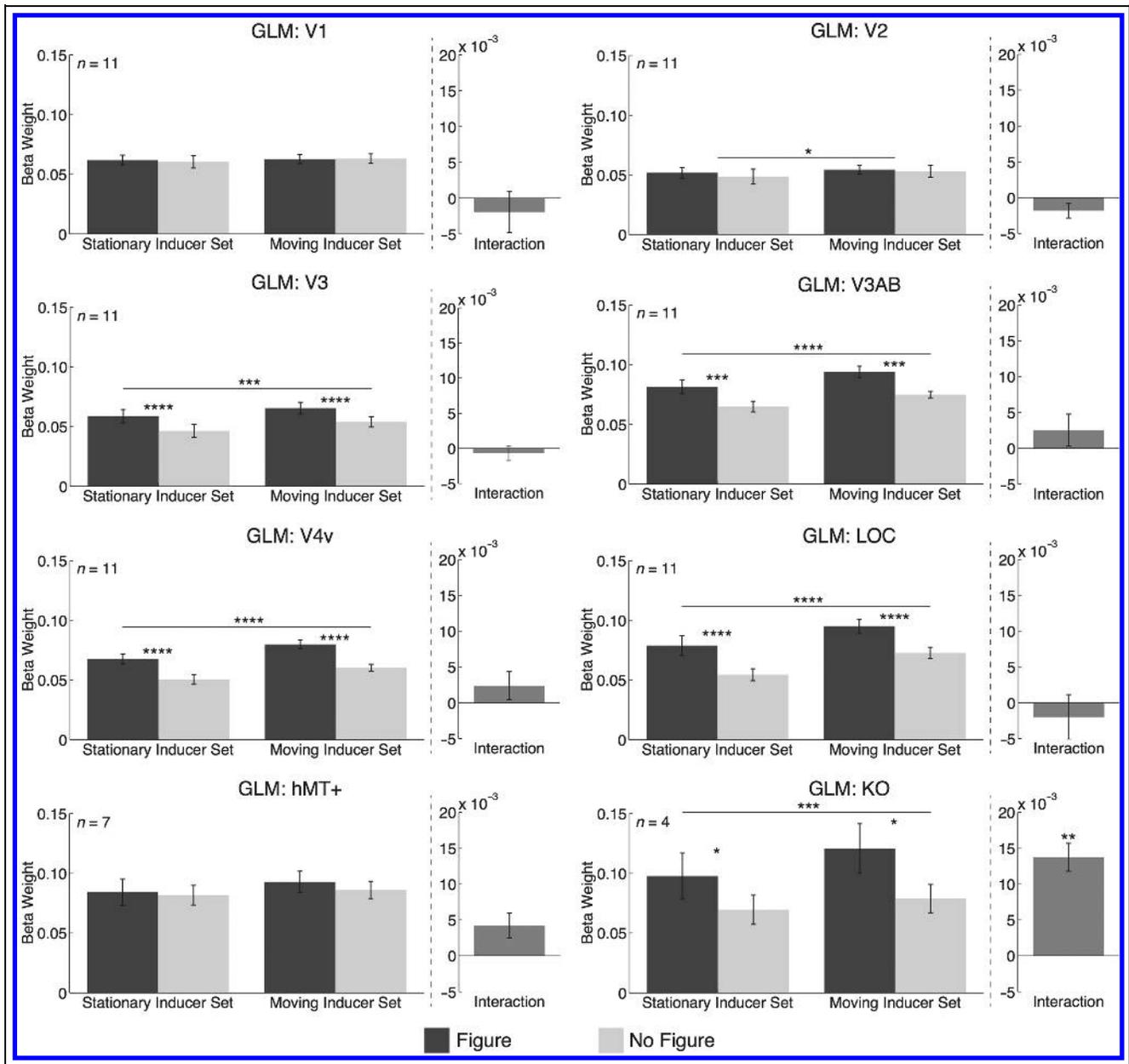


Figure 4. (A) Beta weights averaged across participants for each condition in each ROI. No significant differences emerged in V1. Main effects of Inducer Set with larger BOLD responses to the moving figure emerged in all other visual areas shown with the exception of hMT+. However, an inspection of the time course of the BOLD signal suggests that lower responses to the static figure were likely due to adaptation. Importantly, a main effect of Configuration was found in V3, V3AB, V4v, LOC, and KO. Each of these areas preferentially responded to the presence of a figure, suggesting that they play a role in STFI. Moreover, an interaction with a larger response to the rigidly rotating figure emerged in KO and possibly hMT+, consistent with the view that the neural correlates of position updating occur in these ROIs. The interaction bar was computed by the following formula: (Motion Figure – Static Figure) – (Motion Nonfigure – Static Nonfigure). Error bars indicate ± 1 SEM. Black asterisks indicate significance at the following levels: * $p < .05$, ** $p < .01$, *** $p < .005$, **** $p < .001$.

$\eta^2 = 0.75$), and KO ($F(1, 3) = 94.36, p < .0023, \eta^2 = 0.97$). In each of these areas, the BOLD signal response was larger for the rotation inducer set compared to the stationary inducer set, regardless of whether or not they formed a figure. However, because this analysis collapses across both figure and nonfigure responses, this effect is likely not indicative of a specialized role in representing surfaces, but rather of inherent differences between the two stimulus sets. Specifically, in the stationary inducer sets, the same inducers are presented at the same locations in the visual field throughout the stimulus block. In contrast, the shapes of individual inducers in the rotation inducer sets are always changing. Therefore, the differential responses likely reflect differential degrees of local adaptation resulting from repeated presentations of identical stimuli (for reviews, see Grill-Spector & Malach, 2004; Grill-Spector, Kourtzi, & Kanwisher, 2001).

Directly related to our goal of identifying a neural correlate of the position updating of persistent local form information, a significant interaction between figure and inducer set emerged in area KO ($F(1, 3) = 49.99, p = .006, \eta^2 = 0.94$), and the interaction was nearly significant in hMT+ ($F(1, 6) = 5.80, p = .052, \eta^2 = 0.49$). As can be observed in the interaction bars in Figure 4; in both cases, the interaction is such that these areas responded preferentially to the rigidly rotating figure relative to the other conditions. However, we note that the significant interaction observed in KO is based on only the subset of participants ($n = 4$) for whom KO was functionally localized. Using only this subset of participants, we performed a one-way repeated-measures ANOVA with ROI as a factor to determine whether the interaction term reflecting position updating was indeed greater in KO than in the other ROIs. As expected, the main effect of ROI was significant ($F(7, 21) = 6.87, p < .001, \eta^2 = 0.70$). Moreover, as can be observed in Figure 5, the interaction term is greatest in KO. Pairwise t tests revealed that KO was significantly greater than each of the other ROIs (all $p < .04$, uncorrected) with the exception of V3AB ($t(3) = -1.91, ns$); however, this difference was only significant between KO and V1–V3 as well as the LOC at the Bonferroni-corrected alpha level ($\alpha = 0.05/8$ ROIs = 0.00625). Although this analysis is only based on a small sample size, it supports the conclusion that KO appears to play a specialized role in STFI and position updating.

Whole-brain GLM

The repeated-measures ANOVA in AFNI revealed a main effect of Configuration in bilateral clusters of voxels in the lateral occipital cortex (LOC), posterior parietal cortex, and medial portions of the occipital lobe as well as a small cluster in the left insula (Figure 6). For illustrative purposes, the F statistics derived from the ANOVA were converted to t statistics with the sign assigned to illustrate the direction of the effect (Red = Figure > No Figure;

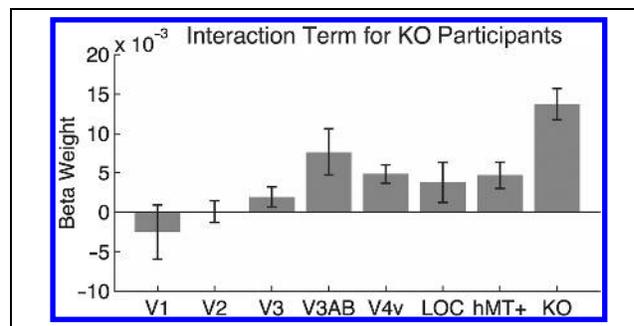


Figure 5. Beta weights for the interaction term averaged across KO participants for each ROI. The interaction bar was computed by the following formula: (Motion Figure – Static Figure) – (Motion Nonfigure – Static Nonfigure). A main effect of ROI emerged with the greatest response in KO. Importantly, the general response trend for this subset of subjects is similar to that observed for all subjects with a slightly elevated response in V3AB, V4v, and the LOC. Although only four subjects completed the KO localizer, the results support the conclusion that KO appears to play a specialized role in STFI and position updating. Error bars indicate ± 1 SEM.

Blue = Figure < No Figure), and the boundaries of a single participant's ROIs are projected onto the surface. The clusters located in the posterior parietal cortex and LOC show increased responses to the conditions in which the figure is formed compared to the no-figure conditions. These clusters are largely consistent with activations commonly observed in studies of object representation in which responses to shapes and objects are contrasted with nonobject or scrambled images (Grill-Spector et al., 2001; Kanwisher, Chun, McDermott, & Ledden, 1996) and the locations of the clusters in the LOC closely match those commonly reported for the LOC. The clusters located along the medial portions of the occipital lobe and the left insula show the opposite effect: increased responses to the no-figure conditions. Taken together, these activations reveal neural correlates of the spatial integration of persistent representations of form that underlie our perception of objects whose form features are revealed over time. The coordinates and results of the ANOVA for each of these clusters are summarized in Table 1.

The ANOVA also revealed a main effect of Inducer Set across a widespread array of areas in which greater activity was observed in response to the rotational conditions compared to the stationary ones (Figure 7). For illustrative purposes, the F statistics were again converted to t statistics with the sign assigned to illustrate the direction of the effect (Red = Motion > Static; Blue = Motion < Static), and the boundaries of a single participant's ROIs are projected onto the surface. As discussed in the results of the ROI analysis, this result likely arises from basic local adaptation and does not inform the goals of the current study.

As stated earlier, we are particularly interested in identifying regions that exhibit a significant interaction with preferential processing of the rotating figure. This is

matched for local features. For instance, areas in retinotopic visual cortex can decode the orientation of gratings due to a global radial bias across these regions (Alink, Krugliak, Walther, & Kriegeskorte, 2013; Freeman, Brouwer, Heeger, & Merriam, 2011; Sasaki et al., 2006). This was also the case using spiral stimuli (Alink et al., 2013) or glass patterns (Mannion, McDonald, & Clifford, 2009) that were balanced for such radial effects. These findings of low-level involvement in global shape processing stand in contrast to our results. Although it is possible that this discrepancy arises due to our use of univariate versus multivariate analyses, we offer the speculative hypothesis that this may be a result of the spatiotemporal nature of our stimuli and suggest that only higher visual areas are capable of spatially integrating form over time.

Evidence supporting the low- and high-level differences observed in our study also comes from work demonstrating that the LOC showed consistent activation in response to faces and objects independent of the image contrast that modulated activity in early visual cortex (Avidan et al., 2002). This has been taken as evidence for a high-level object representation that does not directly depend on the feature level output of early visual areas. Similarly, it has been shown that the formation of global object percepts from the integration of moving local form features inversely modulates activity in early visual cortex and the LOC (Fang, Kersten, & Murray, 2008; Murray, Kersten, Olshausen, Schrater, & Woods, 2002). The conclusion drawn from this observation was that once high-level object representations are formed, inputs arising from early visual areas might be actively inhibited. We found no evidence of a similar inhibition of early visual cortex in our ROI data (i.e., the figure condition does not yield a smaller response than the nonfigure condition in V1 and V2; see Figure 4), but our results do appear to reflect a similar distinction between the types of processing carried out by lower- and higher-level visual areas. Importantly, although all of the studies discussed in this section involve some form of spatial integration, we extend these results by identifying several higher-level visual areas that are capable of integrating form over both time and space.

The whole-brain GLM also revealed an increased response to the figure conditions in posterior parietal cortex. This is consistent with research suggesting a role for this region in surface and boundary formation (Milner, Goodale, & Vingrys, 2006; Tyler et al., 2006; Mendola et al., 1999; Van Oostende et al., 1997; Nakamura, Gattass, Desimone, & Ungerleider, 1993; Rodman & Albright, 1989; Livingstone & Hubel, 1983; Ungerleider & Mishkin, 1982) as well as the representation of objects (Mruczek, von Loga, & Kastner, 2013; Konen & Kastner, 2008; Grill-Spector et al., 2001; Kourtzi & Kanwisher, 2000). The whole-brain analysis also revealed a preference for the figure conditions in several clusters of occipital cortex (see Figure 6 and Table 1), but because those clusters overlap with our functionally defined ROIs to some extent, we limit our discussion of those areas to our ROI analysis.

The whole-brain GLM also revealed significant clusters in which the nonfigure conditions led to a greater response, located in two separate general regions. First, (1) a set of clusters was found in medial occipital regions likely to have some overlap with retinotopically organized cortex. Alternatively, the decreased response to the figure could be a result of competitive interactions, such that neuronal populations representing the illusory figure were enhanced whereas those responding to the inducers were suppressed (Kok & de Lange, 2014). Accordingly, increased activity would be observed for the nonfigure, as no competition would be generated by the presence of an occluding square. However, Kok and de Lange (2014) only investigated competitive interactions within V1, and inspection of the surface maps showing our results in Figure 6 reveals that the preference for nonfigure stimuli is primarily restricted to retinotopic areas V1–V3 in the left hemisphere.

The reduced activity in lower visual areas may also be due to cortical feedback from higher visual areas that represent the figure. Specifically, V1 activity has been shown to decrease when object elements can be grouped into coherent shapes, and this is accompanied by increased activity in the LOC (Murray, Wylie, et al., 2002). It is suggested that when areas such as the LOC can “explain” a visual stimulus, concurrent activity in lower areas decreases because higher-level predictions match and therefore discount the incoming sensory information. Moreover, Likova and Tyler (2008) observed retinotopic suppression of ground regions, but not figure enhancement, in V1 and V2 when their stimuli were consistent with figure-ground organization. Accordingly, the inducer-related activity may have been suppressed in our study when a figure could be spatiotemporally integrated, thereby leading to a reduced response in lower visual areas. Finally, we acknowledge that the stimulus level differences between the figure and nonfigure conditions may be a potential confound. Specifically, because of eccentricity-specific asymmetries, the nonfigure condition may have generated a larger response in retinotopic areas due to the inducers overlapping with what would be the corners of the occluding square in the figure condition. Importantly, although we cannot entirely rule out this possibility for the figural effects observed here, such differences likely do not pose any additional confounds for the case of rigid rotation. Taken together, we conclude that the greater responses to the nonfigure condition in our study can likely be explained by suppressive feedback from higher visual areas representing the integrated figure.

The whole-brain GLM also found a preference for nonfigure stimuli in the left insula, an area not commonly considered to be critical for visual perception, although a small number of recent studies have suggested that the area may be involved in perceptual tasks specific to object processing (Schintu et al., 2014; Volberg & Greenlee, 2014). Most notably, one imaging study found evidence that the left insula plays a role in accumulating evidence

that an object is present in the visual scene (Ploran et al., 2007). Activity in the left insula showed an increase as an object emerged out of noise, reaching an asymptote at the point at which the object became visible. We offer the speculative hypothesis that over the course of a stimulus block, as subsequent inducers are presented, the left insula plays a role in accumulating evidence as to whether or not they represent a figure. In the figure conditions, it presumably takes only a few moments for the figure to be revealed after which it is continuously perceived for the duration of the block. In contrast, in the no-figure condition, evidence is continuously accumulated, albeit in vain, as no figure is ever perceived, thereby leading to increased activity within the left insula during the no-figure condition.

The ROI analysis found interactions with a preferential response to the rigidly rotating figure—the hallmark of position updating—in areas KO and hMT+. Thus, despite the fact that the rotating square contained no net motion energy, our results suggest that hMT+, which is classically considered to be a motion area, and KO support rigidly rotating object percepts based solely on the integration and position updating of local form information revealed over time. These areas have also been implicated in generating other types of motion percepts derived from form-based analyses in the absence of local motion signals (Tse, 2006; Kregelberg et al., 2005). For instance, the involvement of hMT+ has been demonstrated in TAM. In this phenomenon, replacing an object with a different spatially overlapping object can generate motion percepts (Tse, 2006; Tse & Logothetis, 2002). Specifically, observers see the motion of the first object continuously transforming as a sequence of animated scenes rather than two distinct frames each containing a separate object (Tse, 1998). However, activity in V1, V2, V3AB, V4v, and the LOC was also modulated by TAM. Unlike our study, visual area KO was not examined. It has also been reported that both KO/V3B and hMT+ are involved in the representation of implied motion arising from dynamic Glass patterns (Kregelberg et al., 2005). However, as was the case for TAM, several other areas including V1, V2, V3, V3A, V4V, and the LOC are also involved in representing dynamic Glass patterns. In contrast to these examples, increased activation for position updating was only observed in KO and possibly hMT+. This suggests a key difference between position updating and other form-based motion percepts: position updating appears to be a more specialized process, relying on processing in only a few specific areas of cortex.

The whole-brain GLM failed to reveal such an interaction. However, this type of analysis blurs anatomical boundaries, fails to compensate for individual structural variability, and has decreased sensitivity to task-related effects compared to ROI-based analyses (Nieto-Castanon, Ghosh, Tourville, & Guenther, 2003). One possibility is that the interaction between STFI and position updating only occurs in a specific and potentially very small ana-

tomical region. Therefore, any effects that were present within functionally defined regions of individual participants may have been masked when averaging across participants on a standardized anatomical space. It is also possible that the response in the four participants with a functionally defined KO was fundamentally different from the remaining participants. However, inspection of the interaction bars in Figure 4 (all participants) with those in Figure 5 (KO participants) shows that the general trend is similar across both groups with a slightly higher response in areas V3AB, V4v, and the LOC for the KO participants. Importantly, the interaction term in these participants was greatest for KO, lending support to the conclusion that KO appears to play a specialized role in STFI and position updating.

It is important to note that, depending on how KO is defined, it has been suggested to encompass multiple visual areas, showing overlap with areas V3, V3A, V3B, as well as lateral occipital areas 1 (LO1) and 2 (LO2), two retinotopically defined areas that are believed to play distinct roles in extracting boundary and shape information, respectively (Larsson & Heeger, 2006). In defining KO for this study, we explicitly excluded voxels that may have overlapped with V3 and V3AB (for more detail, see Caplovitz & Tse, 2010). It is likely, however, that our functionally defined KO would have some extent of overlap with retinotopically defined LO1 and LO2. But because we did not retinotopically map LO1 and LO2, we cannot know the extent of this overlap.

The fact that KO is observed in both STFI and position updating highlights the notion that, in the context of these stimuli, STFI works in tandem with position updating. Specifically, perception of the globally rotating object depends on both STFI and position updating. KO was originally identified as being selectively responsive to kinetically defined contours (Van Oostende et al., 1997); however, such stimuli also convey depth structure, with each side of the kinetic contour appearing at different depths. Subsequent research has demonstrated that, rather than being specialized for the representation of just kinetic contours, KO plays a more generalized role in encoding depth structure (Tyler et al., 2006). Through the implied occlusion of the inducers, the formation of the STFI figure conveys a segmentation of the stimuli in depth. It is possible that KO plays a key role in maintaining the depth relationships implied by STFI, relationships that are critical for position updating in the case of rigid rotation. In contrast, activity in hMT+ increases only in response to the rotating figure condition. Together, this raises an intriguing hypothesis that the combined interaction of STFI and position updating in KO generates a motion signal that feeds into hMT+, which in turn generates the percept of continuous rigid rotational motion. Interestingly, KO is in a good anatomical position to integrate information between the dorsal and ventral processing streams and support the integration and updating of form information across time.

Conclusions

As we look around our visual environment, many of the objects we see are partially and dynamically occluded. STFI allows the perception of global object shape under these difficult conditions by supporting persistent representations of form features that are revealed over time. When the object itself is in motion, accurate global shape perception also requires position updating of these persistent representations before their integration. We have investigated the neural basis of STFI and position updating using fMRI. Our results show that STFI is mediated by several visual cortical regions, all beyond area V2. Position updating, on the other hand, appears to be restricted to regions KO and possibly hMT+, suggesting that this is a more specialized process, relying on neural mechanisms that are distinct from those supporting STFI. These results address a fundamental gap in the literature and provide an essential step toward a more complete understanding of surface formation and motion perception under natural viewing conditions.

Acknowledgments

This work was supported in part by grants awarded to G. P. C. from the National Institute of Health: NIGMS 5P20GM103650-02 and NEI 1R15EY022775.

Reprint requests should be sent to Gideon Paul Caplovitz, Department of Psychology, University of Nevada, Reno, 1664 N. Virginia St., Reno, NV 89557, or via e-mail: gcaplovitz@unr.edu.

Notes

1. We note that, although this study only investigated the role of position updating as it supports rigidly rotating object representations, this process is also critical for perceiving unified objects when they translate behind an occluding surface. Because of stimulus constraints in the current study, only rigid rotation was investigated here; however, we believe that the current findings have implications for perceiving moving objects under circumstances of occlusion in general.
2. Importantly, however, the stimulus in Tse (2006) was presented 0.3° above fixation. The ventral subregions of areas V1, V2, and V3 process the upper hemifield, and it is possible that a similar pattern of activation would have been observed in dorsal subregions had the stimulus been present in the lower part of the display instead.
3. We encourage readers to watch all Demonstration Videos in loop mode. The Demonstration Videos can be found at www.mitpressjournals.org/doi/suppl/10.1162/JOCN_a_00850.

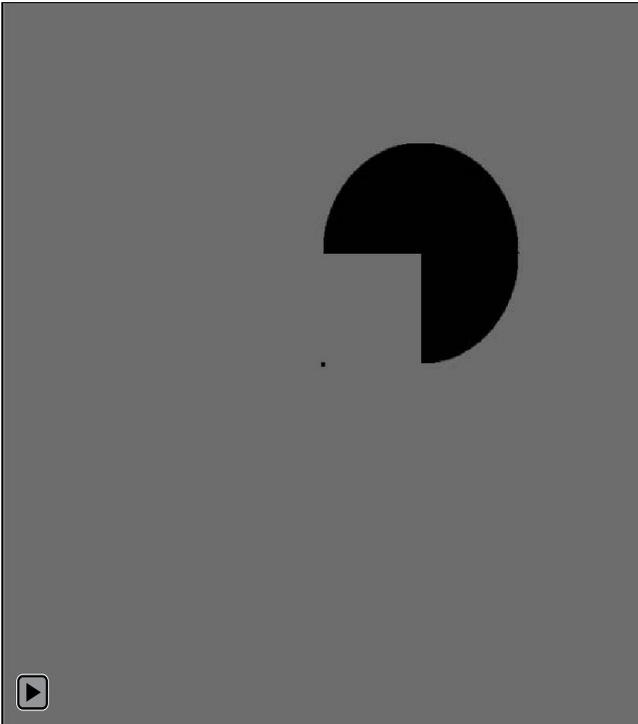
REFERENCES

- Agaoglu, M. N., Herzog, M. H., & Ogmen, 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.
- Alink, A., Krugliak, A., Walther, A., & Kriegeskorte, N. (2013). fMRI orientation decoding in V1 does not require global maps or globally coherent orientation stimuli. *Frontiers in Psychology*, *4*.
- Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., & Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *Journal of Neurophysiology*, *87*, 3102–3116.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Caplovitz, G. P., & Tse, P. U. (2010). Extrastriate cortical activity reflects segmentation of motion into independent sources. *Neuropsychologia*, *48*, 2699–2708.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, *9*, 179–194.
- Day, A. M., & Palomares, M. (2014). How temporal frequency affects global form coherence in Glass patterns. *Vision Research*, *95*, 18–22.
- De Weerd, P., Desimone, R., & Ungerleider, L. G. (1996). Cue-dependent deficits in grating orientation discrimination after V4 lesions in macaques. *Visual Neuroscience*, *13*, 529–538.
- Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*, 1–9.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, *9*, 195–207.
- Freeman, J., Brouwer, G. J., Heeger, D. J., & Merriam, E. P. (2011). Orientation decoding depends on maps, not columns. *Journal of Neuroscience*, *31*, 4792–4804.
- Glass, L. (1969). Moire effect from random dots. *Nature*, *223*, 578–580.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677.
- Grosf, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal “illusory” contours. *Nature*, *365*, 550–552.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*, 632–635.
- Helmholtz, H. (1925). *Treatise on physiological optics* (3rd ed., Vol. 3). New York: Dover Press. (Original work published in 1867)
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*, 215–243.
- Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research*, *5*, 55–67.
- Kellman, P. J., & Cohen, M. H. (1984). Kinetic subjective contours. *Perception & Psychophysics*, *35*, 237–244.
- Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.
- Kojo, I., Liinasuo, M., & Rovamo, J. (1993). Spatial and temporal properties of illusory figures. *Vision Research*, *33*, 897–901.
- Kok, P., & de Lange, F. P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology*, *24*, 1531–1535.
- Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience*, *11*, 224–231.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *The Journal of Neuroscience*, *20*, 3310–3318.

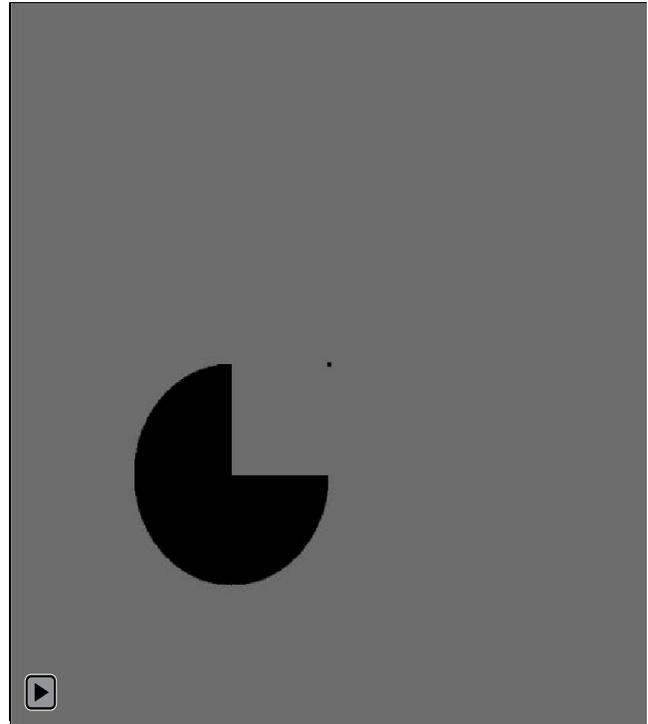
- Krekelberg, B., Vatakis, A., & Kourtzi, Z. (2005). Implied motion from form in the human visual cortex. *Journal of Neurophysiology*, *94*, 4373–4386.
- Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, *26*, 13128–13142.
- Likova, L. T., & Tyler, C. W. (2008). Occipital network for figure/ground organization. *Experimental Brain Research*, *189*, 257–267.
- Livingstone, M. S., & Hubel, D. H. (1983). Specificity of cortico-cortical connections in monkey visual system. *Nature*, *304*, 531–534.
- Mannion, D. J., McDonald, J. S., & Clifford, C. W. (2009). Discrimination of the local orientation structure of spiral Glass patterns early in human visual cortex. *NeuroImage*, *46*, 511–515.
- McCarthy, J. D., Strother, L., & Caplovitz, G. P. (under review). Spatiotemporal form integration: Sequentially presented inducers can lead to representations of stationary and rigidly rotating objects.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *The Journal of Neuroscience*, *19*, 8560–8572.
- Merigan, W. H. (1996). Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. *Visual Neuroscience*, *13*, 51–60.
- Milner, A. D., Goodale, M. A., & Vingrys, A. J. (2006). *The visual brain in action* (Vol. 2). Oxford, UK: Oxford University Press.
- Mruczek, R. E., von Loga, I. S., & Kastner, S. (2013). The representation of tool and non-tool object information in the human intraparietal sulcus. *Journal of Neurophysiology*, *109*, 2883–2896.
- Murray, M. M., & Herrmann, C. S. (2013). Illusory contours: A window onto the neurophysiology of constructing perception. *Trends in Cognitive Sciences*, *17*, 471–481.
- Murray, M. M., Wylie, G. R., Higgins, B. A., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). The spatiotemporal dynamics of illusory contour processing: Combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *The Journal of Neuroscience*, *22*, 5055–5073.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, *99*, 15164–15169.
- Nakamura, H., Gattass, R., Desimone, R., & Ungerleider, L. G. (1993). The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *The Journal of Neuroscience*, *13*, 3681–3691.
- Nieto-Castanon, A., Ghosh, S. S., Tourville, J. A., & Guenther, F. H. (2003). Region of interest based analysis of functional imaging data. *NeuroImage*, *19*, 1303–1316.
- Ostwald, D., Lam, J. M., Li, S., & Kourtzi, Z. (2008). Neural coding of global form in the human visual cortex. *Journal of Neurophysiology*, *99*, 2456–2469.
- Otto, T. U., Ogmen, H., & Herzog, M. H. (2009). Feature integration across space, time, and orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1670–1686.
- Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, *409*, 1040–1042.
- Palmer, E. M., & Kellman, P. J. (2014). The aperture capture illusion: Misperceived forms in dynamic occlusion displays. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 502–524.
- Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General*, *135*, 513–541.
- Pastor-Bernier, A., Tremblay, E., & Cisek, P. (2012). Dorsal premotor cortex is involved in switching motor plans. *Frontiers in Neuroimaging*, *5*, 5.
- Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*, 2490–2502.
- Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *The Journal of Neuroscience*, *9*, 1749–1763.
- Peterhans, E., Von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. *Visual Neuroscience*, *3*, 343–351.
- Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *The Journal of Neuroscience*, *27*, 11912–11924.
- Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, *75*, 53–64.
- Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, *9*, 1421–1431.
- Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., & Cox, R. W. (2004). *SUMA: An interface for surface-based intra- and inter-subject analysis with AFNI*. Paper presented at the IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004.
- Sasaki, Y., Rajimehr, R., Kim, B. W., Ekstrom, L. B., Vanduffel, W., & Tootell, R. B. (2006). The radial bias: A different slant on visual orientation sensitivity in human and nonhuman primates. *Neuron*, *51*, 661–670.
- Schintu, S., Hadj-Bouziane, F., Monte, O. D., Knutson, K. M., Pardini, M., Wassermann, E. M., et al. (2014). Object and space perception—Is it a matter of hemisphere? *Cortex*, *57*, 244–253.
- Seghier, M. L., & Vuilleumier, P. (2006). Functional neuroimaging findings on the human perception of illusory contours. *Neuroscience and Biobehavioral Reviews*, *30*, 595–612.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.
- Slotnick, S. D., & Yantis, S. (2003). Efficient acquisition of human retinotopic maps. *Human Brain Mapping*, *18*, 22–29.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Stuttgart, Germany: Thieme.
- Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., et al. (1997). Functional analysis of V3A and related areas in human visual cortex. *The Journal of Neuroscience*, *17*, 7060–7078.
- Tse, P. U. (1998). The role of parsing in high-level motion processing. In T. Watanabe (Ed.), *High-level motion processing: Computational, neurobiological, and psychophysical perspectives* (pp. 249–266). Cambridge, MA: MIT Press.
- Tse, P. U. (2006). Neural correlates of transformational apparent motion. *NeuroImage*, *31*, 766–773.
- Tse, P. U., & Logothetis, N. K. (2002). The duration of 3-d form analysis in transformational apparent motion. *Perception & Psychophysics*, *64*, 244–265.

- Tyler, C. W., Likova, L. T., Kontsevich, L. L., & Wade, A. R. (2006). The specificity of cortical region KO to depth structure. *Neuroimage*, *30*, 228–238.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. A. Goodale, & R. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1997). The kinetic occipital (KO) region in man: An fMRI study. *Cerebral Cortex*, *7*, 690–701.
- Volberg, G., & Volberg, M. W. (2014). Brain networks supporting perceptual grouping and contour selection. *Frontiers in Psychology*, *5*.
- von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *The Journal of Neuroscience*, *9*, 1731–1748.
- von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260–1262.

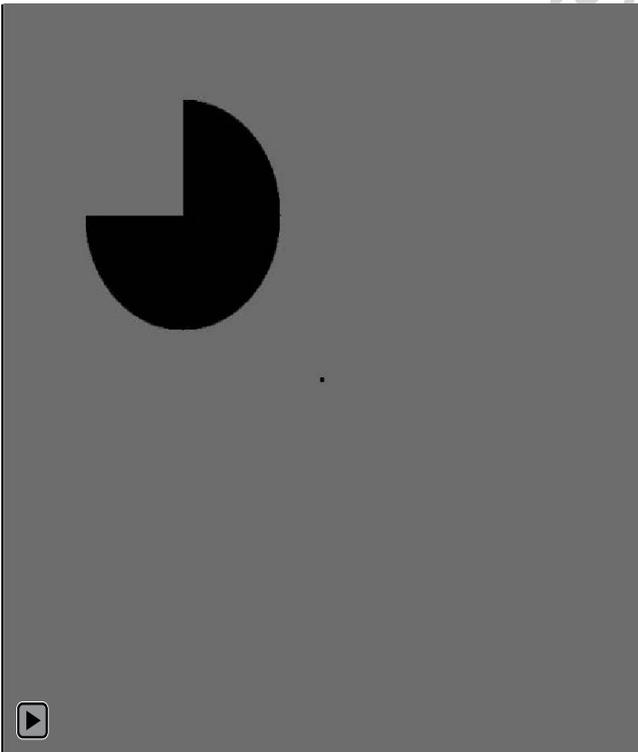
Uncorrected Proof



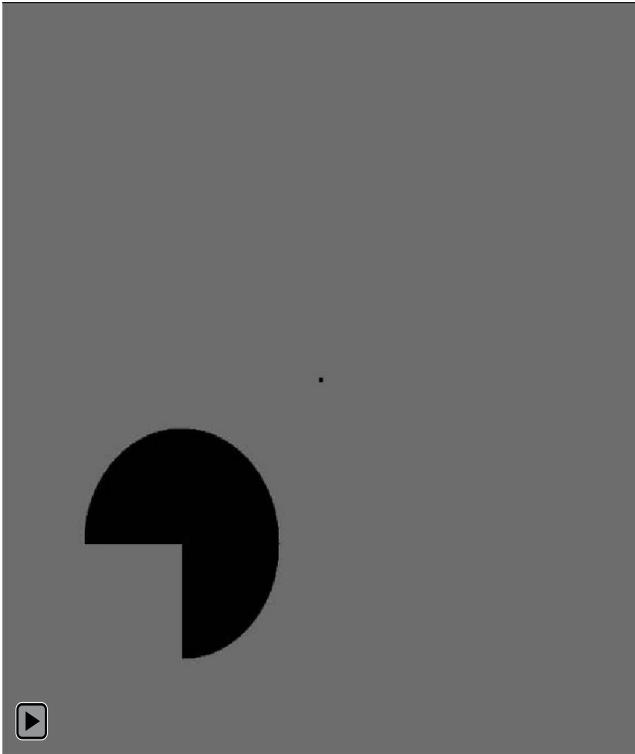
Demonstration Video 1. Static inward stimulus demonstrating STFI. Sequentially presented pac-men inducers give rise to the percept of an illusory square.



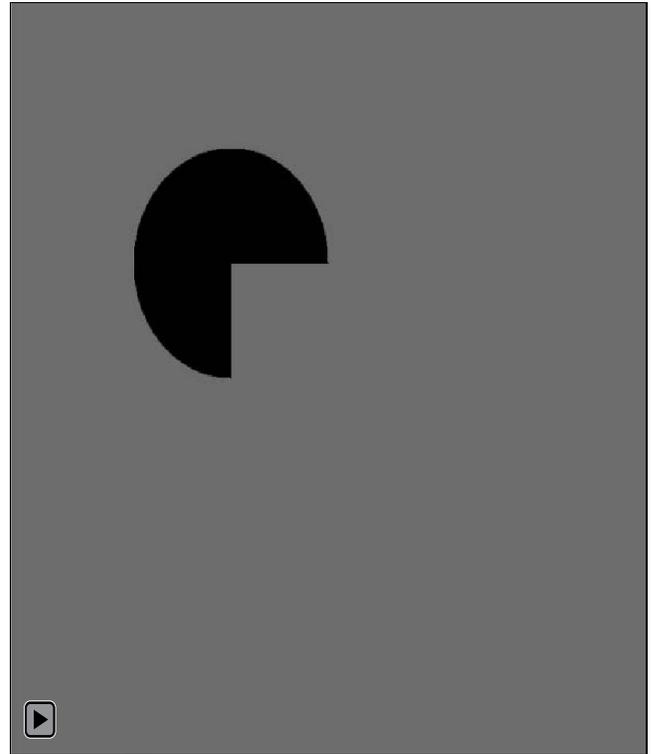
Demonstration Video 3. Moving inward stimulus demonstrating STFI and Position Updating. The occluding square is rotated between each successive inducer. The explicit features defined by the inducers are spatiotemporally integrated and positionally updated to generate the percept of a rigidly rotating rectangle. This video illustrates a 3° rotation between each inducer oscillating between $\pm 20^\circ$



Demonstration Video 2. Static outward control stimulus. Inducers were rotated 180° outward so that no illusory square was present.



Demonstration Video 4. Moving outward control stimulus. The occluding 'square' is rotated 3° between each successive inducer and oscillates between $\pm 20^\circ$. Inducers were rotated 180° outward so that the same local information is available in each inducer is identical to the inward case but a globally rotating square is not present.



Demonstration Video 5. Moving inward condition with an increased rotational velocity. Increasing the angle of rotation between inducers leads a breakdown of position updating and a stationary deformed figure is perceived. This video illustrates a 12° rotation between each inducer oscillating between $\pm 20^\circ$.

Uncorrected