

# Spatiotemporal Form Integration: Sequentially presented inducers can lead to representations of stationary and rigidly rotating objects

J. Daniel McCarthy<sup>1</sup> · Lars Strother<sup>2</sup> · Gideon Paul Caplovitz<sup>2</sup>

Published online: 13 August 2015  
© The Psychonomic Society, Inc. 2015

**Abstract** Objects in the world often are occluded and in motion. The visible fragments of such objects are revealed at different times and locations in space. To form coherent representations of the surfaces of these objects, the visual system must integrate local form information over space *and* time. We introduce a new illusion in which a rigidly rotating square is perceived on the basis of sequentially presented Pacman inducers. The illusion highlights two fundamental processes that allow us to perceive objects whose form features are revealed over time: Spatiotemporal Form Integration (STFI) and Position Updating. STFI refers to the spatial integration of persistent representations of local form features across time. Position updating of these persistent form representations allows them to be integrated into a rigid global motion percept. We describe three psychophysical experiments designed to identify spatial and temporal constraints that underlie these two processes and a fourth experiment that extends these findings to more ecologically valid stimuli. Our results indicate that although STFI can occur across relatively long delays between successive inducers (i.e., greater than 500 ms), position updating is limited to a more restricted temporal window (i.e., ~300 ms or less), and to a confined range of spatial (mis)alignment. These findings lend insight into the limits of

mechanisms underlying the visual system's capacity to integrate transient, piecemeal form information, and support coherent object representations in the ever-changing environment.

**Keywords** Form perception · Surface perception · Motion perception · Form-motion interactions · Illusory contours

## Introduction

Our ability to see the surfaces of objects relies on mechanisms that integrate form information across the visual field (i.e., spatial integration). Because they so elegantly demonstrate the process of surface formation, Pacman-induced illusory contours (Fig. 1A; Kanizsa, 1955, 1979; Kellman & Shipley, 1991) have been used widely to investigate how spatial integration processes allow us to perceive unified surfaces of objects even when they are not fully visible. In addition, there is evidence that spatial integration processes also can operate across time (i.e., spatiotemporal integration). For instance, similar surface percepts can be formed through the sequential presentation of Pacman inducers (Fig. 1B; Demonstration Video 1; Kojo, Liinasuo, & Rovamo, 1993).<sup>1</sup> This example highlights mechanisms that integrate local form information over both space and time. It also has been demonstrated that these spatiotemporal integration processes can support percepts of translating objects (Kellman & Shipley, 1991). In these cases, however, not only does local form information have to be integrated over space and time, but the position of previously viewed form information has to be updated before the integration (Kellman & Shipley, 1991; Palmer, Kellman, & Shipley, 2006). We used a novel variant

---

**Electronic supplementary material** The online version of this article (doi:10.3758/s13414-015-0967-5) contains supplementary material, which is available to authorized users.

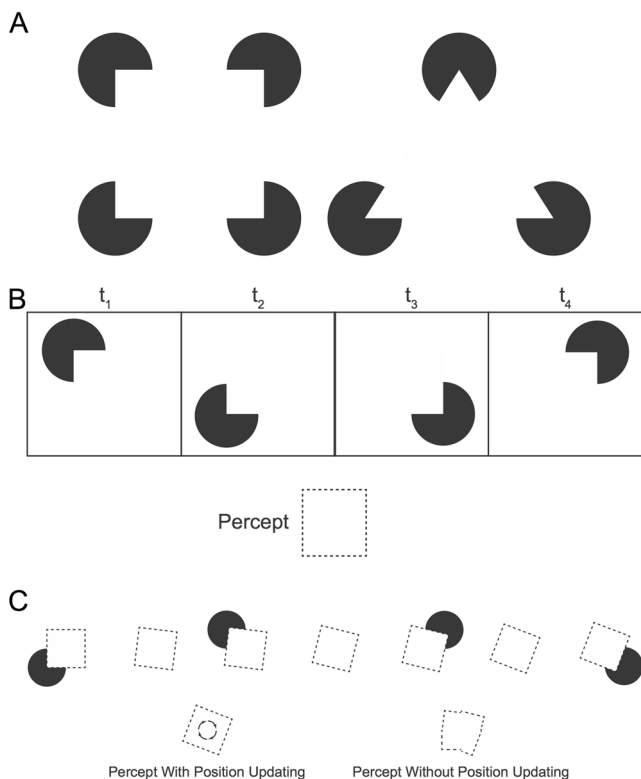
---

✉ Gideon Paul Caplovitz  
gcaplovitz@unr.edu

<sup>1</sup> Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, RI, USA

<sup>2</sup> Department of Psychology, University of Nevada, Reno, Reno, NV, USA

<sup>1</sup> We encourage readers to watch all videos in loop mode.



**Fig. 1** Examples of spatial integration and Spatiotemporal Form Integration (STFI). **A** Examples of a Kanizsa square and triangle (Kanizsa, 1955, 1979). Local form information provided by the inducers is spatially integrated to generate the percept of an occluding shape. **B** Sequentially presented inducers are spatiotemporally integrated to generate the percept of an illusory square. **C** STFI and Position Updating supporting motion perception. Inducers are presented sequentially and the occluding square is rotated between each successive inducer. The explicit features defined by the inducers are accumulated and spatiotemporally integrated and position updated to generate the percept of a rigidly rotating square. The above example illustrates a  $7^\circ$  clockwise rotation between each inducer

of these stimuli to demonstrate that spatiotemporal form integration (STFI) and position updating can support percepts of rigidly rotating surfaces.

As illustrated in the bottom left of Fig. 1C, sequential inducers can be presented such that they are consistent with a rigidly rotating square. Importantly, each inducer itself is static and no changes in configuration occur while they are visible. The stimulus can be conceptualized as a square with the same color as the background that transiently rotates during the inter-inducer-interval and remains stationary when each inducer is present. The inducers can be thought of as filled circles that become partially occluded by the square. Critically, if all four successive inducers were to be presented simultaneously, they would not represent the shape of a square. Instead, because of the rotation that occurs between inducers, they would reveal an irregular, misaligned polygon (Fig. 1C, bottom right). However,

if the inter-inducer-interval is brief and the angular velocity is not too great, a rigidly rotating square can be perceived (Demonstration Video 2).

The rigid-rotation percept highlights the delicate interplay between STFI and position updating. Why might position updating occur in the first place? After all STFI alone could represent a surface, albeit an irregular misaligned one. As can be seen in Demonstration Video 3, if the angular velocity is too great, that is precisely what is perceived. In this case, STFI allows an integrated figure to be formed, yet the increased misalignment of the edge information provided by the inducers prevents position updating from occurring. This demonstration is important, because it illustrates that in the case of rigid rotation the perceived global shape is not being “inferred” from single inducers but rather is constructed from the local information through STFI and position updating. Moreover, if the inter-inducer-interval is too great, neither STFI nor position updating occur and no global figure is perceived (Demonstration Video 4).

The goal of this study was to elucidate the spatial and temporal ranges over which STFI and position updating occur as well as the nature of the local form information that is integrated. We also investigated STFI and position updating using more ecologically valid stimuli and examined the potential role prior knowledge may play in representing spatiotemporally integrated surfaces. To identify basic-level spatial and temporal constraints underlying STFI and position updating (Experiments 1-3), we used simple cases of stationary and rigidly rotating squares like those shown in Demonstration Videos 1 and 2. To examine the external validity of these findings and potential contributions of prior knowledge in STFI and position updating, we used upright and inverted rigidly rotating silhouettes of recognizable animals (Experiment 4).

## General methods

### Participants

Ten observers (6 females; mean age = 25) participated in all three psychophysical experiments. In addition, ten observers (4 females; mean age = 25.8; 5 original participants) participated in the control experiment in experiment 1. All observers were naïve to the underlying aims of the experiments, reported normal or corrected-to-normal vision, and received course credit for participating. Before participating, each observer provided informed consent according to the guidelines of the Department of Psychology and the institutional review board of the University of Nevada, Reno.

## Apparatus and display

Stimuli were presented on a Dell Trinitron P991 monitor (19 inches,  $1024 \times 768$ ) with an 85-Hz refresh rate. The stimulus computer was a 2.4-GHz Mac Mini with an NVIDIA GeForce 320M graphics processor (256 MB of DDR2 SDRAM). Stimuli were created and presented with the Psychophysics Toolbox (Brainard, 1997) for MATLAB (Mathworks Inc., Natick, MA). Stimuli were presented on a gray background. Participants placed their head in a chin rest and viewed the stimuli binocularly from a distance of 57 cm.

## Experiment 1. Stationary STFI figures are nearly as robust as other classic illusory figures

This experiment was designed to determine the durations over which form information can be maintained and integrated to represent an occluding surface. We applied adaptive staircase procedures to compare the representations of STFI completed figures to standard illusory figures using a wide range of timing parameters. Specifically, participants made orientation judgments on induced illusory rectangles in which the inducers were presented simultaneously as well as rectangular surfaces generated by STFI in which the inducers were presented sequentially.

### Stimuli and procedures

As illustrated in Fig. 2A, on any given trial the stimuli consisted of four Pacman inducers with a diameter of  $3^\circ$  visual angle. The inducers were positioned so that their inner contours were consistent with the corners and edges of a horizontally or vertically oriented rectangle. As such, the stimuli can be thought of as consisting of four circular inducers partially occluded by a rectangle sharing the same color as the background. Each trial began with a central fixation point ( $0.1^\circ$  visual angle). Participants were instructed to maintain fixation and refrain from moving their eyes while the stimuli were on the screen. The fixation point was visible for  $\sim 500$  ms (43 frames at 85 Hz) at which point it disappeared for  $\sim 500$  ms (43 frames at 85 Hz) before the onset of the stimulus. The fixation point remained off for the duration of the trial until participants had indicated the orientation of the occluding rectangle via a button press. The absence of a fixation point during the trial was used to prevent participants from using a local central cue to make their judgment.

Adaptive staircases were used to determine the minimum orientation threshold participants were able to distinguish. At the beginning of each staircase, the size of the occluding rectangle started out at  $3^\circ \times 4.5^\circ$  corresponding to an aspect ratio of 1.5. If the participant correctly reported the orientation of the rectangle, the aspect ratio of the rectangle was reduced

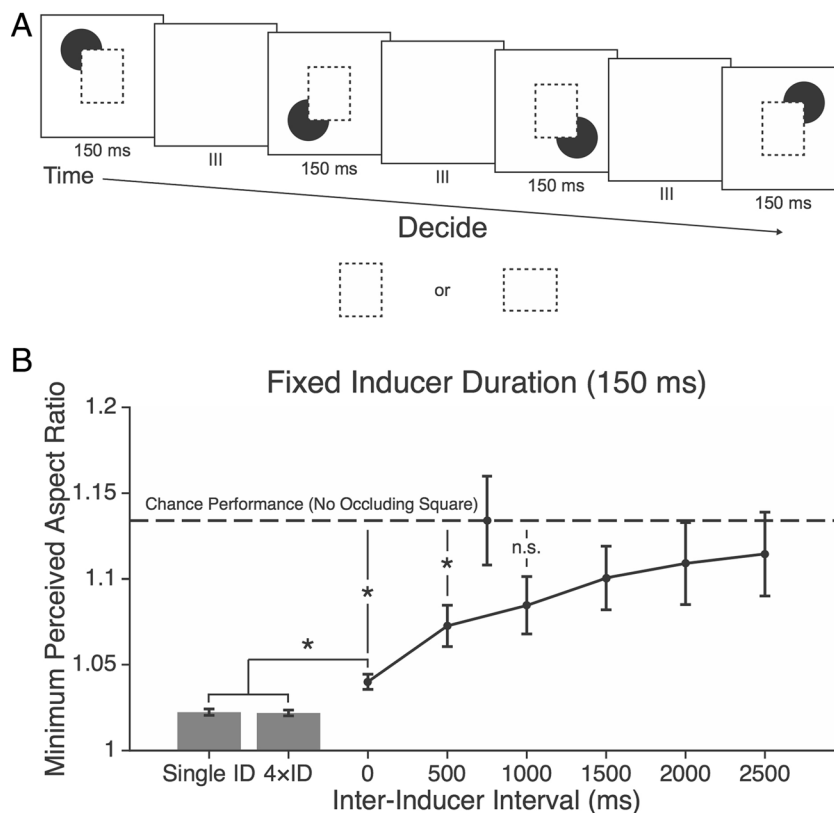
making it closer to a square. Subsequent aspect ratios were drawn from the following list: 1.4, 1.35, 1.3, 1.25, 1.2, 1.15, 1.1, 1.05, 1.025, and 1.0125, corresponding to occluding rectangle sizes of  $3^\circ \times 4.2^\circ$ ,  $3^\circ \times 4.05^\circ$ ,  $3^\circ \times 3.9^\circ$ ,  $3^\circ \times 3.75^\circ$ ,  $3^\circ \times 3.6^\circ$ ,  $3^\circ \times 3.45^\circ$ ,  $3^\circ \times 3.3^\circ$ ,  $3^\circ \times 3.15^\circ$ ,  $3^\circ \times 3.075^\circ$ ,  $3^\circ \times 3.0375^\circ$ , respectively.

In the sequential STFI trials, each inducer was presented for  $\sim 150$  ms (13 frames, 85 Hz) in pseudorandom order until each of the four inducers was presented, resulting in the inducers being present on the screen for a total of 600 ms (51 frames, 85 Hz). This inducer duration (ID) was chosen based on previous work demonstrating that the strength of illusory contours begins to asymptote approximately  $\sim 120$  ms (Guttman & Kellman, 2004) and modulate the N170 ERP component peaking between  $\sim 150$ – $200$  ms (Murray & Herrmann, 2013). Because we sought to compare figures completed by STFI to classic illusory figures as described below, we reasoned this should provide ample time for illusory contours to form in the control stimuli. Thus, we chose a  $\sim 150$  ms ID so that each inducer remained on the screen for a duration that closely mirrors the time course of illusory contour formation.

To minimize the formation of afterimages, the inducers alternated once between black and white so that roughly half of the time ( $\sim 150$  ms ID:  $\sim 75$  ms;  $\sim 60$  ms ID:  $\sim 30$  ms) it was black ( $\sim 150$  ms ID: 7 frames, 85 Hz;  $\sim 60$  ms ID: 3 frames, 85 Hz) and the other half of the time it was white ( $\sim 150$  ms ID: 6 frames, 85 Hz;  $\sim 60$  ms ID: 2 frames, 85 Hz). The black-white order of each inducer was randomly determined on each trial. If participants responded correctly, the aspect ratio of the occluding rectangle was reduced; if the participant responded in error, the aspect ratio on the subsequent trial would be increased. The staircase procedure ended when five reversals were recorded.

For the sequential conditions, the inter-inducer-interval (III) used for each staircase was chosen from the following list: 0 ms,  $\sim 500$  ms (43 frames, 85 Hz), 1000 ms (85 frames, 85 Hz),  $\sim 1500$  (128 frames, 85 Hz) ms, 2000 ms (170 frames, 85 Hz),  $\sim 2500$  ms (213 frames, 85 Hz). This corresponded to individual trial durations of 600 ms (51 frames, 85 Hz),  $\sim 2100$  ms (179 frames, 85 Hz), 3600 ms (306 frames, 85 Hz),  $\sim 5100$  ms (434 frames, 85 Hz), 6600 ms (553 frames, 85 Hz), and  $\sim 8100$  ms (689 frames, 85 Hz), respectively. The procedures for the simultaneous conditions were identical to the above with the exception that all inducers were presented at the same time for a total of  $\sim 150$  ms (single inducer duration) or 600 ms (total inducer duration). In total, each observer completed four staircases of each of the eight distinct conditions (six different IIIs and two simultaneous durations) in pseudorandom order.

To ensure that participants were basing their judgments on the global shape of the induced rectangle, its orientation was randomly determined on each trial and its center randomly



**Fig. 2** Sample trial and the results for experiment 1. **A** Inducers are presented sequentially in random order for ~150 ms with variable IIIs and observers report the orientation of the occluding rectangle via a key press (this example shows a vertical rectangle with an aspect ratio of 1.25 as indicated by the black dotted rectangle). The orientation of the occluding rectangle was randomly determined on each trial. **B** Results of experiment 1. Gray bars indicate the average orientation discrimination across participants for simultaneous trials in which all four inducers are presented at once for the duration of a single inducer (~150 ms) or the

duration of four sequential inducers (600 ms). The points on the solid black line indicate the average orientation discrimination across participants for the sequential condition with various IIIs. The dashed line indicates chance performance across participants in which they made orientation judgments when no occluding rectangle was present in the display. Lines with asterisks indicate main effects of trial type (simultaneous vs. sequential; left) and IIIs at which participants performed significantly better than chance (right). Error bars indicate standard error of the mean

jittered within a 1.5° visual angle radius surrounding the center of the screen. In addition, the centers of the inducers were randomly jittered along the diagonal axis according to the corner of the occluding rectangle with a maximum displacement of 1° visual angle inward or outward (1/3 of the total inducer diameter). Despite these efforts, when the aspect ratio is high, it is possible that observers may be able to use information about the relative positions of the inducers themselves rather than the form-specific information about the occluding rectangle. To further address this concern and create a baseline against which the results derived from the main experiment could be compared, we also included a condition to compute a “chance performance” level. Specifically, each participant completed four sequential-condition staircases with an inducer duration of ~150 ms and a 0 ms III following the exact procedures described above, except that no occluding rectangle was present in the display. In this case, only four large circles appeared and the distance between them was the only information available to make an orientation judgment.

**Analyses**

For each staircase, the aspect ratios at which each of the last four reversals occurred were averaged together. For each participant, the threshold at which the orientation task could be performed was operationally defined as the average of this value across the four repetitions of the staircase for each condition. Thus, for each participant, eight distinct threshold values were computed—one for each of the six IIIs in the sequential conditions and one for each of the two simultaneous conditions.

First, we performed one-sample *t* tests to determine whether the derived thresholds represented better than chance performance (as determined in the experiment in which no occluding square was present). A 1 × 6 one-way repeated measures ANOVA with inter-inducer-interval as the factor was performed on the sequential conditions. An *a priori* linear contrast also was performed on these data to determine whether any main effect of inter-inducer-interval is parametric in nature (i.e., longer intervals lead to increased thresholds). Paired *t* tests were performed

to compare the thresholds derived at specific IIIs in the sequential conditions to that induced in the simultaneous conditions as well as to compare the two simultaneous conditions

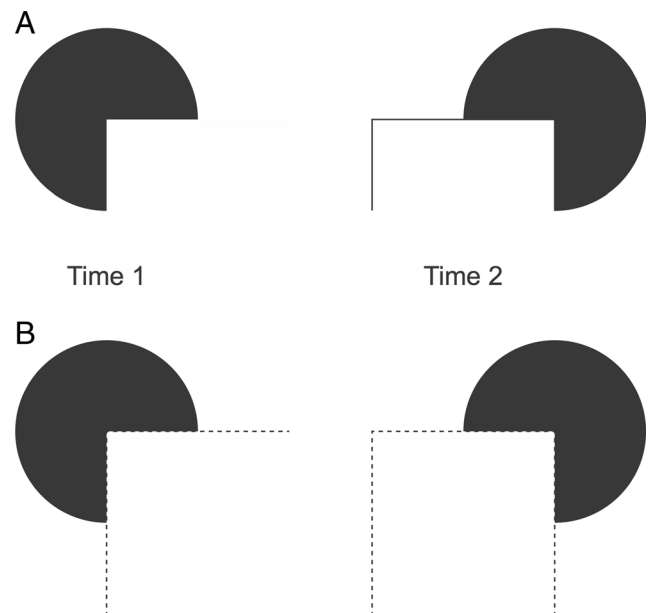
## Results

Figure 2B illustrates the mean thresholds averaged across observers for each condition along with the measured chance performance. As illustrated, participants performed best when the inducers were presented simultaneously compared with when they were presented sequentially. This was true even when the inter-inducer-interval was 0 ms, which had the lowest threshold of the sequential conditions (Single ID:  $t(9) = -3.93$ ,  $p = 0.003$ ;  $4 \times$  ID:  $t(9) = -3.80$ ,  $p = 0.004$ ). No difference was observed between the two simultaneous conditions ( $t(9) = 0.52$ , n.s.). In the sequential conditions, performance thresholds systematically increased as III increased (Main Effect of III:  $F(5,45) = 6.11$ ,  $p < 0.001$ ,  $\eta^2 = 0.40$ ; a priori linear contrast:  $F(1,9) = 9.26$ ,  $p = 0.014$ ,  $\eta^2 = 0.51$ ). Mauchly's test for sphericity confirmed that the assumptions of the ANOVA were not violated. Although the reported  $p$  values are uncorrected for the six independent tests performed, and in some instances would not satisfy a conservative Bonferroni correction, the consistent overall pattern of the data suggests that these results are unlikely to be due to random chance.

We note that across all conditions participants performed the task remarkably well, with the separation of thresholds between the simultaneous conditions (aspect ratio =  $\sim 1.025$  corresponding to a rectangle subtending  $3.0^\circ \times 3.08^\circ$  visual angle) and the longest inter-stimulus-interval sequential condition (aspect ratio =  $\sim 1.125$  corresponding to a rectangle subtending  $3.0^\circ \times 3.38^\circ$  visual angle) being only  $0.3^\circ$  visual angle. However, the results of the chance condition demonstrate that much of this performance can be accounted for by position information intrinsic to the inducers themselves as compared to the form of the induced rectangle. That said, at IIIs of 0 ms ( $t(18) = 3.76$ ,  $p = 0.001$ ) and 500 ms ( $t(18) = 2.24$ ,  $p = 0.038$ ) participants performed better than chance indicating that the form of the induced rectangle was indeed contributing to the orientation judgment. For IIIs of 1000 ms and above, performance was no different from chance (all  $p > 0.1$ ), suggesting that at these delays, participants may not have truly experienced the representation of a rectangle but rather relied on other cognitive strategies based on the inducer positions to make their judgments.

## Experiment 2. STFI integrates local features, not illusory contours

A question arises as to what is being integrated during the STFI process. One hypothesis illustrated in Fig. 3A is that



**Fig. 3** Two alternative hypotheses of how STFI leads to object representations. **A** Corners defined by each inducer persist and are integrated with form information at a different location and time to support object representations. **B** Corners defined by each inducer generate illusory contours extending beyond the explicitly defined region linking form information at a different location and time to support object representations

after an inducer is removed from the image, there is a persistent representation of the local form information. This persistent representation is then available to be integrated with subsequently presented inducers to generate the surface representation of an occluding square. Alternatively, it may be that illusory contours, generated by each single inducer, also contribute to the integration process serving as a perceptual “glue” for integrating subsequently presented explicit and illusory contours (Fig. 3B). Indeed, previous work has demonstrated that illusory contours can be formed from single inducers (Halko, Mingolla, & Somers, 2008). However, in the case of simultaneously presented inducers, illusory contours are not required for the representation of an induced surface (Dresp & Grossberg, 1997). Instead, local form features such as corners or regions of high curvature provide important shape information (Attneave, 1954; Biederman, 1987), without generating illusory contours.

Previous work on illusory figures has demonstrated that the subjective clarity of an illusory contour is dependent on the ratio of the physically specified edge of the inducer to the total edge length of the illusory figure (Shipley & Kellman, 1992). Specifically, the strength of illusory figure percepts degrades linearly as the ratio between the specified to total edge length is reduced. We exploit the fact that the subjective clarity of illusory contours degrades with the loss of spatial support to determine if surface representations generated by STFI

are a result of illusory contour formation or the integration of local form features across space and time. If illusory contour formation is an integral component of STFI, we would expect performance to decline when the spatial support ratio is reduced; alternatively, a finding that performance does not deteriorate with the loss of spatial support would suggest that STFI supports surface representations through local feature integration without the explicit formation of illusory contours. We compared discrimination ability for the orientation of the occluding rectangle at two separate inter-inducer-intervals when the inducers, on average, provided full or half support of the total length of the spatiotemporally completed illusory edge.

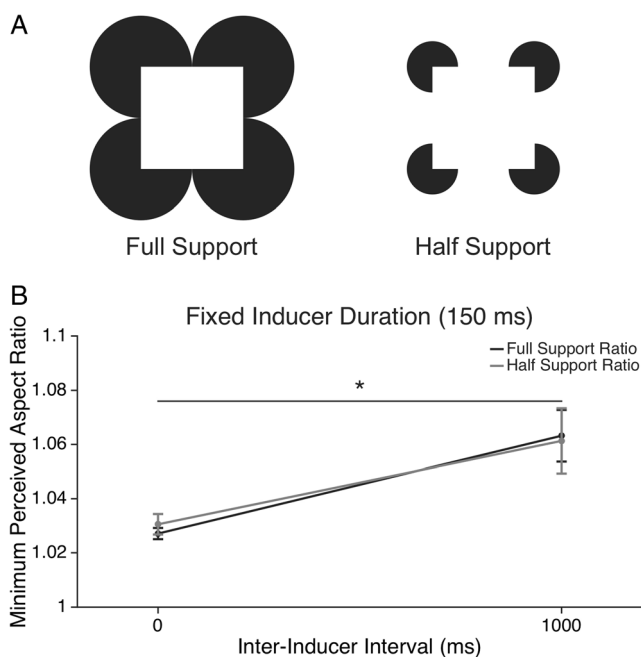
### Stimuli and procedures

The basic stimuli were the same as those in experiment 1 with the exception that the inducers in the half support condition had a diameter of  $1.5^\circ$  visual angle (Fig. 4A; Demonstration Video 5). The centers of the inducers were randomly jittered along the diagonal axis relative to the corner of the occluding

rectangle with a maximum displacement of  $0.5^\circ$  visual angle inward or outward (1/3 of the total inducer diameter). The amount of jitter was randomly determined on each trial so that on average the two conditions had support ratios of 1 and 0.5. We used inter-inducer-intervals of 0 ms and 1000 ms for each spatial support condition. Observers again made a 2AFC orientation judgment (horizontal or vertical) on the occluding rectangle and completed four staircases for each trial type resulting in 16 total staircases.

### Analyses

A 2 (III)  $\times$  2 (Support Ratio) two-way repeated measures ANOVA was performed to identify any potential main effects and interaction between the temporal and spatial components of STFI. In addition, we note that the two full support ratio conditions are nearly identical to the corresponding 0 ms and 1000 ms conditions in the first experiment. Thus, as a check of internal consistency we performed a 2 (III)  $\times$  2 (experiment) repeated measures ANOVA using only the full support ratio data.



**Fig. 4** Stimuli and results for experiment 2. **A** The size of the inducers was varied so that, on average, they provided full spatial support for the contour to be spatiotemporally completed (left) or half spatial support (right). The centers of the inducers were jittered in the actual experiment to prevent participants from making orientation judgments based solely on local cues provided by the inducers. **B** Results of experiment 2. The black line indicates the average orientation discrimination across participants for the full spatial support condition. The grey line indicates the average orientation discrimination across participants for the half spatial support condition. The line with the asterisk indicates a main effect of III. Error bars indicate standard error of the mean

### Results

The mean data averaged across subject are shown in Fig. 4B. Consistent with the results of the first experiment, the thresholds in the 0 ms condition are lower than in the 1000 ms condition (main effect of III:  $F(1,9) = 12.3$ ,  $p = 0.007$ ,  $\eta^2 = 0.58$ ). Mauchly's test for sphericity confirmed that the assumptions of the ANOVA were not violated. As is evident in the figure, decreasing the support ratio had little or no effect on participants' ability to identify the rectangle's orientation (main effect of support ratio:  $F(1,9) = 0.03$ , n.s.; interaction:  $F(1,9) = 0.62$ , n.s.). We note that the lack of an effect for support ratio stands in contrast to past work on the subjective clarity of illusory contours and therefore supports the hypothesis that it is the local form information intrinsic to the inducers that is being accumulated and integrated into the percept of a global figure over time.

As expected, the between experiment repeated measures ANOVA revealed a significant main effect of III duration in which performance with the 0 ms III was better than with the 1000 ms III ( $F(1,9) = 20.97$ ,  $p < 0.001$ ,  $\eta^2 = 0.70$ ). Although we note that in both cases the mean thresholds were slightly lower in the second experiment, perhaps reflecting a small practice effect, the main effect of experiment was not significant ( $F(1,9) = 3.07$ , n.s.) nor was the interaction between III and experiment ( $F(1,9) = 0.21$ , n.s.). Again, Mauchly's test for sphericity confirmed that the assumptions of the ANOVA were not violated. This internal replication further indicates the robust representation of STFI figures.

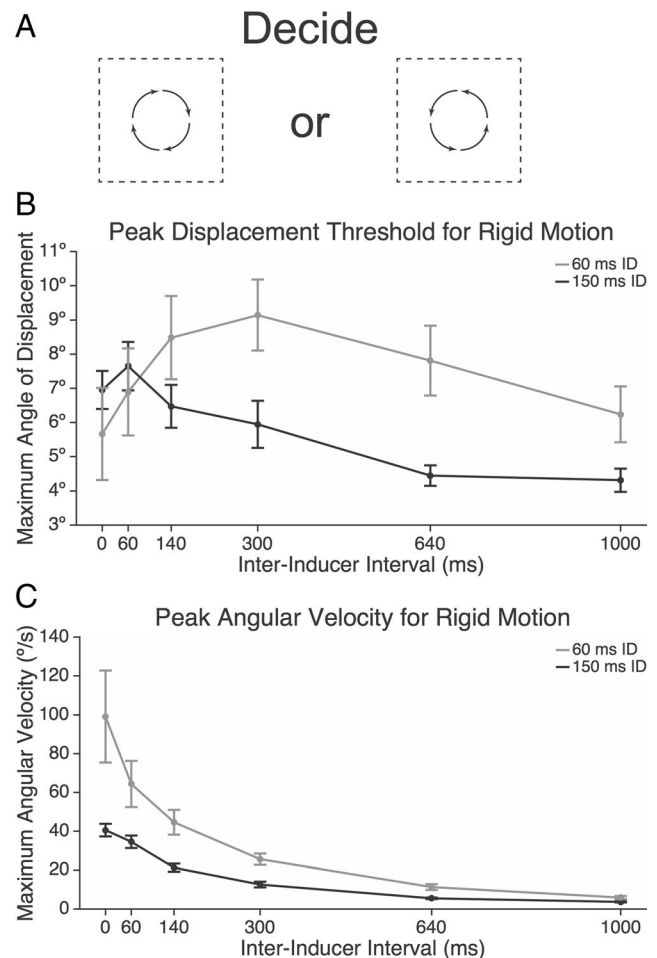
### Experiment 3a and 3b. STFI and position updating can support the perception of a rigidly rotating surface

As demonstrated in the introduction, STFI allows the position of form features to be updated over space and time to support the perception of a rigidly rotating surface. However, the spatial constraints as well as the temporal parameters that lead to position updating remain unknown. The following experiment was designed to identify the restrictions under which STFI can support rigidly rotating surface representations. Specifically, we examined the maximum degree of angular rotation the illusory square could undergo between successive inducers as well as the temporal delays between inducing stimuli that lead to a rigidly rotating square percept. We displayed sequential-inducer stimuli analogous to that shown in Demonstration Video 2 & 6 and varied the degree of rotation between successive inducers using different IDs and IIIs.

#### Stimuli and procedures

The stimuli were similar to those used in the first experiment with the exception that the occluding figure was always a  $3^\circ \times 3^\circ$  square that rotated either clockwise or counterclockwise on a given trial. During a given trial the shape of each inducer was chosen to be consistent with the presence of an occluding square that had been rotated by a fixed amount during the preceding III (Fig. 1C). Each trial consisted of the presentation of eight inducers pseudorandomly ordered such that each corner of the induced figure would be presented twice (2 cycles). In experiment 3a the inducer duration was fixed at  $\sim 150$  ms (13 frames, 85 Hz). In experiment 3b, which was completed by another group of 10 naïve participants, the ID was fixed at  $\sim 60$  ms (5 frames, 85 Hz). As with the other experiments individual inducers alternated between black and white. During the time each inducer was present on the screen, the occluding square remained stationary (i.e., the rotation occurred only during the III). The IIIs for each staircase were chosen from the following list: 0 ms,  $\sim 60$  ms (5 frames, 85 Hz),  $\sim 140$  ms (12 frames, 85 Hz),  $\sim 300$  ms (24 frames, 85 Hz),  $\sim 640$  ms (54 frames, 85 Hz), 1000 ms (85 frames, 85 Hz).

In the first trial of each staircase, the occluding square was rotated  $3^\circ$  between each successive inducer. Observers were first asked: “Did you see a rigidly rotating square.” If they responded “yes,” they were then asked to indicate the direction of rotation via a key press (Fig. 5A). Upon correct responses, the amount of rotation between successive inducers was increased ascending through the following list:  $3.5^\circ$ ,  $4^\circ$ ,  $4.5^\circ$ ,  $5^\circ$ ,  $6^\circ$ ,  $7^\circ$ ,  $9^\circ$ ,  $11^\circ$ ,  $13.5^\circ$ ,  $16^\circ$ . If the participant responded “no” to the first question or made an incorrect rotation judgment, the



**Fig. 5** Task and results for experiment 3. **A** Observers first indicated whether they perceived rigid rotation and, if so, the direction of rotation (clockwise or counterclockwise) for the spatiotemporally integrated square. Stimuli were generated as illustrated in Fig. 2. **B** Rotational thresholds for each III. Points on the lines indicate the maximum angle of displacement the occluding square could undergo between successive inducers before participants could no longer determine the direction of rotation or the percept of a rigidly rotating square was no longer perceived. The black line illustrates  $\sim 150$  ms inducer duration condition and the gray line indicates the  $\sim 60$  ms inducer duration condition. **C** Peak angular velocity for each III. Points on the lines indicate the maximum average angular velocity at which a rigidly rotating square was perceived. The black line illustrates  $\sim 150$  ms inducer duration condition and the gray line indicates the  $\sim 60$  ms inducer duration condition. Error bars indicate standard error of the mean

staircase reversed. On each trial, the starting orientation of the occluding square was randomly chosen to be between either  $20\text{--}35^\circ$  clockwise or  $20\text{--}35^\circ$  counter-clockwise from vertical. Four staircases per trial type were again used resulting in 24 total staircases.

Unlike the first experiment with the static display, we did not include a control baseline due to the nature of the task. As is readily apparent in Demonstration Video 3, the absence of reliable or nearly reliable features makes the directional judgment virtually nonsensical.

## Analyses

The threshold for the maximum amount of rotation that can occur between inducers and peak angular velocity that still lead to percepts of rigid rotation was computed by averaging together the last four reversals in each staircase. These values were then averaged across the four repetitions per condition.

## Results

The data shown in Fig. 5B were averaged across participants using the same procedures as experiments 1 and 2. A repeated measures ANOVA revealed a main effect of III for both the IDs ( $F(5,90) = 5.00, p < 0.001, \eta^2 = 0.29$ ). The main effect of experiment was not significant ( $F(1,18) = 2.24, n.s.$ ); however, the interaction between III and experiment was significant ( $F(5,90) = 5.91, p < 0.001, \eta^2 = 0.25$ ) as can be observed in the peak III difference for the two IDs. Mauchly's test for sphericity confirmed that the assumptions of the ANOVA were not violated. At an ID of ~150 ms, the maximum mean rotation allowed between successive inducers peaked between  $6.5^\circ$ – $7.7^\circ$  for IIIs between ~0–140 ms with the maximum angular displacement occurring at a ~60 ms III. As the III was increased between successive inducers, the percept of rigid motion began to deteriorate as IIIs exceeded ~300 ms. Closer inspection of the data revealed that at such long IIIs, participants often responded “no” to the question “Did you see a rigidly rotating square.” Although the data suggest an inverted u-shaped trend, the quadratic contrast was not significant ( $F(1,9) = 0.40, n.s.$ ). For the ~60 ms ID, the maximum mean rotation allowed between successive inducers peaked between  $8.5^\circ$ – $9.1^\circ$  for IIIs between ~140–300 ms with the maximum angular displacement allowed at an III of ~300 ms. Again, beyond an III of ~300 ms, the percept of rigid motion deteriorated. Interestingly, not only did participants commonly report not seeing a rigidly rotating square at such long IIIs, but also when there was no temporal delay between successive inducers (i.e., 0 ms III). This is supported by a quadratic contrast that reached significance for the ~60 ms ID ( $F(1,9) = 5.97, p = 0.037, \eta^2 = 0.40$ ). Moreover, the maximum mean rotation was greater for the ~60 ms ID at IIIs of ~300 ms ( $t(18) = -2.71, p = 0.014$ ), ~640 ms ( $t(18) = -3.33, p = 0.004$ ), and 1000 ms ( $t(18) = -2.29, p = 0.034$ ).

Although the maximum angle of displacement between inducers is informative to identify spatial constraints as a function of temporal delays, it does not conceptualize STFI and position updating as a combined, spatiotemporal phenomenon. Therefore, we also calculated the peak angular velocity at each III to combine the spatial and temporal factors into a singular concept. As illustrated in Fig. 5C, the maximum angular velocity occurred at a 0 ms III for both the 150 ms ( $M =$

$40.56^\circ/s$ ) and 50 ms ( $M = 99.09^\circ/s$ ) conditions and then rapidly falls off as the III increases. The main effect of III was significant for both IDs ( $F(5,90) = 30.83, p < 0.001, \eta^2 = 0.63$ ), such that increasing the III reduced the maximum angular velocity. In addition, the main effect of experiment was significant ( $F(1,18) = 10.16, p = 0.005, \eta^2 = 0.36$ ). Specifically, the faster ID of ~60 ms increased the rate of angular velocity across all IIIs. Finally, the interaction between III and experiment was significant ( $F(5,90) = 5.15, p < 0.001, \eta^2 = 0.22$ ) and can be accounted for by the non-linear decay of angular velocity with increasing IIIs. Overall, these effects are not surprising as increasing either the ID or III places intrinsic limits on the maximum angular velocity that can be achieved.

These results demonstrate that although STFI can support representations of stationary surfaces across long durations with little spatial constraint, mechanisms leading to the position updating of form features to support rigidly rotating surface representations are much more constrained both spatially and temporally, violating the principle of relatability<sup>2</sup> within strict spatial limits. Moreover, the finding that the peak angle of displacement allowed between successive inducers shifts to a longer III when the ID is decreased demonstrates that spatial and temporal factors mutually interact to support rigid rotation percepts. Importantly, this suggests that STFI and position updating are collectively a spatiotemporal phenomenon. We also conducted a version of Experiment 1 for static stimuli using an ID of ~60 ms. The results, summarized in the Supplemental Material, were quite similar to those obtained with ~150 ms.

## Experiment 4. Prior knowledge influences STFI and position updating

The current experiment used silhouettes of familiar animals instead of an occluding square to investigate the role of prior knowledge in STFI using more ecologically valid stimuli. Occluding silhouettes were presented upright and inverted to determine the influence of prior knowledge on the integration process. If prior knowledge influences the strength of STFI, we expect rigid silhouette percepts to be experienced more often and the direction of rotation to be more readily discriminable when they are upright compared to when they are inverted. Alternatively, if STFI is a purely bottom-up, stimulus driven process, we would not expect differences between the upright and inverted conditions.

<sup>2</sup> The formal definition of relatability can be found in (Kellman & Shipley, 1991) or (Kellman, Garrigan, & Shipley, 2005). Briefly, two contours are said to be relatable if they can be connected by a smooth curve that bends monotonically by no more than  $\sim 90^\circ$ .



## Participants

Seven participants (4 females; mean age = 25 years) participated in the experiment. All participants reported normal or corrected-to-normal vision and provided informed consent before participation according to the guidelines of the Department of Psychology and the IRB of the University of Nevada, Reno.

## Stimuli and procedures

The stimuli and procedures were the same as those described in experiment 3 with the following exceptions: silhouettes of a cat, dog, and rabbit were used instead of an occluding square (Fig. 6A; Demonstration Videos 7-9). The borders of the images were constrained to a  $3^\circ \times 3^\circ$  square and the black silhouettes changed to match the gray background. Square inducers ( $3^\circ \times 3^\circ$ ) were used instead of circles to account for the variation in the outline of the more complex stimuli. Based on the results Experiment 3a, an ID of  $\sim 150$  ms and an III of  $\sim 60$  ms was used. On the first trial of each staircase, the silhouette was rotated  $2^\circ$  between each successive inducer. On a given trial, a random silhouette was presented upright or inverted at a starting angle either  $30\text{--}45^\circ$  clockwise or  $30\text{--}45^\circ$  counter-clockwise from vertical. Participants were first asked to indicate if they perceived a rigidly rotating object and indicated the direction of rotation if they responded “yes.” Upon a correct response, the amount of rotation between successive inducers was increased ascending through the following list:  $3^\circ$ ,  $4^\circ$ ,  $5.5^\circ$ ,  $7^\circ$ ,  $9^\circ$ ,  $11^\circ$ ,  $14^\circ$ ,  $17^\circ$ ,  $21^\circ$ ,  $25^\circ$ ,  $30^\circ$ ,  $35^\circ$ ,  $40^\circ$ . Four staircases for each upright and inverted silhouette were used resulting in 24 total staircases.

## Analyses

A difference index  $\left[ DI = \frac{\text{Upright} - \text{Inverted}}{\text{Upright} + \text{Inverted}} \right]$  was calculated for each participant for the threshold of rotation and percent of trials participants saw a rigidly rotating object to determine the effect of inversion on STFI. This was computed by averaging together the last four reversals in each staircase for each silhouette. These values were then averaged across the four repetitions per silhouette in each condition and collapsing the mean across silhouettes. One-sample t-tests were performed to statistically test differences between the upright and inverted conditions.

## Results

The mean rotation threshold data averaged across participants is shown in Fig. 6B. On average across for upright and inverted figures, the mean angular displacement for which

rigid percepts can be perceived was  $\sim 5^\circ$  of visual angle. Figure 6C shows that the average peak angular velocity that supported a rigid motion percept for upright and inverted figures was  $25\text{--}30^\circ/\text{sec}$ . This is somewhat lower than the results of Experiment 3 which found a threshold of  $\sim 7^\circ$  of visual angle and angular velocity of  $\sim 35^\circ/\text{sec}$  for the same ID and III using simple square stimuli. These results indicate that although STFI and position updating can occur for naturalistic silhouettes, the complex nature of their contours likely imposes greater constraints than with the simple square stimuli.

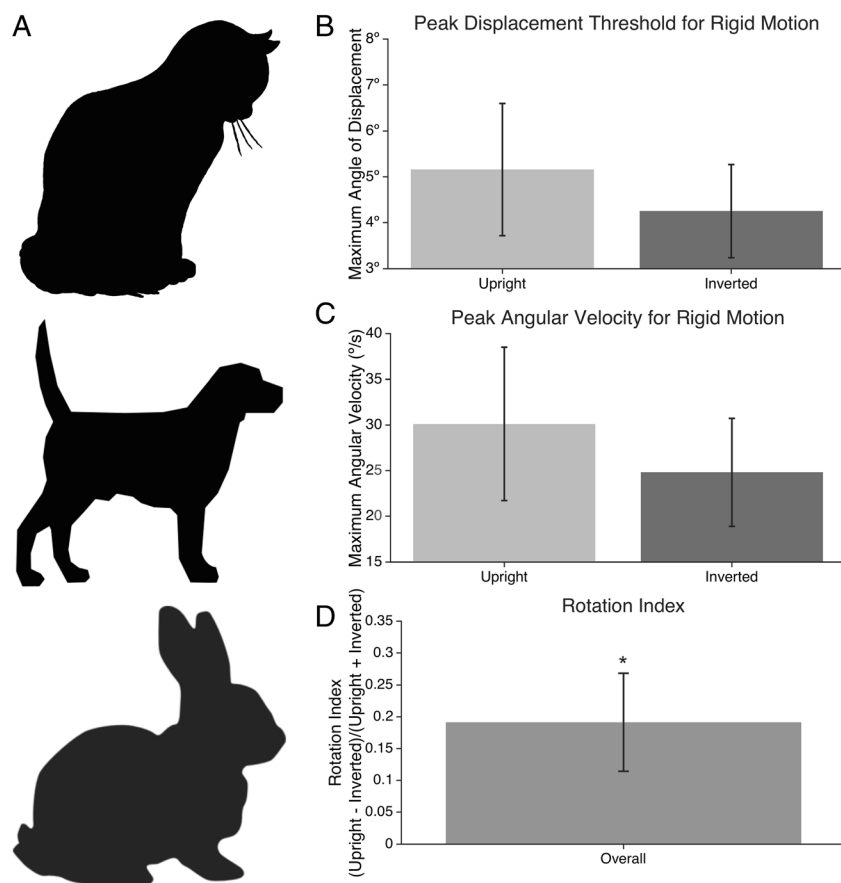
The statistical analyses comparing upright versus inverted silhouettes found the upright stimuli were more readily perceived as rigidly rotating than the inverted ones. Figure 6D illustrates that participants were more accurate at reporting the direction of rotation when silhouettes were upright ( $t(6) = 2.48$ ,  $p = 0.048$ ). A follow-up analysis revealed no differences between silhouettes of cats, dogs, or rabbits ( $F(2, 12) = 0.48$ , n.s.).

## Discussion

We describe a novel illusion in which sequentially presented inducers can lead to percepts of a rigidly rotating surface. The illusion highlights two important processes—spatiotemporal form integration and position updating—that allow us to perceive objects whose features are revealed over time. This set of experiments examined the spatial and temporal parameters under which STFI can support the representation of both stationary and rigidly rotating surfaces that are either simple or naturalistic in nature. We discuss four primary conclusions that can be drawn from the results.

### The integration of local form information over space and time can lead to global representations of stationary objects

The persistence of local form information over relatively long temporal intervals leads to these percepts. Participants performed the task surprisingly well at all IIIs and though their performance was significantly worse than when all inducers were simultaneously presented, the data demonstrate that STFI can support representations of static surfaces when the features are separated across periods exceeding 500 ms. At longer IIIs, although it may be possible to discern some sense of orientation, it is unlikely that this is based solely on an integration of local form features. Indeed, at IIIs greater than or equal to 1000 ms participants did not perform the task significantly better than when local features were not present. To illustrate this point, we encourage readers to view Demonstration Videos 1, 10, and 11 that illustrate examples of trials in which the III increases in duration. As is readily apparent, short IIIs of 0, 50 ms, and 500 ms (Demonstration Videos 1, 10, and 11, respectively) lead to representations of a square. In contrast, at the long III of 2.5 s shown in



**Fig. 6** Stimuli and results for experiment 4. **A** Silhouettes used as occluding surfaces in the experiment. **B** Average rotational thresholds that supported a rigid motion percept for all silhouettes. **C** Average peak angular velocity that supported a rigid motion percept for all silhouettes. **D** Sensitivity indexes for upright vs. inverted conditions.

The left panel shows the index of the difference between participants’ ability to discriminate the direction of rotation for upright and inverted silhouettes. The right panel shows the index of the difference between participants reporting seeing a rigidly rotating object for upright and inverted silhouettes. Error bars indicate standard error of the mean

Demonstration Video 12, there is very little if any sensation of a square.

The neural mechanisms that underlie persistent representations of local form and their integration across space and time remain unknown. Stimulus driven responses of neurons within early areas of visual cortex seldom persist long after the stimulus has been removed, let alone for several hundreds of milliseconds (Hubel & Wiesel, 1968; Peterhans & von der Heydt, 1989; Peterhans, von der Heydt, & Baumgartner, 1986; von der Heydt & Peterhans, 1989; von der Heydt, Peterhans, & Baumgartner, 1984). Moreover, under natural viewing conditions, individuals are constantly moving their eyes, often resulting in a fresh set of bottom-up stimulus driven neural activity. Thus, it is somewhat hard to conceptualize how a purely bottom-up representation of local form information could persist for very prolonged periods of time.

At a higher level, a number of experiments have shown that full object representations can persist for extended periods of time (Ferber, Humphrey, & Vilis, 2003; 2005; Large, Aldcroft, & Vilis, 2005; Strother, Lavell, & Vilis, 2012;

Wong, Aldcroft, Large, Culham, & Vilis, 2009). These experiments exploit the gestalt principle of common fate to generate percepts of objects by having subsets of individual elements move in a coherent fashion. Intriguingly, observers report the perception of coherent objects even after the elements have stopped moving and can do so for intervals exceeding one to two seconds. Neuroimaging data suggests that these persistent representations are being stored at least in part within the relatively high-level lateral occipital cortex (Ferber et al., 2003, 2005). While beyond the scope of the behavioral data presented here, one may speculate a similar role for such high-level neural mechanisms in the persistence observed with STFI.

**STFI integrates features, not illusory contours** Unlike the subjective clarity of illusory contours that degrades linearly as the ratio between specified to total edge length is reduced (Shipley & Kellman, 1992), participants’ ability to perform the orientation task in Experiment 2 did not depend on the spatial support of the inducers. This observation suggests that

the mechanisms underlying the formation of STFI surface representations may not rely strictly upon the same contour propagation mechanisms that underlie illusory contour formation (Peterhans & von der Heydt, 1989; Peterhans et al., 1986; von der Heydt & Peterhans, 1989; von der Heydt et al., 1984; von der Heydt, Zhou, & Friedman, 2000). This is somewhat surprising given the finding that illusory contours can be formed from single inducers (Halko et al., 2008; see their Movie 2). Such contour-extrapolation seems like a natural candidate for mechanisms underlying STFI and likely contributes in some way. However, the result of Experiment 2 is more consistent with the idea that STFI relies on mechanisms that integrate persistent representations of the position and identity of local form features themselves (e.g., corners, contour discontinuities, or regions of high curvature) and not the propagation of illusory contours.

It has long been known that such form features are critical for the representation of an object's shape (Attneave, 1954), and there is a great deal of evidence for neural mechanisms that detect them (Brincat & Connor, 2004; Fujita, Tanaka, Ito, & Cheng, 1992; Hubel & Wiesel, 1965, 1968; Pasupathy & Connor, 1999, 2001, 2002; Tanaka, Saito, Fukada, & Moriya, 1991). Much in the way that co-occurring local form features can be integrated into unified wholes as illustrated with Attneave's cat (Attneave, 1954) or Biederman's cups (Biederman, 1987), STFI seems to reflect a similar level of integration that occurs across time. If the representations of such local form information can persist for prolonged periods, then presumably they can be integrated across space just as they are when simultaneously present in the image.

**STFI can support representations of rigidly rotating surfaces** We demonstrate for the first time that STFI can generate rigidly rotating surface percepts solely from the position updating of form features in the absence of any motion energy in the display. The requisite position updating, however, only occurs within short temporal windows and is subject to strict spatial constraints. Violations of these constraints lead to a breakdown of position updating or of STFI altogether. For the simple square-shaped stimuli we tested, our data indicate that rigid rotation can only be perceived for small angular displacements (max < 8–11°) between inducers presented within ~300 ms of each other and peaking at an III of ~60 ms for ~150 ms IDs. This temporal constraint is extended to a peak III of ~300 ms for ~60 ms IDs. The finding that such non-zero delays allow for greater angular displacement may be due to the fact that real motion takes time to occur. As Demonstration Video 13 demonstrates, providing a small window between the updated form information could be more easily interpreted as smooth rotation compared to an abrupt change of the occluding square's orientation upon the immediate onset of a subsequent inducer. We note, however, that this is only a speculative explanation. In any case, at large

angular displacements occurring within ~300 ms of one another, the misaligned inducers can still lead to a representation of a deformed stationary surface (Demonstration Video 3), and at longer IIIs STFI will break down altogether (Demonstration Video 4). In all, though constrained, these results speak to the importance of spatiotemporally revealed form features in representing the motion of objects.

Our findings concerning spatiotemporal constraints are consistent with a study demonstrating that apparent motion in Kanizsa figures was only perceived when the time interval between successive frame was less than 500 ms (Mather, 1988). In addition, based on the original results of Palmer et al. (2006), Palmer and Kellman (2014) estimated that the visual persistence of translating occluded objects lasts between ~170–270 ms. Thus, these results are in line with our observation that position updating decayed somewhere beyond the ~300 ms III used in our study.

The spatiotemporal constraint of rotation we observed using these stimuli, however, is more limited than previous work. While there is a reasonable amount of research investigating the role of STFI in translating objects (Guttman & Kellman, 2004; Keane, Lu, & Kellman, 2007; Kellman & Shipley, 1991; Palmer & Kellman, 2014; Palmer et al., 2006), to our knowledge Kellman and Cohen (1984) is the only study to examine the STFI using rotational motion. They found that kinetic subjective figures could be perceived by generating interruptions in inducers over time that were consistent with a globally occluding object rotating at a rate of 144°/s. Using short IDs of ~60 ms and IIIs of 0 ms in the current study, the peak angular displacement of the occluding square corresponds to a rotation rate of 99.09°/s—roughly 2/3 of the speed reported in Kellman and Cohen (1984). Critically, however, their study differed in two important ways: 1) all of the inducing elements in their display were always present while the illusory figure underwent rotation, and 2) changes to the occluding figure were continuous such that real motion was present when the occluding figure passed over each inducer.

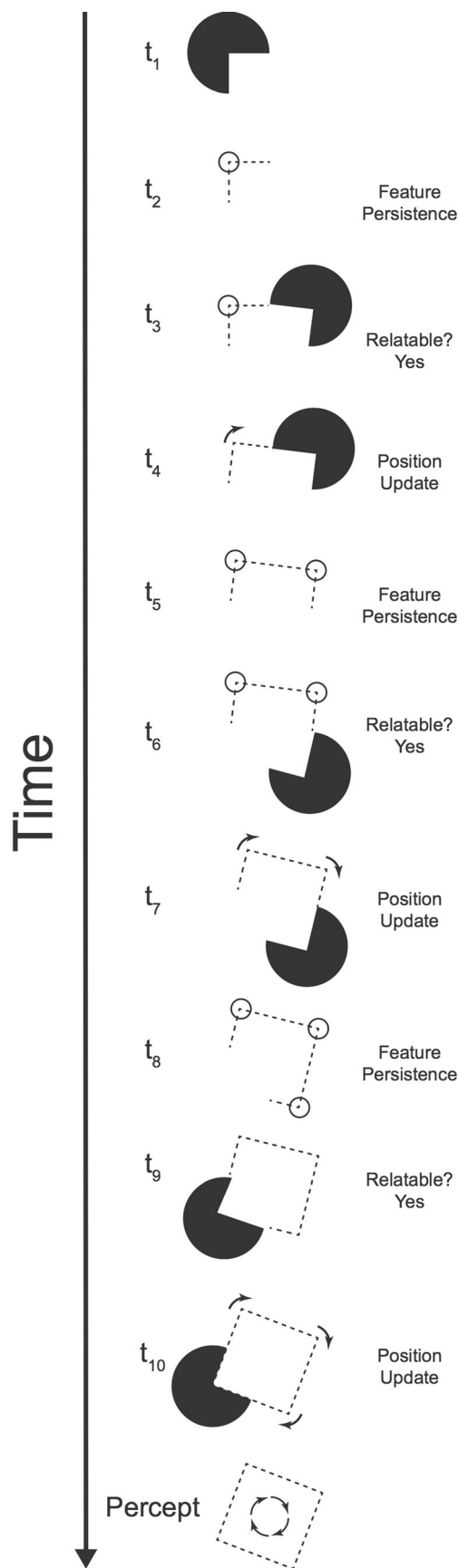
We hypothesize that one reason an increased spatiotemporal constraint was observed with our stimuli is due to the absence of any motion energy in the display. Although Kellman and Cohen (1984) used a rotational velocity of 144°/s, it is likely that similar percepts could have been generated using slower or faster rates, as long as new information provided by changes in the inducers was continuously updated within the temporal windows described above. Because the rotation in our stimuli relied solely on form processing, we would expect the inferred motion that can be derived to be highly dependent on the reliability of successive features. Indeed, Palmer et al. (2006) found that that participants were sensitive to misalignments as small as 5 arcmin or ~0.083° between relatable elements of translating objects. Thus, the spatiotemporal constraint that position updating will

**Fig. 7** Conditions necessary for STFI to support rigidly rotating objects. ▶ The representation of an explicit form feature provided by an inducer at a given moment in time (first row) persists perceptually (second row). If the subsequent inducer provides information about another form feature that is geometrically relatable (third row) and this occurs within a critical time window, the position of the previously represented feature is updated and integrated with the current form information that is explicitly available (fourth row). These newly integrated form features are again accumulated perceptually and continue to be position updated with subsequent form information so long as this new information is geometrically relatable and is revealed within the critical duration under which position updating occurs (less than ~300 ms). These features become integrated together over space and time to support the percept of a rigidly rotating square (last row). If either of these conditions is violated, position updating does not occur and rigid motion is no longer perceived. The above example illustrates the maximal degree of rotational displacement allowed between successive inducers (7°) under which STFI can support rigidly rotating object representations

only occur for small angular displacements within strict temporal windows is consistent with the idea of relatability proposed in other investigations of position updating (Palmer et al., 2006). If the angular displacements between inducers are too great within small temporal windows, the local form features will not be geometrically relatable and position updating will not occur. However, the features will be relatable if the displacement is small enough and persisting representations will be position updated if presented within a brief time-window (Fig. 7). Moreover, Palmer and Kellman (2014) demonstrated that occluded regions of translating objects persist; however, they are perceived to move slower than visible portions of the object. Taken together, these findings seem like a plausible explanation for the increased rotational constraints observed with our stimuli.

A question arises as to the degree to which these constraints are stimulus-dependent. For example, if similar constraints would be derived if the stimuli were smaller/larger or closer/farther apart. Such dependencies are observed across a wide range of stimulus domains. For example, apparent motion can be observed between two objects at shorter ISIs the closer they are together (Korte’s third law of apparent motion: Gepshtein & Kubovy, 2007; Korte, 1915). In addition, we observed that short and long IDs had an impact on the maximum angle of displacement that is allowed between successive inducers and/or the ideal III that leads to rigidly rotating percepts. This suggests that there is interplay between the duration of visible form features and gaps between them that impacts the ideal ratio supporting rigidly rotating percepts. Thus, using various other inducer durations would likely impact the duration over which form features could be maintained and integrated in into rigidly rotating figure representations. Additional experimentation will be required to investigate both of these issues further.

A fundamental question is why position updating would occur at all. If the angular displacement is small and the



features are relatable, why not just spatially integrate them to form a stationary object? Indeed, illusory figures can be formed (often with curved contours) with simultaneously presented inducers whose features are not perfectly collinear. We suggest the fact that STFI can support motion percepts at all speaks to the fundamental challenges the visual system is presented with when trying to represent the motion of objects in the environment. Given the ambiguity of motion detection (Adelson & Movshon, 1982; Nakayama & Silverman, 1988a, 1988b) and the difficulties it presents in the context of partial occlusion, having a mechanism by which position updating of form information can occur enables important aspects of the world around us to be represented in a more accurate way. Indeed, research suggests that the visual system is capable of constructing non-retinotopic figural representations using information generated from the motion signal (Agaoglu, Herzog, & Ogmen, 2012; Nishida, 2004; Otto, Ogmen, & Herzog, 2009), suggesting an important contribution of the motion system in form perception. Recent research suggests that the visual system may accomplish this task by constructing non-retinotopic spatial representations of the object generated from its motion (Agaoglu et al., 2012; Nishida, 2004; Otto et al., 2009).

This is the fundamental argument that is given for other visual phenomena that rely upon mechanisms similar to STFI. In these cases, the perception of a moving object has been demonstrated to rely in part upon an analysis that integrates form information (i.e., corner or surface) present at one location and time with form information at another location and time. For example, anorthoscopic perception (Fendrich, Rieger, & Heinze, 2005; Helmholtz, 1867/1925; Parks, 1965; Zöllner, 1862) demonstrates that form features revealed over time at a given location in space can be integrated to form a coherent percept of a moving object—a necessary condition for perceiving objects through an aperture (i.e., a window or a door that is ajar). Importantly, the detectable form features present at one moment in time are different from those at any other moment in time.

It is not always the case, however, that the locations at which features are detected will correspond to the positions at which those features will be located at a later moment in time. As such, not only does form information need to be spatially integrated at any specific moment in time, but the positions of previously integrated form information must be retained and updated so that they can be spatiotemporally integrated with what is currently visible. Like the stimuli used here, spatiotemporal boundary interpolation (Kellman & Shipley, 1991; Palmer et al., 2006) provides an example of form features being integrated over time at different locations in space. Critically, the locations at which different features are detected will not correspond to the positions at which those features will be located at a later moment in time. To generate a coherent object representation, persistent representations of

form information must be maintained and position updated consistent with the speed and direction of the moving object.

### **Prior knowledge provides an advantage in solving the problem of spatially integrating form features across time**

More robust rotating STFI percepts were generated when occluding silhouettes of familiar animals were oriented upright compared to when they were inverted. For stationary non-sequential stimuli, this inversion effect is well established for the recognition of faces (Schwaninger, Carbon, & Leder, 2003) and to a lesser extent objects (Diamond & Carey, 1986; Tanaka & Curran, 2001). Although STFI and position updating were more constrained for the animal silhouettes, the observed inversion effect in Experiment 4 suggests that object identity facilitates the STFI and position updating processes. This makes intuitive sense when considering how the visual system constructs representations of moving objects in the real world. For example, when seeing a cat run across the room, it may be unrecognizable at first as only portions of the cat may be visible. If one of these partial views contains highly diagnostic features, it may be sufficient to enable recognition that the moving object is indeed a cat. This in turn may ease the task of matching features across time and help overcome basic spatiotemporal challenges in integrating the cat's complex contours.

As stated earlier, local form features such as corners or regions of high curvature provide important cues to the shape of objects (Attneave, 1954; Biederman, 1987). These features also play an important role in determining the speed of moving objects (Blair, Goold, Killebrew, & Caplovitz, 2014; Caplovitz, Hsieh, & Tse, 2006; Caplovitz & Tse, 2007; Ullman, 1979). Specifically, the angular velocity of rotating objects is more accurately estimated when they contain salient contour regions such as contour discontinuities and regions of high contour curvature (Blair et al., 2014). This has important implications for perceiving moving objects under conditions of occlusion, especially considering the finding that occluded regions of objects are perceived to move more slowly (Palmer & Kellman, 2014) and such features may aid in extracting an accurate motion signal. Once a feature has been identified as belonging to an object, it can be maintained and matched with the same features at subsequent locations and points in time. This “configural matching” allows the visual system to construct surface percepts of rigidly rotating and deforming objects (Caplovitz, 2011). Accordingly, the finding that the rotational direction of silhouettes was more easily reported when they were upright may be due to the identification of features indicating global object identity that could subsequently be integrated and used to track the global object thereby facilitating the STFI and position updating. In sum, this suggests STFI is not a purely bottom-up process and prior knowledge can play an important role in our ability to accurately represent the shape and motion of moving objects.

## Conclusions

The results presented demonstrate that STFI is one of many processes that provide a crucial foundation for the visual perception of objects and surfaces. Moreover, local and transiently visible features can be maintained and position updated to support the perception of not only stationary, but also moving surfaces despite the absence of any correlated motion energy in the image. Although in both of these cases the ability to form these representations is dependent on highly constrained spatial and temporal parameters, the fact that stationary and moving surfaces can be perceived at all speaks to the importance of these processes in supporting coherent object percepts. These findings thus contribute to our understanding of the limits under which the visual system is capable of generating object representations through the integration form information over space and time. Together, STFI and position updating help unify the perceptual bits of our visual world to overcome the pervasive problems of motion and occlusion.

**Acknowledgments** This work was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health (1P20GM103650) and a grant from the National Eye Institute of the National Institutes of Health (1R15EY022775). The authors thank Kyle Killebrew and Gennadiy Gurariy for their contributions to this project.

## References

- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525.
- 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.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*(3), 183–193.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115.
- Blair, C. D., Goold, J., Killebrew, K., & Caplovitz, G. P. (2014). Form features provide a cue to the angular velocity of rotating objects. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 116–128.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.
- Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, *7*(8), 880–886.
- Caplovitz, G. P. (2011). Visual form motion interactions. In A. M. Columbus (Ed.), *Advances in Psychology Research* (Vol. 82, pp. 133–152). New York: Nova Science.
- Caplovitz, G. P., & Tse, P. U. (2007). V3A processes contour curvature as a trackable feature for the perception of rotational motion. *Cerebral Cortex*, *17*(5), 1179–1189.
- Caplovitz, G. P., Hsieh, P. J., & Tse, P. U. (2006). Mechanisms underlying the perceived angular velocity of a rigidly rotating object. *Vision Research*, *46*(18), 2877–2893.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107.
- Dresp, B., & Grossberg, S. (1997). Contour integration across polarities and spatial gaps: From local contrast filtering to global grouping. *Vision Research*, *37*(7), 913–924.
- Fendrich, R., Rieger, J. W., & Heinze, H. J. (2005). The effect of retinal stabilization on anorthoscopic percepts under free-viewing conditions. *Vision Research*, *45*(5), 567–582.
- Ferber, S., Humphrey, G. K., & Vilis, T. (2003). The lateral occipital complex subserves the perceptual persistence of motion-defined groupings. *Cerebral Cortex*, *13*(7), 716–721.
- Ferber, S., Humphrey, G. K., & Vilis, T. (2005). Segregation and persistence of form in the lateral occipital complex. *Neuropsychologia*, *43*(1), 41–51.
- Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, *360*(6402), 343–346.
- Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision*, *7*(8), 9.
- Guttman, S. E., & Kellman, P. J. (2004). Contour interpolation revealed by a dot localization paradigm. *Vision Research*, *44*(15), 1799–1815.
- Halko, M. A., Mingolla, E., & Somers, D. C. (2008). Multiple mechanisms of illusory contour perception. *Journal of Vision*, *8*(11), 1–17.
- 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. (1965). Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat. *Journal of Neurophysiology*, *28*, 229–289.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243.
- Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia*, *49*(1), 7–30.
- Kanizsa, G. (1979). *Organization in Vision: Essays on Gestalt Perception*. New York: Praeger.
- Keane, B. P., Lu, H., & Kellman, P. J. (2007). Classification images reveal spatiotemporal contour interpolation. *Vision Research*, *47*(28), 3460–3475.
- Kellman, P. J., & Cohen, M. H. (1984). Kinetic subjective contours. *Perception & Psychophysics*, *35*(3), 237–244.
- Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*(2), 141–221.
- Kellman, P. J., Garrigan, P., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review*, *112*(3), 586–609.
- Kojo, I., Liinasuo, M., & Rovamo, J. (1993). Spatial and temporal properties of illusory figures. *Vision Research*, *33*(7), 897–901.
- Korte, A. (1915). Kinematoskopische untersuchungen. *Zeitschrift für Psychologie*, *72*, 193–296.
- Large, M. E., Aldcroft, A., & Vilis, T. (2005). Perceptual continuity and the emergence of perceptual persistence in the ventral visual pathway. *Journal of Neurophysiology*, *93*(6), 3453–3462.
- Mather, G. (1988). Temporal properties of apparent motion in subjective figures. *Perception*, *17*(6), 729–736.
- Murray, M. M., & Herrmann, C. S. (2013). Illusory contours: A window onto the neurophysiology of constructing perception. *Trends in Cognitive Sciences*, *17*(9), 471–481.
- Nakayama, K., & Silverman, G. H. (1988a). The aperture problem—I. Perception of nonrigidity and motion direction in translating sinusoidal lines. *Vision Research*, *28*(6), 739–746.
- Nakayama, K., & Silverman, G. H. (1988b). The aperture problem—II. Spatial integration of velocity information along contours. *Vision Research*, *28*(6), 747–753.
- Nishida, S. (2004). Motion-based analysis of spatial patterns by the human visual system. *Current Biology*, *14*(10), 830–839.

- 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(6), 1670–1686.
- 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(2), 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(4), 513–541.
- Parks, T. E. (1965). Post-Retinal Visual Storage. *The American Journal of Psychology*, 78, 145–147.
- Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, 82(5), 2490–2502.
- Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, 86(5), 2505–2519.
- Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, 5(12), 1332–1338.
- Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *The Journal of Neuroscience*, 9(5), 1749–1763.
- Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In J. D. Pettigrew, K. J. Sanderson, & W. R. Levick (Eds.), *Visual neuroscience* (pp. 343–351). Cambridge: Cambridge University Press.
- Schwaninger, A., Carbon, C.-C., & Leder, H. (2003). Expert face processing: Specialization and constraints. In G. Schwarzer & H. Leder (Eds.), *Development of face processing* (pp. 81–97). Göttingen: Hogrefe & Huber.
- Shipley, T. F., & Kellman, P. J. (1992). Strength of visual interpolation depends on the ratio of physically specified to total edge length. *Perception & Psychophysics*, 52(1), 97–106.
- Strother, L., Lavell, C., & Vilis, T. (2012). Figure-ground representation and its decay in primary visual cortex. *Journal of Cognitive Neuroscience*, 24(4), 905–914.
- Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, 12(1), 43–47.
- Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, 66(1), 170–189.
- Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge: MIT Press.
- 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(5), 1731–1748.
- von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, 224(4654), 1260–1262.
- von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, 40(15), 1955–1967.
- Wong, Y. J., Aldcroft, A. J., Large, M. E., Culham, J. C., & Vilis, T. (2009). The role of temporal synchrony as a binding cue for visual persistence in early visual areas: An fMRI study. *Journal of Neurophysiology*, 102(6), 3461–3468.
- Zöllner, F. (1862). Über eine neue Art anorthoskopischer Zerrbilder. *Annalen der Physik*, 193(11), 477–484.