

# Occlusion and the solution to visual motion ambiguity: Looking beyond the aperture problem

**Maarten J. van der Smagt**

Experimental Psychology,  
Helmholtz Institute and Utrecht University,  
Utrecht, The Netherlands



**Gene R. Stoner**

Vision Center Laboratory,  
The Salk Institute for Biological Studies,  
La Jolla, CA, USA



A horizontally moving grating viewed within a diamond-shaped aperture can be made to appear to move obliquely by introducing appropriate depth-ordering cues (R. O. Duncan, T. D. Albright, & G. R. Stoner, 2000). It is commonly assumed that the depth cues in such displays determine which line terminators are seen as *intrinsic* to the grating and which are seen as resulting from occlusion and hence *extrinsic* to the grating. The ambiguous motion of the grating (arising from the aperture problem) is then supposed to be overcome by selectively pooling motion signals arising from the intrinsic terminators with those arising from the grating while discounting the motion of the extrinsic terminators. In our first experiment, we tested the sufficiency of this explanation. Observers reported the direction of motion of ambiguously moving random dots viewed through a diamond-shaped aperture defined by four panels. Binocular disparity was used to simulate occlusion: two panels occluded the virtual surface upon which the dots were positioned and two panels were occluded by that surface. Reports were significantly biased toward the direction of the occluding panels. Since none of the moving features abutted the surrounding panels, none should have been classified as extrinsic and hence this result cannot have relied on terminator classification. In a second experiment, we tested the hypothesis that depth-ordering cues selectively gate the propagation of motion signals so that the representation of the moving surface extends behind the occluders. This was tested by asking observers to report the direction of moving dots viewed through a briefly “opened” probe window within either *occluding* or *occluded* panels. Consistent with our hypothesis, evidence of motion propagation was only found for probe windows within *occluding* panels. Surprisingly, however, this propagation was only observed when the dots in the inducer window moved away from the probe window, suggesting a “pull,” and not a “push” mechanism.

Keywords: barber-pole, barber-diamond, depth–motion interaction, occlusion, line terminators

Citation: van der Smagt, M. J., & Stoner, G. R. (2008). Occlusion and the solution to visual motion ambiguity: Looking beyond the aperture problem. *Journal of Vision*, 8(2):4, 1–12, <http://journalofvision.org/8/2/4/>, doi:10.1167/8.2.4.

## Introduction

When viewed through a rectangular aperture, an obliquely oriented moving grating is usually seen to move in the direction of the long axis of the rectangle. The orientation of the aperture (for instance vertical versus horizontal) thus determines the perceived direction of the moving grating. This is the well-known barber-pole illusion. A common explanation for this and related motion phenomena rests on the observation that the velocity of a one-dimensional (1-D) feature such as a grating is ambiguous (the aperture problem), whereas two-dimensional (2-D) features, such as the terminations of the grating, provide unambiguous motion information (Wallach, 1935; Wuerger, Shapley, & Rubin, 1996). Since there are more “line terminators” moving along the long axis compared to the short axis of the rectangular aperture, motion along this axis dominates the percept. Findings consistent with the hypothesis that line terminators serve to overcome the ambiguity inherent in the motion of 1-D

features have come from both psychophysical (Lorencean & Shiffrar, 1992; Shiffrar, Li, & Lorencean, 1995; Shiffrar & Lorencean, 1996) and neurophysiological studies (Pack, Born, & Livingstone, 2003).

Natural visual images typically contain multiple objects. If the recovery of motion is to be veridical, then the motion of 2-D features cannot therefore be indiscriminately applied to the recovery of an object’s motion—only those 2-D features that “belong to” an object should be applied to the recovery of that object’s motion. Consistent with this observation, Shimojo, Silverman, and Nakayama (1989) found evidence that the visual motion system uses depth information to distinguish between *intrinsic* line terminators, those that correspond to the physical endpoints of 1-D features, and *extrinsic* line terminators, those that result from occlusion by another object. Specifically, they found that the barber-pole illusion is markedly reduced (i.e., the grating is seen to move orthogonal to its orientation) when the aperture surrounding the moving grating is placed stereoscopically in front of the grating such that the terminators are seen to result

from occlusion by the aperture. Conversely, when the aperture is placed stereoscopically behind the grating, the illusion returns. One explanation for this and related findings is that the depth information is used by the visual system to identify extrinsic terminators so that the corresponding motion signals can be suppressed: Intrinsic terminators are given more weight than are extrinsic terminators (e.g., Lidén & Pack, 1999).

Van der Smagt and Stoner (2002) reported psychophysical results that appear inconsistent with explanations relying on depth-based classification of terminators. They used the “barber-diamond” stimulus originally devised by Duncan, Albright, and Stoner (2000) to study the neuronal mechanisms underlying depth–motion interactions. Rather than using depth cues to render all of a grating’s terminators either intrinsic or extrinsic as Shimojo et al. (1989) had done, Duncan et al. simulated partial occlusion so that half of the grating terminators were seen as intrinsic and half as extrinsic. In these stimuli, a vertical square-wave grating moved horizontally within a diamond-shaped aperture. While the grating itself had zero disparity, two of the four panels defining the aperture were placed in front of the grating (crossed disparity), while the remaining two were situated behind (uncrossed disparity). As a result of these manipulations, the grating terminators at the *Far* panels were perceived as intrinsic to the grating, whereas those at the *Near* (occluding) panels were perceived as extrinsic. Under these conditions, the motion of the grating was generally perceived to follow the oblique path of the intrinsic terminators (see Figure 1 top-left panel).

Van der Smagt and Stoner (2002) replicated Duncan et al.’s (2000) psychophysical findings and then extended them by showing that depth ordering also modulated the perceived direction of the motion aftereffect (MAE). They had observers adapt to a grating that was perceived to move unambiguously either leftward or rightward within a diamond-shaped aperture with zero disparity. Observers then viewed a static barber-diamond stimulus with *Near* and *Far* panels. It was found that the stationary grating appeared to move obliquely along the axis defined by the *Far* panels just as if the grating were physically moving. Critically, van der Smagt and Stoner (2002) found this result even when the aperture boundaries of the test pattern were well outside the motion-adapted region. Since under these circumstances, the *Near* and *Far* borders were spatially distant from the moving terminators of the adapting stimulus, this result seems difficult to reconcile with depth-based classification of the moving terminators (whether their motion is real or illusory). In the temporal domain, a study by Graf, Adams, and Lages (2004) hints at a similar conclusion. In that study, prior information about depth ordering, without any moving stimulus, biased perceived subsequent motion without depth information (i.e., without specific extrinsic nor intrinsic borders) towards the borders that were *Near* prior to the motion.

Although the findings discussed above appear inconsistent with explanations of depth–motion interactions

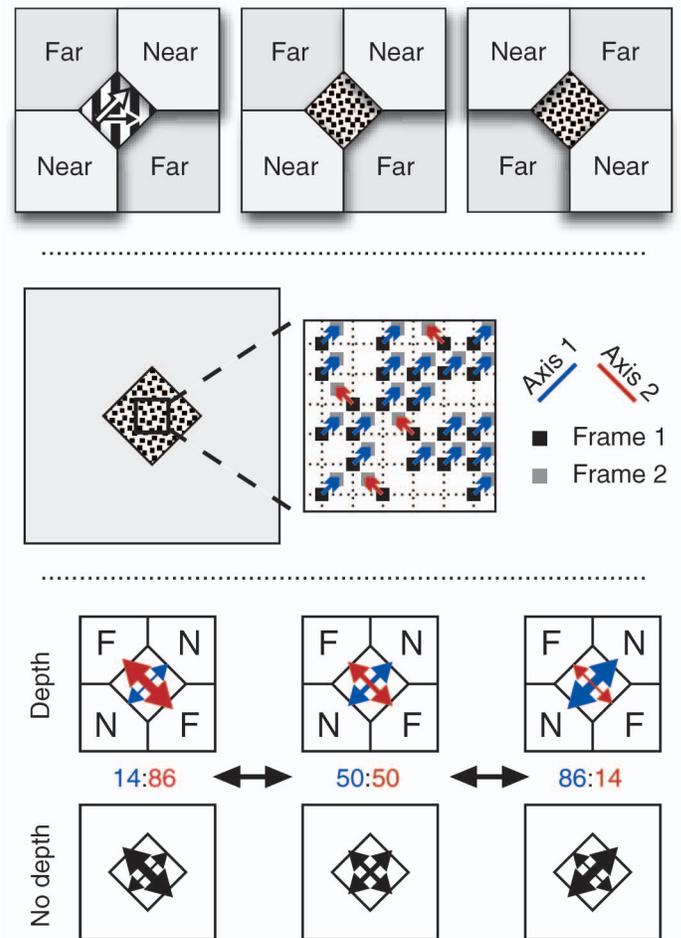


Figure 1. Schematic of stimuli used in Experiment 1. Shadows convey depth in these illustrations, whereas for actual stimuli depth was conveyed via binocular disparity. Top: Comparison of original “barber-diamond” stimulus (only a single depth-ordering is shown) of Duncan et al. (2000) (top-left; black and white arrows depict physical and perceived direction, respectively) with the two depth-ordered stimuli used in Experiment 1 (top-centre and top-right). Middle: two-frame illustration of dot positioning and motions of our stimuli. Dotted lines depict the invisible lattice upon which individual dots were positioned. Bottom: Illustration of how probability of dot motion along the two axes was varied. See text for details.

based on terminator classification and suppression, those experiments were not specifically designed to address that issue.

To explicitly address the viability of terminator-based explanations we designed a variation of the barber-diamond stimulus in which the grating was replaced with 2-D features (dots), whose motion was ambiguous due to the correspondence problem (rather than the aperture problem associated with oriented features<sup>1</sup>). None of these features abutted the surrounding aperture and hence depth-based terminator classification did not apply. We found that depth ordering did indeed affect the perceived direction of these 2-D features. This finding argues against

the sufficiency of explanations based on terminator classification.

In a second experiment, we tested the hypothesis that depth-ordering cues gate the propagation of motion signals. The results partially support this hypothesis, with qualifications, which will be discussed.

## General methods

### Observers

Six paid, naive observers participated in each experiment. All had normal or corrected-to-normal acuity and could easily distinguish between the different depth-ordering conditions. They were required to place their heads in a chin and forehead rest, situated 57 cm away from the stimulus monitor.

### Apparatus and stimuli

For these experiments, the “barber-diamond” stimulus (Duncan et al., 2000) was modified to frame moving dots rather than a moving grating. As in van der Smagt and Stoner (2002), red-green anaglyph glasses were used to create the different binocular disparities that define the three depth planes in this stimulus.

All stimuli were generated on a Pentium-class computer and presented on a CRT with a frame-rate of 85 Hz. Moving dots were viewed through a diamond-shaped aperture (2.25 deg *va*, as measured horizontally or vertically from corner to corner) and were always at zero disparity. The aperture was surrounded by a square textured region, subtending 8.5 degrees *va*. This surround was divided into four quadrants, each of which could be placed at zero disparity, in front of, or behind the dots in the aperture. This yielded three depth-ordering conditions: One “no depth” condition with the entire stimulus at zero disparity, and two mirror-symmetric “depth” conditions (see Figure 1, top-centre and top-right panels), where the two diagonally opposed textured quadrants flanking the aperture had 0.12 deg crossed disparity (in front) while the two complementary quadrants had 0.12 deg uncrossed disparity (behind). These depth manipulations also introduced monocular half-occlusions of the texture elements of the surrounding panels (e.g., Anderson, 1999; Castet, Charton, & Dufour, 1999). Unlike the case for previous depth–motion experiments of this type (e.g., Anderson, 1999; Castet et al., 1999; Shimojo et al., 1989), however, all the *moving* features within the aperture were completely visible to both eyes.

The dots (1.75 min arc) in the aperture were randomly placed on the intersections of an invisible lattice (7 min arc apart) and from frame to frame this lattice was shifted

such that the lattice’s intersections lay exactly in between the intersections of the previous frame. Only 90% of the lattice intersections were occupied by dots (else these stimuli would be completely ambiguous consistent with motion in one of the four directions associated with the two diagonal axes). From frame to frame, the dots were displaced such that they remained positioned on the lattice’s intersections.

This resulted in four possible apparent-motion directions (i.e., towards one of the quadrants mentioned above) for each dot.

The magnitude of each dot’s displacement was the same (see Figure 1, middle panel). Unlike conventional random-dot stimuli, these stimuli have the nice quality that they yield a compelling sense of motion even when there is no coherent “signal”. This is due to the four-direction correspondence ambiguity presented by shifting the lattice by half its horizontal and vertical periods (Figure 1). As a result, motion energy is not evenly distributed across all directions but instead is highest in the four diagonal directions. These stimuli are thus perceptually “quadrable.”<sup>2</sup> The partial occupancy of the lattice intersections yields the additional advantage that perception can be biased by varying the percentage of signal dots (see below).

## Experiment 1

In this experiment, we examined the effects of depth-ordering cues on the perceived direction of moving 2-D features, none of which abutted the surrounding stationary panels. If the type of depth–motion interactions studied here were exclusively mediated by suppression of motion signals arising from “extrinsic” line terminators, then depth-ordering should not affect the perceived motion of these displays.

### Stimuli

All the dots in the aperture were displaced along one of the two diagonal axes from frame to frame. The probability of dots moving along one axis versus the other was varied from 14% to 86% in 25 equal steps. On every trial, the dots were displaced 12 times (every 10 frames, i.e., at 8.5 Hz) yielding a stimulus duration of about 1.4 seconds. Each trial contained two direction reversals, resulting in oscillatory motion with a period of about 0.5 seconds (see Figure 1, bottom panels).

### Procedure

Upon fixating a small annulus in the center of the stimulus the observers initiated each trial with a key-press.

After the motion of the stimulus had ended, a line appeared in the center of the screen, which observers rotated (using the arrow keys on the keyboard) to either  $-45$  degrees or  $+45$  degrees relative to vertical. This task was therefore a 2-alternative forced-choice task: Subjects had to report which axis of oscillatory motion dominated their perception. The observers' choice was subsequently logged by pressing the enter key. A subsequent key-press started the next trial. Conditions were presented in random order, and there were 60 trials per condition for each observer.

## Results

Upon debriefing, none of the observers reported transparent motion. Moreover, one axis of motion was always more dominant perceptually. For the purposes of illustration and analysis, results from the two mirror-symmetric depth-ordering conditions (Figure 1, *top-centre* and *top-right*) were mapped onto a single depth-ordering configuration in which *Near* and *Far* panels are aligned along the  $+45$  and  $-45$  degree axes, respectively. We refer to motion in the direction of the *Near* panels as “axis 1” and motion in the direction of *Far* panels as “axis 2.” For *No-depth* stimuli, axes 1 and 2 map simply onto the  $+45$  and  $-45$  degree axes. The mapping of both Depth and No-Depth conditions is shown in Figure 1. The results of individual observers are presented in Figure 2, where the percentage of perceived motion along axis 1 is plotted as a function of the percentage of signal along that axis. A normal cumulative distribution function (cdf) was fitted through the data for both Depth and No-Depth conditions. As can be seen from Figure 2, curves for the Depth condition are shifted leftward compared to that for the No-Depth condition. This shift is significant for all observers (Table 1).

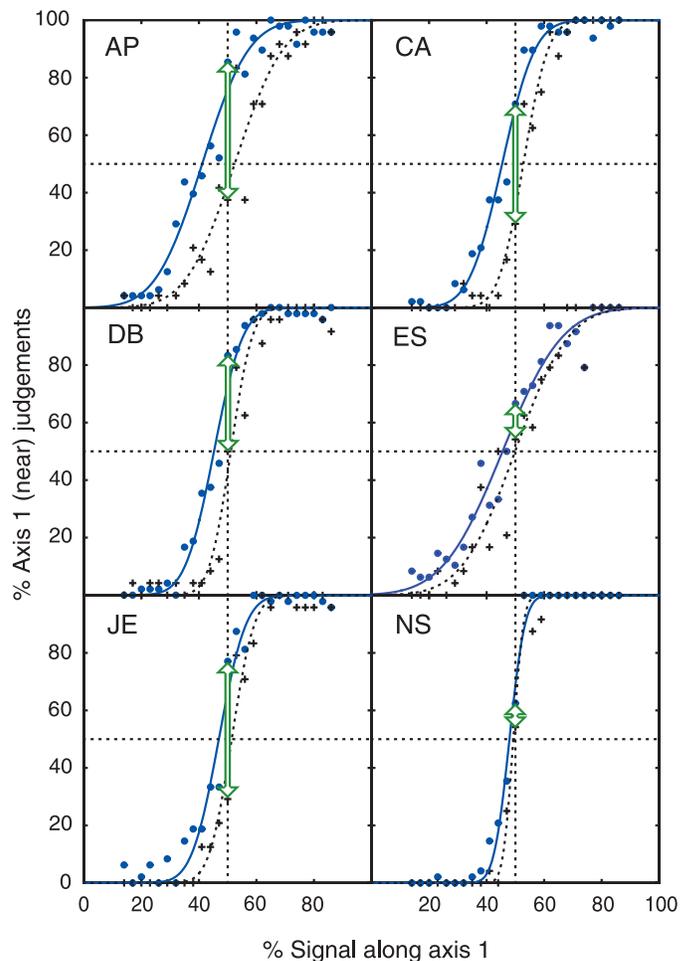


Figure 2. Percentage of each subject's motion reports along axis 1 as a function of percent signal along that axis. Black crosses correspond to reports in the no-depth condition. For stimuli with depth (blue circles) axis 1 reports correspond to motion along the axis defined by the *Near* panels. Lines (dashed for no depth, blue for depth) represent normal cumulative distribution functions fitted through the data. Open arrows indicate the effect of depth ordering on judgments of ambiguous motion.

	Mean				Slope			
	Depth		No depth		Depth		No depth	
AP	39.6	<b>42.7</b>	<b>50.1</b>	<b>54.7</b>	10.5	15.0	10.0	16.5
CA	<b>44.4</b>	<b>46.4</b>	<b>51.8</b>	<b>53.8</b>	7.7	10.4	5.4	8.2
DB	<b>44.4</b>	<b>46.1</b>	<b>50.0</b>	<b>52.2</b>	<b>6.7</b>	<b>9.1</b>	<b>4.3</b>	<b>7.5</b>
ES	<b>43.7</b>	<b>47.4</b>	<b>47.7</b>	<b>52.1</b>	13.8	19.0	11.4	17.7
JE	<b>45.8</b>	<b>48.1</b>	<b>50.4</b>	<b>52.4</b>	5.8	9.0	4.8	7.5
NS	<b>47.4</b>	<b>48.5</b>	<b>48.8</b>	<b>49.7</b>	<b>3.5</b>	<b>5.0</b>	<b>2.1</b>	<b>3.4</b>

Table 1. Ninety-five percent confidence intervals for mean and slope parameters of the fitted normal cumulative distribution functions for Depth and No-Depth conditions. Left column: individual observers. Values in bold type indicate non-overlapping ranges between depth and no-depth conditions.

In the absence of a bias in the motion signal (i.e., 50% signal along axis 1 and consequently 50% signal along axis 2), all observers reported motion along axis 1 (towards *Near* panels) more often than along axis 2 (towards *Far* panels) in the Depth condition, and also more often than along either axis in the No-Depth condition.

Figure 3 (left panel) shows the same data averaged across observers. The right panel shows the mean difference between motion along axis 1 in the Depth and No-Depth conditions. For ambiguous-motion stimuli (50% signal along either axis), this difference reaches nearly 30%. A demonstration of this effect can be found [here](#). It is clear from these data that depth ordering modulates the perceived direction and that this modulation is strongest for ambiguous stimuli. Since, unlike the

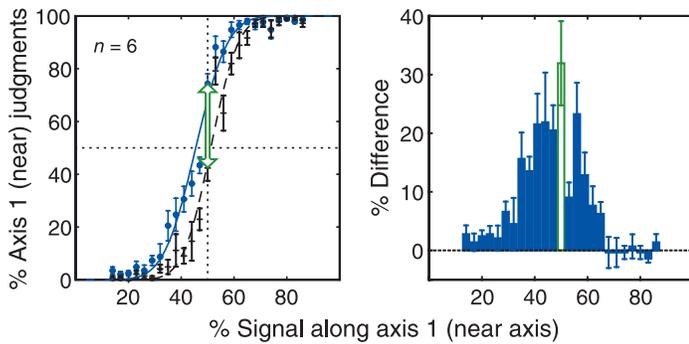


Figure 3. *Left panel:* Average percentage axis 1 reports for depth (blue line) and no-depth conditions (dashed line) as a function of percent signal along that axis. *Right panel:* Average difference between depth and no-depth conditions. Open arrow and bar indicate the difference between depth and no-depth conditions that presented ambiguous motion. Error bars are  $\pm 1$  SEM.

barber-diamond stimuli used by Duncan et al. (2000), the stimuli used here present neither moving 1-D features nor line-terminators subject to depth-based classification, this result is inconsistent with explanations based on terminator classification.

## Experiment 2

The previous experiment showed that terminator-based explanations of depth–motion interactions are insufficient. We next considered alternative mechanisms. Although we realize that the visual system might utilize a number of computational mechanisms, two (related) frameworks that exist in the literature are especially straightforward and intuitive. One is the perceptual phenomenon of amodal completion (Kanizsa, 1979) and amodal continuation<sup>3</sup> (Anderson, 2007; Gillam, 2003), and the other is the propagation of signals through cortex from nearby like-tuned neurons, which can be modulated (or “gated”) by differential depth ordering. Selective propagation of orientation signals within V2 and (to a lesser extent) V1 has been demonstrated by Bakin, Nakayama, and Gilbert (2000). It was found that a neuron’s responses to an optimally oriented contour in the classical receptive field (CRF) increased when a collinear contour was presented outside the CRF. This increase, however, was blocked by an orthogonal line placed between the contours but recovered when the orthogonal line was placed in the near depth plane (thus presumably enabling amodal completion of the contours<sup>4</sup>). It thus appears that orientation signals can propagate “amodally” from one location in the cortical map to another collinear location.

Our second experiment was designed to probe for a motion analogue of this depth-gated collinear facilitation: Depth-gated *co-directional* facilitation (see Figure 4, top).

To illustrate, the 50% signal stimulus used in Experiment 1 would be expected to excite direction-selective units tuned to up-left as well as those tuned to up-right. We postulate that these directly stimulated units facilitate like-tuned units that lie along the motion path consistent with the preferred direction (hence *co-directional* facilitation). In the case of the stimuli of Experiment 1, these facilitated units might include those with receptive fields positioned over one of the four panels surrounding the diamond-shaped aperture. We assume, however, that this facilitation is blocked where depth-ordering cues do not support the existence of occluded motion so that only neurons with receptive fields positioned over the *Near* panels are facilitated (hence *gated co-directional* facilitation). This assumption is consistent with the findings of McDermott, Weiss, and Adelson (2001), who demonstrated that the

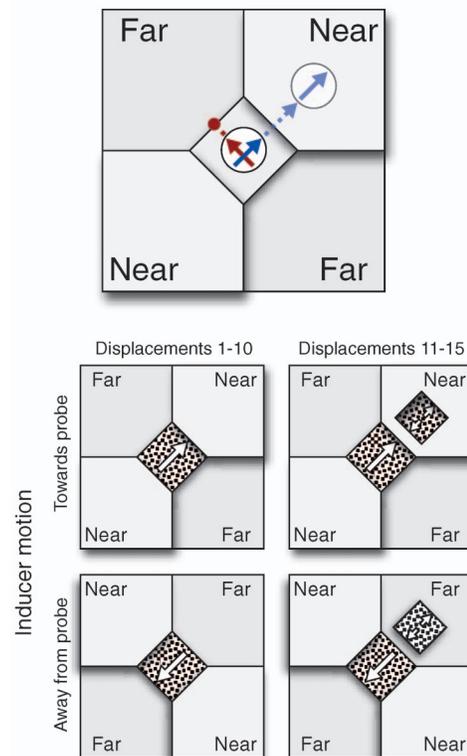


Figure 4. *Top:* Illustration of depth-gated co-directional facilitation scheme. Circles represent cell receptive fields tuned to a certain motion direction (arrows). Dotted arrows depict depth-gated facilitation such that (motion) signals from the cells in the centre selectively propagate to like-tuned cells with receptive fields at the location of the *Near* panels, but not the *Far* panels. *Bottom:* Schematic of stimuli used in Experiment 2. Inducer motion was either towards the probe (TO conditions, top panels) or away from the probe (FROM conditions, bottom panels). Moreover, the probe could appear within a *Near* or a *Far* panel. Signal in the probe was varied. Note: The condition with inducer motion towards the probe in the *Far* panel and the condition with inducer motion away from the probe in the *Near* panel are not shown.

perceived motion of occluded features depends critically upon the extent that *amodal* propagation is supported by occlusion cues. As illustrated in Figure 4 (top), the directional preferences of the facilitated neurons should be those consistent with the “co-directional” propagation of the visible motion underneath the *Near* panels. Our final assumption, then, is that the neurons with receptive fields over the abutting panels (thus encoding *amodal* motion) interact with the neurons that encode the (visible) motion in the central aperture. This would tilt the perceptual balance in favor of motion in the direction of the *Near* panels.

To look for further evidence of these facilitated neurons, we created stimuli with “probe windows” positioned within the *Near* and *Far* panels through which ambiguous motion was presented (Figure 4, bottom). These probe windows were briefly opened after amodal propagation had time to occur. We predicted that amodally encoded motion signals “underneath” the *Near* panels would bias the perception of ambiguous visible motion at the same location, whereas this effect would be absent for *Far* panels where amodal signals are absent. These probe windows were thus meant to give us a glimpse of what was happening underneath the abutting panels.

To summarize, our *gated co-directional propagation* account supposes that (1) motion signals propagate co-directionally, (2) this propagation is gated by occlusion cues, and (3) the neurons activated by this *amodal* propagation can influence the perception of visible motion. Experiment 2 looks for evidence of this last assumption.

## Stimuli

In this experiment, we varied motion coherence such that a given percentage of (signal) dots moved in one direction, while all other (noise) dots were displaced in one of the three remaining directions. Signal dots were displaced in one of two directions, either up-right or down-left. Dots within the central aperture (which will be called the inducer) were displaced 15 times (every 10 frames, i.e., at 8.5 Hz). After 10 displacements of the inducer dots, a second (smaller, 1.7 deg) aperture “opened” within the top-right flanking region (this second aperture will be called the probe). Dots within both the inducer and probe windows then underwent 5 displacements. Signal in the probe was varied in eleven equal steps, from 50% of the dots moving down-left to 50% of the dots moving up-right. Probe-motion signal strength thus varied from being maximally consistent with inducer motion (i.e., 50% signal dots moving in the same direction as the dots in the inducer) to being maximally inconsistent to the motion of the inducer (i.e., 50% signal dots moving in the direction opposite to that of the inducer). Signal strength in the inducer was held constant at 50%.

As in Experiment 1, two depth-ordering configurations were used such that the probe could be opened either in a *Near* region of a *Far* region (both probe and inducer dots were at zero disparity). In addition, the probe windows could be positioned such that inducer motion was either towards the probe (Figure 4, top panels) or away from the probe (Figure 4, bottom panels). A flat version (the entire stimulus at zero disparity) of the stimulus was also included.

## Procedure

Observers initiated each trial with a key-press. Their task was to determine the motion direction of the probe, while retaining fixation on a small annulus in the center of the stimulus. Subjects thus fixated the center of the inducer but had to momentarily shift their attention to the probe window when it appeared. After the stimulus stopped moving, an arrow appeared over the probe window, which the observers adjusted, via arrow keys of keyboard, so that it pointed either down-left or up-right. The observers’ choice was subsequently logged by pressing the enter key. Another key-press then started the next trial. All conditions were presented semi-randomly and were repeated 8 times per block. Eight blocks were run per observer, yielding 64 trials per condition.

Each observer first practiced, for at least 24 trials per condition (11 in total, see above) with a “flat” stimulus, which presented only noise dots in the inducer. A psychometric function (normal cumulative distribution function) was fitted through the results of each observer, and only those observers for whom the  $r^2$  of that function was at least 0.9 were included in the experiment. This ensured that all subjects in this experiment could reliably report the dominant physical motion within the probe windows.

## Results

There were two inducer motion directions, two depth-orderings (in addition to the flat condition), and two choices for reporting probe direction. We plotted the percentage of perceptual reports that agreed with the inducer motion (% responses in inducer direction) as a function of the percentage of signal (in the probe) in that direction. Probe signal considered this way ranged from –50 to 50%.

Figure 5 shows data averaged across the 6 observers. As can be seen, depth ordering appears to have had only a modest impact on perceived motion: The 95% confidence intervals for the means of the fitted psychometric functions for the probe in *Near* and *Far* quadrant overlap (distribution-mean *Near*: –9.7 to –2.7; *Far*: –3.6 to 2.3).

This finding thus provides only weak support for the co-directional facilitation prediction considered above. That prediction, however, was modeled on orientation mechanisms and orientation differs from motion in that the later has a periodicity of 360 degrees rather than 180 degrees. As a consequence, there is a “sidedness” to motion that is lacking in the orientation domain: The two panels along the directional axis differ in that the inducer motion is either *toward* or *away from* them. To determine whether motion propagation was dependent upon motion direction, rather than merely motion axis, we divided the data into TO and FROM categories (See Figure 4). This resulted in 32 trials for a given probe-signal level (per observer, per depth-ordering configuration) for both TO and FROM conditions.

Divided into these categories, an interesting pattern emerges from the data (Figure 6). Consistent with Figure 5, for FROM conditions, the observers’ responses are biased towards the direction of motion of the inducer dots when the probe is in a *Near* panel but not when the probe was in a *Far* panel (Figure 6, lower-left). This bias is significant as the 95% confidence intervals for the distribution-mean of the fitted curves for *Near* and *Far* do not overlap (*Near*:  $-15.8$  to  $-4.7$ ; *Far*:  $-3.8$  to  $5.4$ ). No such effect is seen, however, for the *to* conditions (Figure 6, upper-left): The shift seen in Figure 5 was driven solely by the FROM conditions.

The right panels of Figure 6 show the average difference between *Near* and *Far* configurations for the TO and FROM inducer-motion conditions. Whereas there is no consistent difference for the TO conditions, the FROM conditions reveal a clear trend favoring the *Near* configuration. The largest average difference for the FROM condition is about 19%. The distribution of the

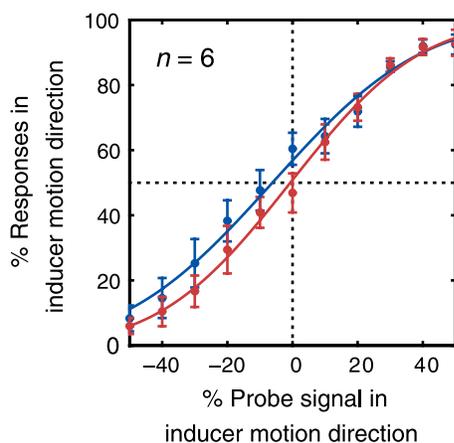


Figure 5. Percent responses in inducer direction (averaged across observers) as a function of the percent probe signal moving in inducer motion direction. Blue and red plots correspond to the probe occurring in the *Near* or *Far* quadrants, respectively. The condition without depth ordering (not shown) largely overlaps with the *Far* condition. Error bars are  $\pm 1$  SEM.

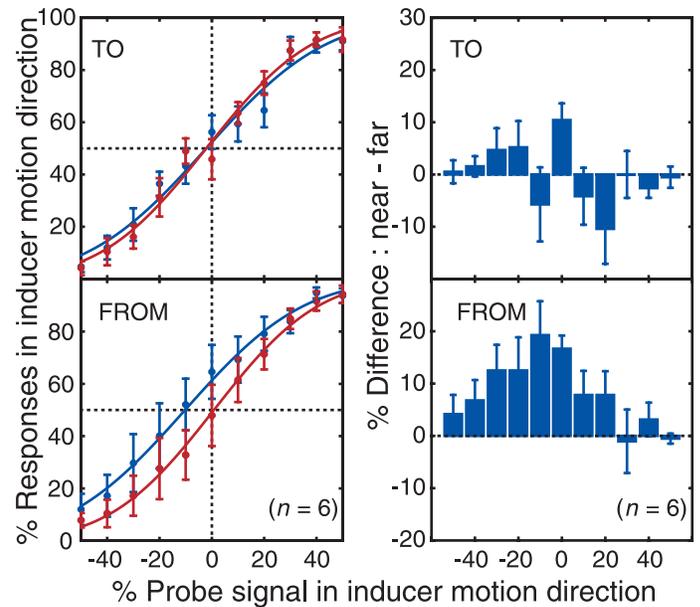


Figure 6. *Left panels*: Percent responses in the direction of the inducer dots (y-axis), averaged across observers, as a function of the percent probe signal moving in the same direction as the inducer dots. Blue and red plots correspond to the probe occurring in the *Near* or *Far* quadrants, respectively. Data are divided into conditions for which inducer motion was directed towards (TO) the probe and away from (FROM) the probe. The conditions without depth ordering (not shown) largely overlap with both the TO and FROM *Far* conditions. *Right panels*: Average difference in percent responses between *Near* and *Far* configurations (see text and Figure 4). Error bars are  $\pm 1$  SEM.

depth-ordering effect on direction discrimination was shifted slightly leftward, indicating a slightly stronger effect when the probe signal dots predominately moved away from the inducer.

The results from individual observers are in line with the group data, though noisy. Ninety-five percent confidence intervals for the fit coefficient (for the “mean” parameter) for probe *Near* and probe *Far* conditions are found in Table 2. For 3 out of the 6 observers, the confidence intervals for these conditions do not overlap, indicating a significant effect of depth ordering for those individuals. For the other three observers, the differences between the two curves are in the same direction, but without significant differences in the fit coefficients. The performance of one of these observers (PL, see Table 2), however, appeared to be near ceiling as almost all probe motions were reported as moving in the inducer direction.

These results support our depth-gated co-directional facilitation hypothesis: Motion signals amodally propagate to the location of the probe window and subsequently bias the perceived motion direction when the probe window opens. However, this appears to occur with the unanticipated additional refinement that facilitation

	Near		Far	
	Lower	Upper	Lower	Upper
AG	−26.5	−17.9	−10.2	0.1
DN	−29.0	−20.0	−11.5	−0.6
JL	−1.7	4.1	1.0	5.8
MM	−4.9	4.3	6.2	15.5
NS	10.0	19.2	15.7	21.5
PL	−98.9	−18.3	−46.1	−11.5

Table 2. 95% confidence intervals for “mean” parameter of the fitted normal cumulative distribution function for *Near* and *Far* probe conditions. Left column: individual observers. Values in bold type indicate data from subjects with non-overlapping ranges between the two depth-ordering conditions. The other subjects showed trends in the same direction but were not individually significant.

applies only to motion detectors with receptive fields positioned where the inducer motion came *from* not where that motion was going *to*.

## Discussion

We have shown that differential depth ordering can significantly bias the perceived direction of moving dots. Since terminators were absent in our stimuli, our results argue against the sufficiency of explanations of depth–motion interactions that rely on classification of terminators as intrinsic or extrinsic. In our second experiment, we found partial support for an alternative account based upon amodal continuation of motion. In what follows, we first relate these findings to previous studies and then offer some tentative conclusions.

### Terminator- and non-terminator-based models of motion processing

Explanations of depth–motion interactions can be divided into those that do and do not invoke depth-based classification of terminators as intrinsic and extrinsic. The validity of terminator classification explanations rests in part on whether terminators play a role in disambiguating the motion of 1-D features even when depth-order is not manipulated. Several lines of evidence do indeed support a role for terminators in overcoming motion ambiguity.

Manipulations of terminator salience have been shown to affect the perception of the barber-pole illusion (e.g., Lorenceau & Shiffrar, 1992; Shiffrar et al., 1995; Shiffrar & Lorenceau, 1996). Furthermore, neurophysiological experiments have revealed neurons in visual cortex, specifically hyper-complex (aka end-stopped) cells (Hubel & Wiesel, 1968), which appear suited to extract the motion of line terminators (Chey, Grossberg, & Mingolla,

1998; Lorenceau, Shiffrar, Wells, & Castet, 1993; van den Berg & Noest, 1993). This suggestion has recently received direct experimental support (Pack, Livingstone, Duffy, & Born, 2003): End-stopped cells in V1 can signal motion direction irrespective of line orientation.

Although, as reviewed above, there is evidence that end-stopped neurons are involved in overcoming the aperture problem, Löffler and Orbach (1999) have advanced a motion model that explains the influence of terminators on motion processing without the need for end-stopped neurons. Likewise, Power and Mouldon (1992) (see also Power, 1993) explained the barber-pole illusion in terms of a population response of ensembles of motion sensors tuned to different directions and to different spatial and temporal correspondences. Their explanation does not depend upon local terminator motion signals, although they are implicitly represented in the population response.

To summarize, although there is compelling evidence of cortical mechanisms specialized in detecting the motion of terminators, it is possible to construct motion models that explain the standard (i.e., without depth manipulations) barber-pole illusion without terminator-based mechanisms.

### Stereoscopic occlusion and the solution to the aperture problem

Shimojo et al. (1989) discovered that the barber-pole illusion was greatly attenuated when the rectangular aperture was placed stereoscopically in front of the moving grating but returned when the aperture was placed stereoscopically behind the moving grating. This finding led to the now well-accepted conclusion that occlusion cues are used to classify terminators as intrinsic or extrinsic, and that extrinsic terminators are discounted in overcoming the aperture problem.

Subsequent studies have taken issue with the idea that stereoscopic depth *per se* is primarily responsible for this classification and concluded instead that interocularly unpaired stimulus regions is the key factor (Anderson, 1999; Castet et al., 1999). Anderson (1999) emphasized the importance of having the terminators themselves (rather than the parts of the surrounding aperture) being interocularly unpaired. Since the stimuli in our studies lacked terminators, our results demonstrate that terminator classification explanations of any type can only be part of the story. This conclusion is in fact consistent with that of Anderson (1999), who concludes that multiple mechanisms likely underlie depth–motion interactions.

### Neuronal basis of depth–motion interactions

To explore the neurophysiological basis of depth–motion interactions, Duncan et al. (2000) created the “barber-diamond” stimulus upon which the stimuli in the

current study were based. In the original barber-diamond stimuli, two of the edges of the diamond-shaped aperture were placed stereoscopically in front of a moving grating and two were placed behind. Perception was found to strongly follow the motion of the intrinsic terminators. For the purposes of neurophysiological investigations, these stimuli had the advantage that depth manipulations could alter perceived motion by 90 degrees (rather than 45 degrees observed with the barber-pole). Moreover, the diamond-shaped aperture allowed the terminators and depth cues to be positioned outside the classical receptive field (CRF). It was found that the directional responses of many area MT neurons reflected the motion of intrinsic over extrinsic terminators. This was true even when the terminators, and depth cues were outside the CRF. These findings thus confirmed that direction-selective neurons were sensitive to depth-ordering cues. It should be stressed, however, that these findings did not demonstrate that this sensitivity occurred by differential weighting of intrinsic versus extrinsic motion signals.

In support of the generality of Duncan et al.'s (2000) findings, Pack, Gartland, and Born (2004) reported that MT neurons are more strongly influenced by the motion of intrinsic versus extrinsic terminators when monocular occlusion cues support that classification. However, these monocular occlusion cues intruded into the CRF and introduced contrast differences between intrinsic and extrinsic terminators. In recognition of the possibility that MT neurons were differentially sensitive to the contrast of occluded and unoccluded terminators rather than to occlusion per se, Pack et al. conducted control experiments, the results of which supported the occlusion interpretation.

Where might sensitivity to occlusion cues arise? In their study of area MT, Pack et al. (2004) also recorded the responses of a small number of area V1 neurons and reported that these neurons respond more to intrinsic than to extrinsic terminators. Their findings suggest therefore that area MT's differential sensitivity to intrinsic versus extrinsic terminators might be inherited from area V1 (though a role for luminance contrast was not, however, ruled out in the case of V1). Whether this sensitivity was restricted to end-stopped neurons is also unclear. Relevant to this discussion are the recent findings of Guo et al. (2006), who found evidence that some area V1 neurons could overcome the aperture problem but that this ability was not restricted to end-stopped neurons. Taken together, these data provide tantalizing support for a role for V1 in depth–motion interactions but the precise role of end-stopped neurons (and hence terminator classification) remains unresolved.

### Amodal continuation as an alternative to terminator-based explanations

Our results demonstrate the insufficiency of explanations that rely upon depth-based terminator classification.

Although other explanations might, of course, be conceived, we find that our results are particularly consistent with amodal continuation accounts. Amodal continuation could underlie depth–motion interactions in three non-mutually exclusive ways: one based on terminators, and two not. First, amodally continued visual features could be used in establishing motion correspondence between occluded and unoccluded form tokens (Shimojo & Nakayama, 1990). This mechanism requires that oriented features abut occluders to form extrinsic terminators and hence does not apply to our current study. In a study similar to ours, however, Watanabe and Cole (1995) showed that occluding surfaces (simulated with monocular cues) disambiguated the perceived direction of a horizontally moving group of dots. In their displays, unlike in ours, the dots on the margins were partially occluded. Hence it is plausible that in Watanabe and Cole's study, amodal continuation of the dots may have served to establish correspondence between amodally continued dots and fully visible dots in the next frame.

A second way that amodal continuation could be involved is by facilitating apparent motion trajectories that appear to occur behind occluders. The occluders, by this account, provide ecological “cover” for the fact that no physical motion appeared between subsequent temporal frames. Shimojo and Nakayama (1990) found evidence against this explanation, which in any case would not apply to [Experiment 1](#) as correspondence tokens (i.e., the dots) did not straddle the panels in our display. Third, motion itself, independent of form information, could be amodally continued so that additional motion signals are introduced, which when pooled with the motion signals arising from visible parts of the surface, alter motion interpretation (Duncan et al., 2000). Our results favor this third mechanism though they do not rule out a contribution from other mechanisms in other types of displays. Psychophysical experiments can, of course, only provide indirect evidence of amodal continuation (or completion) and hence further neurophysiological studies are needed if the underlying neuronal mechanisms are to be resolved.

### Depth-gated co-directional propagation of motion signals

The results of our second experiment provide additional evidence that motion propagated amodally behind the *Near* panels (though only in the FROM condition, [Figure 6](#) bottom panels). Of course [Experiment 2](#) only offers *indirect* evidence for amodal propagation in [Experiment 1](#), and therefore its results do not conclusively rule out other accounts. Nevertheless, as we have reviewed in this paper, there are several other psychophysical lines of evidence in support of the importance of amodal completion in motion interpretation (e.g., McDermott et al., 2001).

Amodal continuation can, in principle, be explained physiologically by propagation of signals through cortex

from nearby like-tuned (in this case direction selective) neurons, which can be modulated (or “gated”) by differential depth ordering. Such a propagation of signals has been demonstrated by Bakin et al. (2000) for static features in V2 and V1: The response of a neuron to an optimally oriented contour in its receptive field was facilitated by a flanking, collinear contour outside the CRF. An orthogonal bar in the same or *Far* depth plane blocked the flank-induced facilitation, whereas the facilitation was not blocked when this bar was placed in the *Near* depth plane. Interestingly, Sugita (1999) reported that some cells in V1 respond to the non-visible part of an amodally continued contour (moving in the cell’s preferred direction). Bakin et al.’s and Sugita’s findings provide compelling evidence of area V1’s involvement in depth-gated collinear facilitation. Given Pack et al.’s (2004) evidence that occlusion cues also alter directional responses in area V1, it is tempting to conclude that V1 may support depth-gated co-directional facilitation as well.

### Pull versus push propagation

The co-directional facilitation hypothesis underlying [Experiment 2](#) was inspired in part by the phenomena of collinear facilitation, which has been shown to be gated by depth cues (Bakin et al., 2000). Orientation of course differs from direction in that the latter has a periodicity of 360 degrees rather than 180 degrees. As a consequence, there is a “sidedness” to motion that is lacking in the orientation domain: The two panels along the directional axis differ in that the inducer motion is either *toward* or *away from* them. Although we had no *a priori* reason to postulate an asymmetry between TO and FROM, to our surprise that is what we found: Depth ordering influenced the perceived motion direction in the probe, but only when the inducing motion was directed *away from* the probe. As yet, we can only speculate as to the explanation of this finding. There are other direction-related asymmetries in the literature that could conceivably be related to our findings. For example, motion within a stationary envelope (e.g., drifting gratings in a stationary Gabor patch) can induce a bias in the perceived envelope position in the direction of motion (De Valois & De Valois, 1991; Ramachandran & Anstis, 1990; Whitaker, McGraw, & Pearson, 1999). This has been shown to be consistent with a spatial asymmetry in the synaptic connections of direction-selective cells (Fu, Shen, Gao, & Dan, 2004). Unfortunately, this asymmetry appears to be in the opposite direction of that predicted by our results.

It is possible that the push–pull asymmetry we have observed is inherent not to visual motion processing *per se* but to attention. It has been found that spatial attention can reduce the influence of collinear facilitation (Freeman, Sagi, & Driver, 2001; Ito & Gilbert, 1999). If we assume that the same is true for co-directional facilitation, then we might further speculate that our findings can be explained by asymmetrical allocation of attention around TO and

FROM probe windows. If, for instance, subjects attentionally (but involuntarily) tracked the motion of the inducer (even perhaps following the motion as it moved beneath the occluding panels), then the focus of attention would be closer to the probe aperture for TO conditions than for FROM conditions. Assuming that attention reduces co-directional facilitation as has been found for collinear facilitation this would then explain our observation that co-directional facilitation is only significant for the FROM condition. This is of course highly speculative and further experiments are necessary to gain insight to this asymmetry.

## Conclusion

We have shown that depth-ordering information influences the perceived direction of moving dots. Our results cannot be explained by depth-based classification of terminators, as terminators were absent in our displays. Instead, we have offered a hypothesis of depth-gated co-directional propagation of motion signals such that amodally represented motion is pooled with signals arising from visible motion. Our results do not rule out mechanisms involving terminator classification in other displays but rather are consistent with the notion that multiple mechanisms are involved in the ability of depth cues to overcome visual motion ambiguity.

## Acknowledgments

MJvdS was supported by an innovational research incentives grant (veni-scheme nr. 451-02-094) of the Netherlands Organization for Scientific Research (NWO). GRS was supported by the National Eye Institute (grant 2RO1EY12872-05). Part of this research was conducted when MJvdS was supported by a Howard Hughes Medical Institute grant to Thomas D. Albright.

We thank Christiaan Paffen, Francois Klam, and Anja Schlack for valuable comments on an earlier draft of this manuscript.

Commercial relationships: none.

Corresponding author: Gene R. Stoner.

Email: gene@salk.edu.

Address: 10010 North Torrey Pines Rd., La Jolla, CA 93027, USA.

## Footnotes

<sup>1</sup>The aperture problem can, however, be viewed as a special case of the correspondence problem in that one-dimensional

features are “undifferentiated” along their long axis so that motion correspondence along that axis is ambiguous (Wallach, 1935; Wuerger et al., 1996).

A pilot study with this stimulus demonstrated that all (6) observers nearly always reported motion in one of the four oblique motion directions (even with 0% coherence) rather than in any of the cardinal directions.

Amodal completion is here distinguished from continuation in that the former joins two separated features whereas the latter implies continuation of a feature “beneath” an occluder without necessarily linking with a second feature.

The amodal completion interpretation of these results is somewhat problematic given that the orthogonal lines did not abut the collinear lines. Such a stimulus gives a weak perceptual completion. See Albright and Stoner (2002).

## References

- Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, *25*, 339–379. [PubMed] [Article]
- Anderson, B. L. (1999). Stereoscopic occlusion and the aperture problem for motion: A new solution. *Vision Research*, *39*, 1273–1284. [PubMed]
- Anderson, B. L. (2007). The demise of the identity hypothesis and the insufficiency and nonnecessity of contour relatability in predicting object interpolation: Comment on Kellman, Garrigan, and Shipley (2005). *Psychological Review*, *114*, 470–487. [PubMed]
- Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, *20*, 8188–8198. [PubMed] [Article]
- Castet, E., Charton, V., & Dufour, A. (1999). The extrinsic/intrinsic classification of two-dimensional motion signals with barber-pole stimuli. *Vision Research*, *39*, 915–932. [PubMed]
- Chey, J., Grossberg, S., & Mingolla, E. (1998). Neural dynamics of motion processing and speed discrimination. *Vision Research*, *38*, 2769–2786. [PubMed]
- De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research*, *31*, 1619–1626. [PubMed]
- Duncan, R. O., Albright, T. D., & Stoner, G. R. (2000). Occlusion and the interpretation of visual motion: Perceptual and neuronal effects of context. *Journal of Neuroscience*, *20*, 5885–5897. [PubMed] [Article]
- Freeman, E., Sagi, D., & Driver, J. (2001). Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. *Nature Neuroscience*, *4*, 1032–1036. [PubMed] [Article]
- Fu, Y. X., Shen, Y., Gao, H., & Dan, Y. (2004). Asymmetry in visual cortical circuits underlying motion-induced perceptual mislocalization. *Journal of Neuroscience*, *24*, 2165–2171. [PubMed] [Article]
- Gillam, B. (2003). Amodal completion—A term stretched too far: The role of amodal continuation. *Perception*, *32* (suppl), 27.
- Graf, E. W., Adams, W. J., & Lages, M. (2004). Prior depth information can bias motion perception. *Journal of Vision*, *4*(6):2, 427–433, <http://journalofvision.org/4/6/2/>, doi:10.1167/4.6.2. [PubMed] [Article]
- Guo, K., Robertson, R., Nevado, A., Pulgarin, M., Mahmoodi, S., & Young, M. P. (2006). Primary visual cortex neurons that contribute to resolve the aperture problem. *Neuroscience*, *138*, 1397–1406. [PubMed]
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*, 215–243. [PubMed] [Article]
- Ito, M., & Gilbert, C. D. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron*, *22*, 593–604. [PubMed] [Article]
- Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.
- Lidén, L., & Pack, C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: A neural network model. *Vision Research*, *39*, 3301–3320. [PubMed]
- Löffler, G., & Orbach, H. S. (1999). Computing feature motion without feature detectors: A model for terminator motion without end-stopped cells. *Vision Research*, *39*, 859–871. [PubMed]
- Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, *32*, 263–273. [PubMed]
- Lorenceau, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, *33*, 1207–1217. [PubMed]
- McDermott, J., Weiss, Y., & Adelson, E. H. (2001). Beyond junctions: Nonlocal form constraints on motion interpretation. *Perception*, *30*, 905–923. [PubMed]
- Pack, C. C., Born, R. T., & Livingstone, M. S. (2003). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, *37*, 525–535. [PubMed] [Article]
- Pack, C. C., Gartland, A. J., & Born, R. T. (2004). Integration of contour and terminator signals in visual area MT of alert macaque. *Journal of Neuroscience*, *24*, 3268–3280. [PubMed] [Article]

- Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: Two-dimensional motion signals in macaque V1. *Neuron*, *39*, 671–680. [[PubMed](#)] [[Article](#)]
- Power, R. P. (1993). Tests of the dipole model of perceived movement in apertures. *Perception*, *22*, 1099–1110. [[PubMed](#)]
- Power, R. P., & Moulden, B. (1992). Spatial gating effects on judged motion of gratings in apertures. *Perception*, *21*, 449–463. [[PubMed](#)]
- Ramachandran, V. S., & Anstis, S. M. (1990). Illusory displacement of equiluminous kinetic edges. *Perception*, *19*, 611–616. [[PubMed](#)]
- Shiffrar, M., Li, X., & Lorenceau, J. (1995). Motion integration across differing image features. *Vision Research*, *35*, 2137–2146. [[PubMed](#)]
- Shiffrar, M., & Lorenceau, J. (1996). Increased motion linking across edges with decreased luminance contrast, edge width and duration. *Vision Research*, *36*, 2061–2067. [[PubMed](#)]
- Shimojo, S., & Nakayama, K. (1990). Amodal representation of occluded surfaces: Role of invisible stimuli in apparent motion correspondence. *Perception*, *19*, 285–299. [[PubMed](#)]
- Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619–626. [[PubMed](#)]
- Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, *401*, 269–272. [[PubMed](#)]
- van den Berg, A. V., & Noest, A. J. (1993). Motion transparency and coherence in plaids: The role of end-stopped cells. *Experimental Brain Research*, *96*, 519–533. [[PubMed](#)]
- van der Smagt, M. J., & Stoner, G. R. (2002). Context and the motion aftereffect: Occlusion cues in the test pattern alter perceived direction. *Perception*, *31*, 39–50. [[PubMed](#)]
- Wallach, H. (1935). Ueber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, *20*, 325–380.
- Watanabe, T., & Cole, R. (1995). Propagation of local motion correspondence. *Vision Research*, *35*, 2853–2861. [[PubMed](#)]
- Whitaker, D., McGraw, P. V., & Pearson, S. (1999). Non-veridical size perception of expanding and contracting objects. *Vision Research*, *39*, 2999–3009. [[PubMed](#)]
- Wuerger, S. M., Shapley, R. M., & Rubin, N. (1996). “On the visually perceived direction of motion” by Hans Wallach: 60 years later. *Perception*, *25*, 1317–1367.