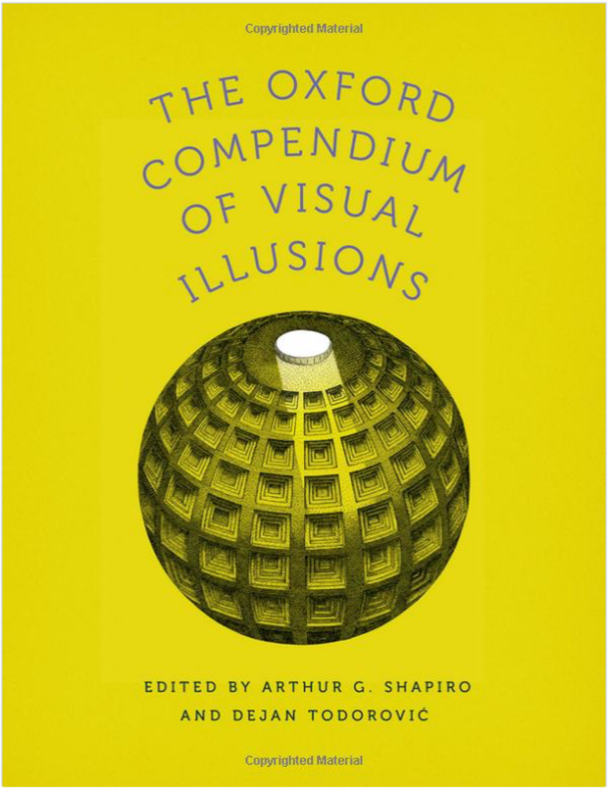
Chapter 71. Second-order Reversed Phi

Zhong-Lin Lu and George Sperling

in



Lu, Z-L., & Sperling, G. (2018). Second-order Reversed Phi. In Shapiro, A. & Tedorovic, D. (Eds). The Oxford Compendium of Visual Illusions. New York, NY: Oxford University Press. Chapter 71, Pp. 522-526.

Chapter 71
Second-Order Reversed Phi

Zhong-Lin Lu and George Sperling

A phi stimulus is composed of a vertical white bar at one location followed after a brief interval by an adjacent white bar. Observers perceive motion from the first to the second bar. The perceived motion in the phi stimulus was famously important to the Gestalt psychologists (e.g., Wertheimer, 1912) as an example of pure motion. When Reichardt (1961) showed a phi stimulus to a beetle (chlorophanus), the beetle also behaved as if it has seen motion from the first to the second bar. However, when the second bar was black instead of white, the beetle responded as if the direction of motion had been reversed. This observation formed the cornerstone of Reichardt's correlation model of the beetle's motion computation, a model that with suitable elaborations also describes human motion computations (van Santen & Sperling, 1984, 1985).

When almost any object is depicted at one location at time 1 (frame 1), and that same object is shown at an adjacent location after a suitable delay (frame 2), it generally produces a vivid sensation of motion. In what has come to be known as a first-order reversed-phi motion stimulus, the black–white contrast of successive frames is reversed; that is, the even frames are the negatives of the odd frames and vice versa. Under suitable conditions, the direction of apparent motion will appear to be reversed (Anstis, 1970; Anstis & Rogers, 1975; Chubb & Sperling, 1989).

In second-order reversed-phi stimuli (Lu & Sperling, 1999; Nishida, 1993), what is reversed between frame 1 and frame 2 is not the black–white contrast of individual pixels but the overall amount of texture-contrast of areas of the stimuli. The stimuli (Fig. IV.71-1) are constructed such that the first-order motion system (the one originally described by Reichardt, 1961) cannot perceive any systematic motion; that is, the stimuli are invisible to first-order motion analysis. In second-order reverse phi, the salient features (areas with a particular texture) that across frames reverse the sign of their difference with the background (e.g., areas that have more texture contrast than the background in frame 1) have less texture contrast than the background in frame 2. Recall that the first-order motion energy model utilizes the contrast of individual pixels including the sign of the contrast, the second-order motion energy model utilizes the total quantity of texture contrast of textured areas without the sign of the contrast, and the third-order motion energy model utilizes the salience of areas (i.e., high-salience features versus low-salience background; Lu & Sperling, 1995a, 1995b, 2001). In the stimuli of Figure IV.71-1, the salient stimulus *features* move in the forward (feature displacement) direction, but the second-order motion energy model predicts motion in the reversed direction.

In all the second-order reversed-phi stimuli shown in Figure IV.71-1, a neutral background texture occupies 75% of the display area. The remaining 25% of the stimulus is filled with a textured rectangle that varies from frame to frame in texture contrast (Fig. IV.71-1a), in texture spatial frequency (Fig. IV.71-1b), or in temporal frequency (i.e., flicker rate; Fig. IV.71-1c).

In Figure IV.71-1a the translating bars are composed of either high-contrast texture or low-contrast texture on a medium-contrast background. In successive frames, the texture bars move one bar width to the right and the high-contrast regions switch to low contrast and vice versa. For the second-order system, increasing and decreasing texture contrast is equivalent to increasing and decreasing luminance in the first-order system (Chubb & Sperling, 1988; Chubb, Sperling, & Solomon, 1989). The stimulus of Figure IV.71-1a is similar to the stimulus of Nishida (1993). Nishida's stimulus involved just two levels of texture contrast; the areas to which these contrast levels were assigned alternated in consecutive frames. The stimulus in Figure IV.71-1a involves three levels of texture contrast, with the middle level remaining the same from frame to frame and the areas of high and low contrast reversing on consecutive frames.

In Figure IV.71-1b the bars are composed of either low spatial frequency sine-wave gratings or high spatial frequency sine-wave gratings on a medium spatial frequency sine-wave grating background. The bars shift one bar width to the right and alternate between high and low spatial frequency textures in successive frames, always with a new random spatial phase. In second-order motion, decreasing spatial frequency is equivalent to increasing texture contrast (Werkhoven, Sperling, & Chubb, 1993, 1994),

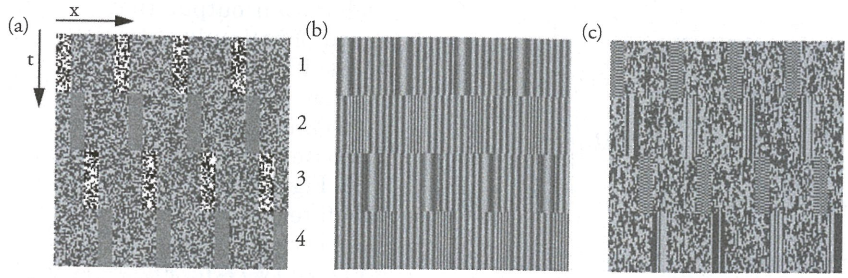


Figure IV.71-1. Three types of second-order reversed-phi stimuli. Each panel shows four consecutive frames (labeled 1, 2, 3, 4). A small rectangular area that occupies 25% of the area is the "figure"; the rest is background. When reproduced properly, figure and background have the same average contrast in all panels; that is, from a sufficiently large viewing distance, the panels look uniformly gray. In each successive frame, the figure translates 90° (one-quarter of the period, the repetition cycle) to the right. For all stimuli, if the texture within the figure produces more overall visual response than the background in one frame, it produces less than the background in the next frame and so on. Panels (a) and (b) illustrate the actual appearance of the stimuli; panel (c) is more schematic. Stimulus (c) consist of vertical lines, somewhat like (b), except that the lines change their contrast from black to white according to the color in the time axis, *t*. (Adopted from Figure 6 in Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics, 61*(6), 1075–1088.)

so these stimuli, too, become reversed-phi stimuli for the second-order system.

Figure IV.71-1c represents flickering stimuli. The vertical axis represents time. All the stimuli are composed entirely of vertical stripes; the *x, t* graph represents the color of a stripe at a particular instant in time. For the first- and third-quarter cycles of the stimulus (rows 1 and 3), the bars represent regions with a maximum rate of pixel flicker (a black–white reversal of every pixel on every frame), superimposed on a background that has a random, medium rate of flicker. In the second- and fourth-quarter cycles, bars remain unchanged during their entire exposure (vertical stripes in panel c, low flicker rate regions). The translating bars displace one bar width to the right, switching from high to low flicker and vice versa from frame to frame.

The second-order reversed phi stimuli of Figure IV.71-1 were shown to subjects at rates between 0.94 and 15 Hz (Lu & Sperling, 1999). In peripheral vision, every observer saw presentations only in the reversed direction (opposite to the bar displacement). In some conditions—for example, at the highest temporal frequencies of the flicker modulation stimuli—there were some trials in which no motion was perceived. When motion was perceived, however, it was universally in the direction of second-order reversed phi.

In central viewing, observers reported perceiving motion in the forward direction at low temporal frequencies and in the reversed direction at high temporal frequencies. The temporal frequency at which the direction of perceived motion changed depended on the stimulus but was similar for all observers. For the contrast modulation stimulus (Fig. IV.71-1a), all observers perceived motion in the reversed direction at 15 Hz and motion in the forward direction at temporal frequencies of 0.94, 1.88, and 3.75 Hz; 7.5 Hz was ambiguous for all observers. For the spatial frequency modulation stimulus (Fig. IV.71-1b), all observers reported forward motion at temporal frequencies of 0.94 and 1.88 Hz and reversed motion at 3.75 and 7.50 Hz. For the flicker modulation stimulus (Fig. IV.71-1c), at temporal frequencies of 0.94 and 1.88 Hz, all observers reported perceiving motion in the forward direction. At the highest resolvable temporal frequency for these stimuli (3.75 Hz), all observers reported perceiving motion in the reversed direction. Moving the observer away from the displays had the same effect as changing from central viewing to periphery viewing: reversed motion became more dominant.

To understand second-order reversed phi, we start with a theory on first-order reversed phi. Figure IV.71-2a shows four frames of a first-order reversed-phi stimulus (Chubb & Sperling, 1989) in which the contrast of the bars alternates polarity (from white to black and from black to white) from frame to frame. Applying motion energy analysis directly to raw stimulus contrast predicts a motion direction that is opposite to the bar displacement (reversed phi) because the dominant Fourier component (indicated by the light and the dark stripes of a superimposed sine-wave grating) is oriented in the leftward direction (Fig. IV.71-2b). When viewing in the periphery or from afar, human observers perceive this so-called reversed motion direction (as predicted by motion energy analysis). In central vision, humans perceive motion in the forward direction, contradicting what is predicted by motion energy analysis. This was explained by a model in which stimulus contrast is first full-wave rectified (absolute value computation) and only then subjected to motion energy analysis (Fig. IV.71-3). After rectification, the dominant Fourier component is in the forward direction (indicated by the superimposed sine-wave grating, Fig. IV.71-2c).

A second-order motion energy model consists of a texture grabber followed by a motion energy detector (Fig. IV.71-3). The texture grabber consists of a spatiotemporal linear filter. The spatial component is represented as a cosine receptive field (which is primarily sensitive to a particular orientation and spatial frequency), and the temporal component is a bandpass filter that is relatively insensitive to very low and very high temporal frequencies. Full-wave rectification (e.g., absolute value or square law rectifier; absolute value is shown) gives the texture grabber an output that depends only on the total quantity of texture in its preferred orientation and frequency, independent of the sign of the outputs of the linear filters, which may be either positive or negative. Motion energy analysis determines the direction of motion of the texture detected by the texture grabber.

The second-order reversed phi stimuli in Figure IV.71-1 are preprocessed by a texture grabber that consists of a low-pass spatial filter and a band-pass temporal filter followed by full-wave rectification (Fig. IV.71-3). The texture grabber

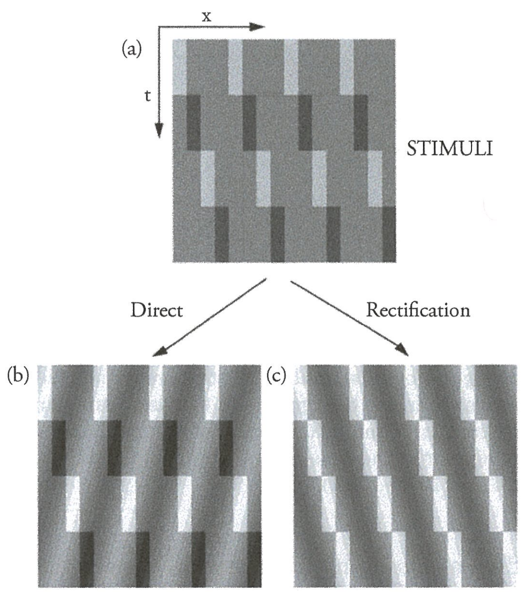


Figure IV.71-2. First-order reversed-phi: stimulus and theory. The horizontal axis represents space; the vertical axis represents time. (a) A reversed-phi square-wave stimulus. (b) The fundamental sine wave from a Fourier analysis of the stimulus (which is the primary stimulus to the motion system) is shown superimposed on the stimulus, illustrating that it signals a direction opposite to the direction in which figure moves. (c) The second-order motion system is indifferent to the sign of contrast, so all stimulus contrasts are shown as equal and positive. The third-order system is sensitive to the figure, so all areas of the figure are shown as positive. The superimposed fundamental sine wave illustrates both higher-order motion systems' report of motion in the direction of the figure. (After Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088, with permission.)

produces an output that is equivalent to a first-order reversed-phi stimulus (Fig. IV.71-4). Rectification alone would have sufficed to reveal the second-order motion of the texture contrast modulation stimulus in Figure IV.71-1a. However, rectification must be preceded by low-pass spatial filtering to expose the motion of the three grating frequencies stimulus of Figure IV.71-1b. Temporal band-pass or temporal low-pass filtering followed by rectification (Chubb & Sperling, 1988) is required to expose the motion of the flicker stimulus of Figure IV.71-1c. Figure IV.71-4 illustrates that, after preprocessing, the dominant motion energies of the reversed-phi stimuli of Figures IV.71-1a, IV.71-1b, and IV.71-1c are in the leftward direction, opposite to the rightward displacement of the bars. The motion that is extracted from second-order reversed-phi stimuli by a texture grabber, followed by a motion energy computation, would be reversed motion.

Because the second-order reversed-phi stimuli are invisible to first-order analysis, and because second-order analysis produces motion in the reversed direction, the perception of forward motion requires yet another mechanism, presumably the third-order salience map mechanism proposed by Lu and Sperling (1995a, 1995b, 2001). The three reversed-phi stimuli in our experiments were constructed so that one-fourth of the area was occupied by the figure, which changed type from frame to frame, and three-fourths of the area was occupied by the background, which remained physically unchanged from frame to frame in the contrast modulation stimuli and unchanged in type for the spatial frequency and flicker stimuli. The third-order motion mechanism is assumed by Lu and Sperling (1995a, 1995b) to compute the moment-to-moment motion of those parts of the visual stimulus that are marked as figure in a neural salience field. So if (a) occupying a small part of the field and changing from frame to frame were to

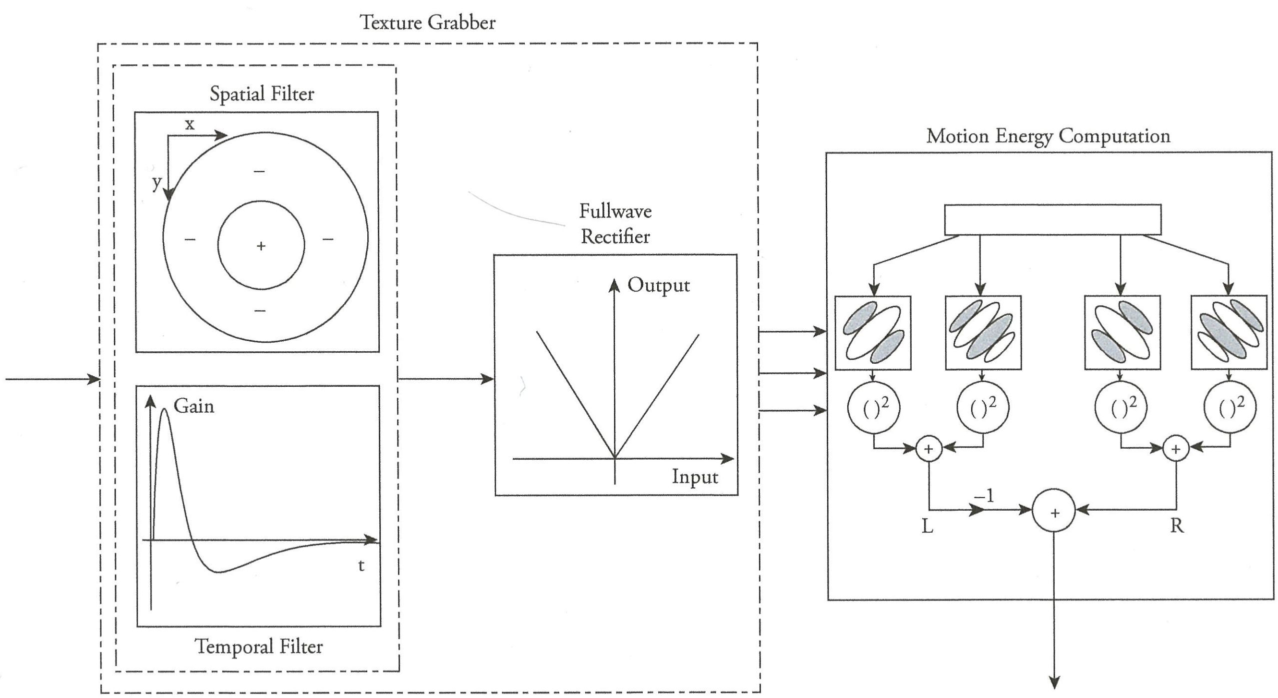


Figure IV.71-3. A second-order motion energy model. A texture grabber is followed by a motion energy detector. A texture grabber consists of a two-dimensional bandpass spatial filter and a one-dimensional bandpass temporal filter followed by rectification. The motion energy computation is tuned to optimally compute motion power in a three-dimensional spatiotemporal frequency band. (After Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088, with permission.)

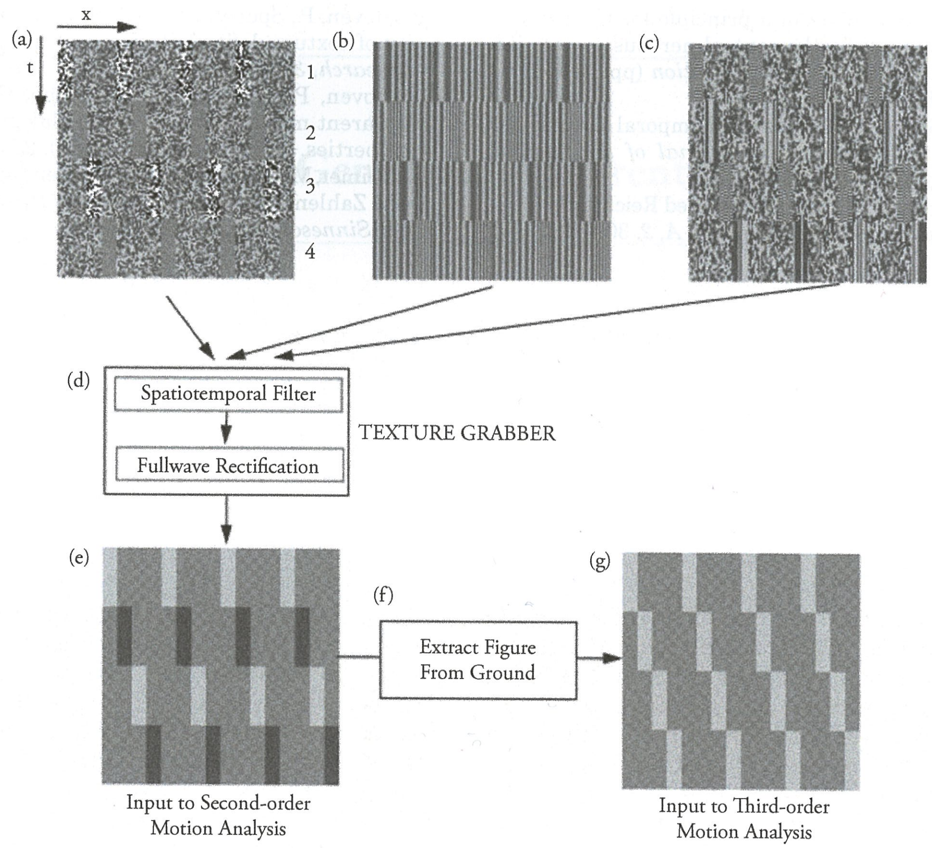


Figure IV.71-4. (a) through (c) Second-order reversed phi stimuli. (d) Spatiotemporal filtering and full-wave rectification (see Fig. IV.71-3). (e) The stimuli, as transformed by (d), can be represented as a space–time graph. For stimuli (a), (b), and (c), (e) can also be interpreted as representing four consecutive frames of texture grabber output. The dominant motion energy in panel (e) is leftward (opposite to the bar displacement direction). Applying motion energy analysis on the output of the texture grabber would predict motion in the reversed direction. (f) Salience computation. (g) A salience system that marked the location of features that differed from the background in each frame would compute motion in the forward (feature) direction. (After Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088, with permission.)

cause that portion of the stimulus to be marked as figure, and (b) occupying a large part of the field and remaining stable were to cause that portion of the stimulus to be marked as ground, and (c) the third-order motion system computed the motion of figure, then the second-order reversed-phi stimuli would support third-order motion in the forward direction. This is precisely what occurred.

The third-order motion system has a much lower cutoff temporal frequency (≈ 3–4 Hz) than does the second-order motion system (≈ 10–12 Hz; Lu & Sperling, 1995a). Similarly, the higher the order of the motion system, the lower its spatial resolution and, therefore, the more likely it will be confined to areas near the fovea, where spatial frequency resolution is better. These results—that the third-order motion direction was (a) never perceived in peripheral viewing and (b) perceived in foveal areas only at low temporal frequencies—are entirely consistent with the previously noted properties of the second- and third-order motion systems.

REFERENCES

Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, *10*(12), 1411–1415.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, *15*(8), 957–956.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli- A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, *5*, 1986–2007.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences USA*, *86*(8), 2985–2989.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences USA*, *86*(23), 9631–9635.

Lu, Z.-L., & Sperling, G. (1995a). Attention-generated apparent motion. *Nature*, *377*(6546), 237–239.

Lu, Z.-L., & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088.

Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A*, *18*(9), 2331–2370.

Nishida, S. Y. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research*, *33*(5), 633–645.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). Cambridge, MA: MIT Press.

van Santen, J. P., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*(5), 451–473.

van Santen, J. P., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*, 300–321.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, *33*(4), 463–485.

Werkhoven, P., Sperling, G., & Chubb, C. (1994). Perception of apparent motion between dissimilar gratings: Spatiotemporal properties. *Vision Research*, *34*(20), 2741–2759.

Wertheimer, M. (1912). Über das Denken von Naturvölker: 1. Zahlen und Zahlengebilde. *Zeitschrift fuer Psychologie und Physiologie der Sinnesorgane*, *60*, 321–389.