



HAL
open science

Gestalt factors modulate basic spatial vision.

Bilge Sayim, G Westheimer, M H Herzog

► **To cite this version:**

Bilge Sayim, G Westheimer, M H Herzog. Gestalt factors modulate basic spatial vision.. Psychological Science, 2010, Psychological Science, 21 (5), pp.641-4. 10.1177/0956797610368811 . hal-03565507

HAL Id: hal-03565507

<https://hal.univ-lille.fr/hal-03565507>

Submitted on 11 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gestalt Factors Modulate Basic Spatial Vision

B. Sayim¹, G. Westheimer², and M.H. Herzog¹

¹Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), and ²Department of Molecular and Cell Biology, University of California, Berkeley

Abstract

Human perception of a stimulus varies depending on the context in which the stimulus is presented. Such contextual modulation has often been explained by two basic neural mechanisms: lateral inhibition and spatial pooling. In the present study, we presented observers with a vernier stimulus flanked by single lines; observers' ability to discriminate the offset direction of the vernier stimulus deteriorated in accordance with both explanations. However, when the flanking lines were part of a geometric shape (i.e., a good Gestalt), this deterioration strongly diminished. These findings cannot be explained by lateral inhibition or spatial pooling. It seems that Gestalt factors play an important role in contextual modulation. We propose that contextual modulation can be used as a quantitative measure to investigate the rules governing the grouping of elements into meaningful wholes.

Keywords

spatial vision, perceptual grouping, verniers, crowding, lateral masking, surround inhibition, object superiority

Contexts can strongly modulate the perception of visual stimuli. In the tilt illusion, for example, observers perceive a vertical grating as tilted when it is embedded in an oblique grating (Fig. 1a). In crowding, observers' ability to recognize a letter in their peripheral vision deteriorates when the letter is flanked by other letters (Fig. 1b). In vernier discrimination, observers have to indicate whether the lower of two vertical lines is displaced to the left or to the right of the upper line (Fig. 1c, top). Observers' performance deteriorates significantly when the vernier stimulus is flanked by two lines (Fig. 1c, bottom), compared with when it is unflanked (Levi, Klein, & Aitsebaomo, 1985; Westheimer & Hauske, 1975). Such contextual modulation is central to vision science because it exemplifies the difference between local physical stimulation on the one hand and perceptual and neural responses on the other hand.

Often, contextual modulation is explained by two basic neural mechanisms: lateral inhibition and spatial pooling (Fig. 2). Propositions based on lateral inhibition suggest that contextual flankers decrease neural activity related to the target when the flankers fall on the inhibitory region of receptive fields of neurons dedicated to target processing. Lateral inhibition has been invoked to explain why flanking lines surrounding a vernier stimulus decrease vernier related neural activity (Figs. 2a and 2b; Westheimer & Hauske, 1975).

Lateral inhibition has also been used to explain the tilt illusion (Solomon, Felisberti, & Morgan, 2004) and the modulation of motion perception (Tadin, Lappin, Gilroy, & Blake, 2003).

Propositions based on spatial pooling posit that information from the target and the flankers is pooled and then averaged; consequently, the neural activity corresponding to the target is reduced. Badcock and Westheimer (1985) have proposed spatial pooling as the mechanism responsible for the deterioration of performance when a vernier stimulus is presented with flanking lines (Fig. 2c). Spatial pooling is also assumed to underlie many other contextual effects, such as those observed in letter identification (Fig. 1b; e.g., Pelli, Palomares, & Majaj, 2004), orientation discrimination (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), and contrast discrimination (Wilkinson, Wilson, & Ellemberg, 1997).

Propositions based on perceptual organization suggest that perceptual grouping is crucial in contextual modulation of foveal stimuli (Malania, Herzog, & Westheimer, 2007). Flankers strongly deteriorate performance on a target when flankers

Corresponding Author:

Bilge Sayim, Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
Email: bilge.sayim@epfl.ch

and target form one perceptual group, as, for example, when they have the same color (Sayim, Westheimer, & Herzog, 2008). Elsewhere, we have proposed that dynamic lateral inhibition could underlie these kinds of grouping effects (Hermens, Luksys, Gerstner, Herzog, & Ernst, 2008; Herzog, Ernst, Etzold, & Eurich, 2003).

In the present study, we investigated contextual modulation using vernier stimuli flanked by different configurations of lines and shapes. We hypothesized that the deleterious influence of single flanking lines would be reduced if the lines were integrated into geometric configurations distinct from the vernier stimulus. Indeed, we found that integrating single flanking lines into rectangles or perspective cuboids significantly improved observers' performance. This finding cannot be explained by lateral inhibition or spatial pooling.

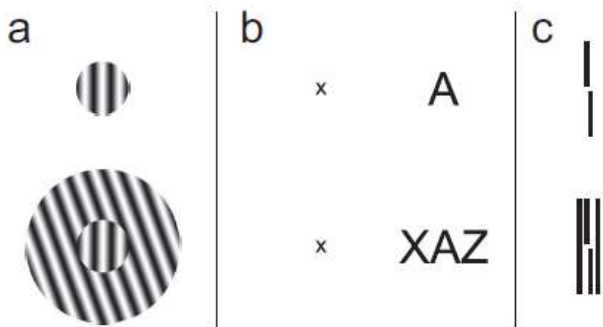


Fig. 1. Three examples of contextual modulation. In the tilt illusion (a), observers perceive a vertical grating as tilted when it is embedded in an oblique grating. Although observers can easily recognize a letter in their peripheral vision (b; the x represents the fixation point), their recognition is impaired when the letter is flanked by other letters (crowding). Observers' offset discrimination of a vernier stimulus deteriorates (c) when the stimulus is flanked by additional lines. (You may need to look at the figure from a distance to experience this effect; see also Fig. 2).

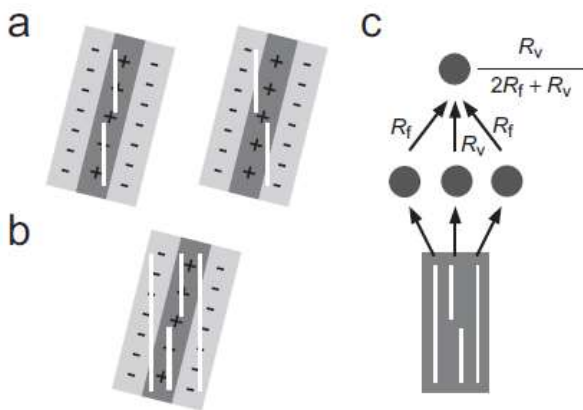


Fig. 2. Proposed explanations for the effect of flankers on vernier offset discrimination. The illustration in (a) shows left and right offset vernier stimuli superimposed on the receptive field of a neuron tuned to an oblique orientation. The left offset vernier stimulus (shown on the left) falls on the region of the receptive field that excites the neuron (plus signs), whereas the right offset vernier stimulus (shown on the right) also falls on the region of the receptive field that inhibits the neuron (minus signs). Neurons with such receptive fields may be involved in discrimination of vernier offset because the left offset vernier stimulus covers a larger area of the excitatory region than the right offset vernier stimulus does (Wilson, 1986). According to the lateral inhibition proposition (b), the addition of flanking lines impairs discrimination performance because these lines fall on the inhibitory region of the receptive field and that inhibits the excitatory response to the vernier stimulus. According to the spatial pooling proposition (c), the vernier stimulus and flanking lines each feed into separate neurons. Responses from these neurons are pooled, which reduces the relative signal from the vernier stimulus. (R_f and R_v denote responses to the flanking lines and the vernier stimulus, respectively).

Method

Participants

Ten observers (3 females, 7 males; ages 19–31 years, mean age = 23 years) participated in the experiment. All participants were students at the École Polytechnique Fédérale de Lausanne (EPFL) or the Université de Lausanne and were paid for participation. They were not told the purpose of the experiment. All observers had normal or corrected to normal visual acuity.

Stimuli

Stimuli were displayed against a black background on an oscilloscope (HP1332A XY with a P11 phosphor; Hewlett Packard, Palo Alto, CA). The stimuli had a luminance of 80 cd/m². Observers viewed the monitor from a distance of 2 m. Observers were presented with five different stimulus configurations (Fig. 3). The first configuration, an unflanked vernier stimulus, served as a baseline. The vernier stimulus consisted of two vertical 10 arcmin long lines separated by a vertical gap of 1 arcmin.

In the second configuration, the vernier stimulus was flanked on each side by one line of the same length as the vernier stimulus (21 arcmin). The distance between these two flanking lines and the vernier stimulus was adjusted for each observer to optimize interference with vernier offset discrimination. We started with a distance of 200 arcsec. If the vernier offset threshold (i.e., the smallest offset size observers can discriminate with 75% accuracy) was not at least 2 times the threshold for the unflanked vernier stimulus, we reduced the distance between the flankers and the vernier stimulus by 50 arcsec. If the threshold was more than 10 times the unflanked vernier threshold, we increased the distance by 50 arcsec. If the threshold was still more than 10 times the unflanked vernier threshold, a distance of 300 arcsec was used. Distances varied as follows—150 arcsec: 1 observer; 200 arcsec: 6 observers; 250 arcsec: 2 observers; and 300 arcsec: 1 observer. The following configurations contained these two flankers (with the individually adjusted vernier to line distances). In the third configuration, each of the two flankers was complemented to form a separate rectangle (width = 46.6 arcmin each). In the fourth configuration, five more lines were added to each rectangle to create perspective cuboids (angle of oblique lines = 135°, length = 16.6 arcmin). In the fifth configuration, all the lines in the cuboids, except the original two single flanking lines, were scrambled to obliterate the cuboids' Gestalt. All flanking configurations were symmetrical and centered on the vernier stimulus.

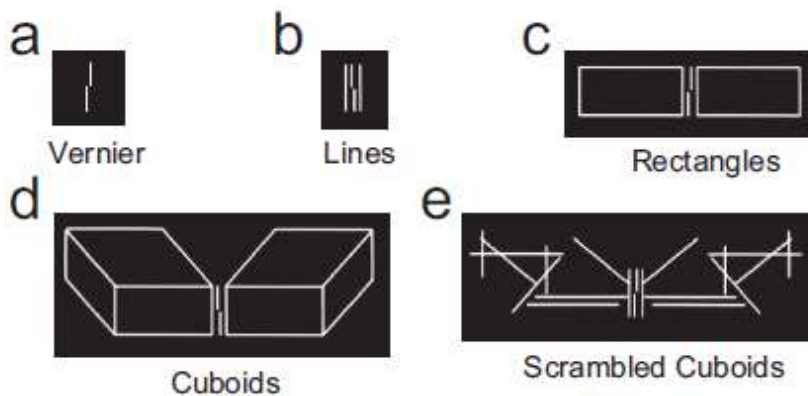


Fig. 3. Stimuli used in the experiment. We presented participants with (a) a vernier stimulus without flankers, (b) a vernier stimulus flanked by two lines, (c) a vernier stimulus flanked by two rectangles, (d) a vernier stimulus flanked by two cuboids, and (e) a vernier stimulus flanked by two scrambled cuboids. The two flanking lines were part of all configurations except (a). The scrambled cuboids contained the same elements as the cuboids. Stimulus pictures are achromatic renderings of photographs of the experimental screen.

Procedure

In each trial, the two lines of the vernier stimulus were randomly offset either to the left or to the right. A block consisted of 80 trials of a single stimulus configuration (e.g., rectangles), and each block had a different pseudorandom sequence of left and right offsets. There were equal numbers of left and right offsets within each block. The observers' task was to indicate the offset direction by pressing a button. Presentation time was 150 ms. After the stimulus disappeared, the screen remained blank for a maximum period of 3,000 ms, during which time the observer was required to respond. After the response, the screen remained blank for 500 ms until the next trial. Errors and omissions were indicated by auditory feedback.

The order of conditions was randomized for each observer. All conditions were presented twice. To compensate for possible learning effects, we reversed the order of conditions after each condition had been presented once.

An adaptive staircase procedure was used to determine the threshold at which an observer reached 75% correct responses. The starting offset was 75 arcsec. A cumulative Gaussian function was fitted to the data, and thresholds were determined using pro-bit and likelihood analyses. The data were analyzed using t tests with Bonferroni correction for multiple comparisons. The adjusted alpha level was .01 per test (.05/5).

Results

As Figure 4 shows, performance was best (i.e., thresholds were lowest) when the vernier stimulus was unflanked ($M = 12.23$ arcsec). When the vernier stimulus was flanked by single lines, performance deteriorated significantly ($M = 41.44$ arcsec), $t(9) = 6.30$, $p < .001$ (Bonferroni-corrected $\alpha = .01$). However, when additional lines were added to each of the single flanking lines to form rectangles and cuboids, observers' performance improved significantly over their performance when the flankers were single lines (rectangles: $M = 22.90$ arcsec; cuboids: $M = 25.0$ arcsec), $t(9) = 6.14$, $p < .001$, and $t(9) = 5.93$, $p < .001$, respectively.

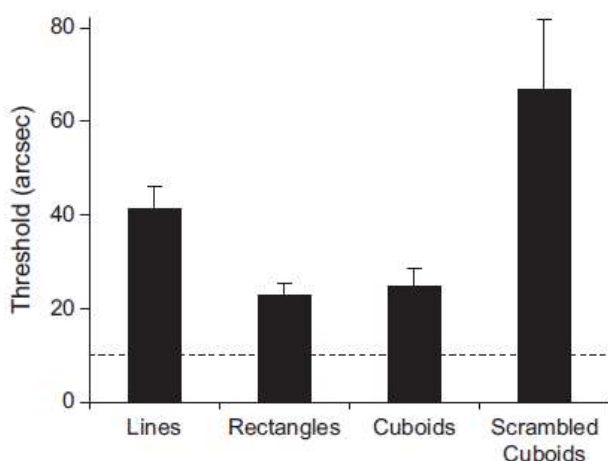


Fig. 4. Vernier offset discrimination in the five conditions. The graph shows the mean 75% threshold (i.e., the smallest offset the observer could discriminate with 75% accuracy) across the 10 observers. The dashed horizontal line marks the threshold for an unflanked vernier stimulus. The bars show results for the other four conditions: the vernier stimulus flanked by two lines, by two rectangles, by two cuboids, and by scrambled cuboids. Error bars indicate standard errors of the mean.

The fifth configuration (the scrambled cuboids) was used as a control condition. The scrambled cuboids contained the same elements as the unscrambled cuboids and had a similar spatial extent; thus, low level properties, such as overall luminance and size, were similar in the two conditions. However, observers' performance with scrambled cuboids was worse than their performance in any other condition ($M = 66.8$ arc sec). We found a significant difference between performance with the scrambled and the unscrambled cuboids, $t(9) = 3.32$, $p = .009$. Moreover, there was a trend toward stronger deterioration of performance with scrambled cuboids than with the single flanking lines, $t(9) = 2.06$, $p = .069$.

The deterioration of performance with single flanking lines and scrambled cuboids (relative to the unflanked condition) can be easily explained by lateral inhibition or spatial pooling (Fig. 2). However, the better performance with the unscrambled cuboids than with the scrambled cuboids cannot be explained by these mechanisms, because the unscrambled and scrambled cuboids contained exactly the same elements. We attribute the reduced flanker interference (i.e., increased observer performance) in the unscrambled cuboid condition to the cuboids' good Gestalt (see also Weisstein & Harris, 1974). The Fourier spectra in the two conditions were very similar.

Conclusion

We propose that perceptual grouping plays a key role in contextual modulation (see also Banks & Prinzmetal, 1976; Banks & White, 1984; Herzog & Fahle, 2002; Livne & Sagi, 2007; Malania et al., 2007; Saarela, Sayim, Westheimer, & Herzog, 2009). We have shown that the deleterious influence of flanking lines is reduced when they are integrated into a figure with a good Gestalt, such as a rectangle or a cuboid. We attribute this effect to the perceptual ungrouping of the flanking lines from the vernier stimulus when the lines are complemented to form part of a figure. Hence, genuine Gestalt factors modulate vernier offset discrimination. Further research will be needed to unravel the neural mechanisms of grouping and Gestalt processing in contextual modulation, which may still be found at an early level of visual processing. These mechanisms may consist of neural circuits dedicated specifically to processing certain Gestalt or figurativeness operations. Alternatively, the influence of Gestalt factors may be mediated by top down processes that modulate early visual processing.

We propose that vernier offset discrimination with different flanking configurations is a good tool for investigating Gestalt laws. Since grouping of pattern elements into configurations (Gestalten) was first proposed by Wertheimer in 1923, the on-going search for the rules governing such grouping processes has been only moderately successful, in part because of the lack of appropriate research tools. The methodology of the present study yielded a quantitative measure of what Wertheimer called the "goodness" of a configuration (Wertheimer, 1923). We therefore suggest that Gestalt laws may be quantified by measuring vernier offset discrimination in different contexts.

Acknowledgments

We thank M. Repnow for technical support.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by the Swiss National Science Foundation project "The Dynamics of Feature Integration" and Pro*Doc "Processes of Perception."

References

- Badcock, D.R., & Westheimer, G. (1985). Spatial location and hyper acuity: The centre/surround localization contribution function has two substrates. *Vision Research*, 25, 1259–1267.
- Banks, W.P., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception & Psychophysics*, 19, 361–367.
- Banks, W.P., & White, H. (1984). Lateral interference and perceptual grouping in visual detection. *Perception & Psychophysics*, 36, 285–295.
- Hermens, F., Luksys, G., Gerstner, W., Herzog, M.H., & Ernst, U. (2008). Modeling spatial and temporal aspects of visual backward masking. *Psychological Review*, 115, 83–100.

- Herzog, M.H., Ernst, U., Etzold, A., & Eurih, C. (2003). Local interactions in neural networks explain global effects in Gestalt processing and masking. *Neural Computation*, *15*, 2091–2113.
- Herzog, M.H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature*, *415*, 433–436.
- Levi, D.M., Klein, A.S., & Aitsebaomo, A.P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*, 963–977.
- Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, *7*(2), Article 4. Retrieved April 4, 2008, from <http://journalofvision.org/7/2/4/>
- Malania, M., Herzog, M.H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision*, *7*(2), Article 1. Retrieved February 13, 2007, from <http://journalofvision.org/7/2/1/>
- Parkes, L., Lund, J., Angelucci, A., Solomon, J.A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*, 739–744.
- Pelli, D.G., Palomares, M., & Majaj, N.J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12), Article 12. Retrieved March 27, 2006, from <http://journalofvision.org/4/12/12/>
- Saarela, T.P., Sayim, B., Westheimer, G., & Herzog, M.H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, *9*(2), Article 5. Retrieved February 6, 2009, from <http://journalofvision.org/9/2/5/>
- Sayim, B., Westheimer, G., & Herzog, M.H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision*, *8*(8), Article 12. Retrieved June 30, 2008, from <http://journalofvision.org/8/8/12/>
- Solomon, J.A., Felisberti, F.M., & Morgan, M.J. (2004). Crowding and the tilt illusion: Toward a unified account. *Journal of Vision*, *4*(6), Article 9. Retrieved June 6, 2009, from <http://journalofvision.org/4/6/9/>
- Tadin, D., Lappin, J.S., Gilroy, L., & Blake, R. (2003). Perceptual consequences of center-surround antagonism in visual motion processing. *Nature*, *424*, 312–315.
- Weisstein, N., & Harris, C.S. (1974). Visual detection of line segments: An object superiority effect. *Science*, *186*, 752–755.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt; II. *Psychologische Forschung*, *4*, 301–350.
- Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research*, *15*, 1137–1141.
- Wilkinson, F., Wilson, H.R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *14*, 2057–2068.
- Wilson, H.R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research*, *26*, 453–46