

Gestalt rules the corner of your eye

Samba Silla

Stuyvesant High School
345 Chambers Street
New York, NY 10282

Mentor: Denis G. Pelli
Professor of Psychology and Neural Science
Department of Psychology
New York University
New York, NY 10003

Advisor: Anne de Sostoa Manwell
Stuyvesant High School
345 Chambers Street
New York, NY 10282

S.Silla (2003) *Gestalt rules the corner of your eye*. Intel Science Talent Search.
Semifinalist. <http://psych.nyu.edu/pelli/#intel>

LAYMAN'S ABSTRACT

We continually recognize objects without apparent effort. We don't know how the brain does it and we can't get machines to do it. Object recognition in the peripheral visual field is more limited and simpler than in central vision. Hoping to make faster progress by studying this simpler system, we tested the role of Gestalt laws in recognizing objects in central and peripheral vision.

ABSTRACT

We continually recognize objects without apparent effort. We don't know how the brain does it and we can't get machines to do it. Form vision in the peripheral visual field is more limited and simpler than in central vision. Perhaps we can make faster progress by studying this simpler system. We tested the role of Gestalt laws in recognizing objects in central and peripheral vision. We measured efficiency for identifying objects as a function of position in the visual field for letters of simple and complex fonts. We also made letters out of strings of striped patches to test the Gestalt law of good continuation. We find that the effect in the periphery is twice as powerful as in the fovea. Studying the periphery allows us to isolate the effects of the Gestalt laws on letter identification.

INTRODUCTION

How people recognize objects is a great mystery. We lack even a tentative explanation. Yet, our visual system excels at object recognition. At any moment, there are innumerable ways that we could interpret what we're seeing. Yet, under normal viewing conditions, we are quite accurate. In central vision, there seem to be several competing feature-integration mechanisms that allow us to effortlessly perform this task. Because several mechanisms are available, the effect of parametric stimulus manipulations may reveal the sensitivity and tuning of a particular mechanism, or provoke a switch to a different mechanism, and it is hard to distinguish the two cases. Studying the periphery proves invaluable with respect to that problem since the periphery contains fewer integration mechanisms than the fovea and thus allows a purer study of each mechanism.

The point in your visual field that you are looking at is called the "fixation" point. Around that point, the central 1 degree (diameter) of visual field is called the "fovea". The rest of our visual field is called the "periphery". Put more simply, the periphery is anywhere that you are not looking at directly but remains in your vision, i.e. in the corner of your eye. A finger is roughly 1° wide at arm's length, so you can see that the fovea represents only a tiny fraction of your visual field. *Eccentricity* refers to how many degrees of visual angle away from the fixation point (center of the fovea) an object is.

We detect motion more efficiently in the periphery than in the fovea, but we are less adept there at identifying objects, often by a significant factor. There is evidence that the periphery may be less efficient at identifying objects largely because it lacks some of the feature-integration mechanisms that are present in the fovea.

A phenomenon referred to as crowding occurs in the periphery. Crowding causes the observer to perceive the features of separate objects as jumbled together. They see objects that are amalgams of the features of the individual objects and are very difficult to identify. Crowding occurs between objects and also within single complex objects. Crowding occurs because we have fewer small integration fields in the periphery. *Critical spacing* is the minimal spacing needed between features in the periphery for crowding not to occur. The critical spacing (in deg of visual angle) turns out to be roughly half the eccentricity (Bouma, 1972). When features of objects in the periphery are closer than the critical spacing, they occupy the same integration field and the visual system integrates them into one perceived object, which may be hard to identify.

One reason that it is important to study the periphery is because of visual impairments like macular degeneration that selectively impair central vision, forcing people to rely more on peripheral vision. Many tasks, like reading, are much harder in the periphery than in the fovea, but current understanding of peripheral vision is very limited. Age-related macular degeneration is the most common cause of vision loss in those over 50 and its prevalence increases with age. It is caused by degeneration of the macula, a region including and slightly larger than the fovea, which is the central, and most sensitive part of the retina at the back of the eye. Macular degeneration is caused by hardening of the arteries that nourish the retina. This deprives the sensitive retinal tissue of oxygen and nutrients that it needs to function and thrive. As a result, these patients lose central vision and retain only peripheral vision. A greater understanding of the periphery might lead to new rehabilitation strategies for patients with central field loss.

Gestalt theory suggests that we perceive an object as a whole rather than just the sum of its parts. Wertheimer and other Gestalt psychologists determined that we tend to perceive the whole of patterns, objects, and scenes, and to give less emphasis to the smaller parts of which they are composed (Ellis, 1938). Wertheimer proposed that the nervous system is innately predisposed to group incoming sensory elements according to certain principles or *factors* of perceptual grouping. Those principles lay down specific rules according to which we group features of a stimulus. A few of those principles follow: The *Factor of Proximity* states that we tend to see stimulus elements that are near each other as part of the same object. The *Factor of Similarity* states that we tend to see stimulus elements that are similar as part of the same object. *The Factor of Direction (or Good Continuation)* observes that when lines intersect we tend to group the line segments in such a way as to form continuous lines with minimal change in direction. *The Factor of Closure* indicates that we tend to see forms as completely enclosed by a border and to ignore gaps in the border. Less tangible than those principles, and underlying all of Gestalt theory, is the principle of good form or *Prägnanz*, which indicates that the perceptual system strives to produce percepts that are elegant — simple, uncluttered, symmetrical, and regular.

Research on Gestalt Principles

Field, Hayes, and Hess (1993) ran experiments on paths, also referred to as “snakes,” consisting of a string of striped patches (gabor) in a field of randomly oriented gabors. The only thing distinguishing the snake from the background is that its gabors are aligned with an invisible path. Gabors are good stimuli for psychometric studies because

they are thought to isolate single feature-detectors in the visual system. A gabor consists of a sinusoidal grating vignetted by a gaussian window (Fig. 1). It is a striped patch with soft edges. Field, Hayes, and Hess found that observers correctly identified which field of gabors contained the snake if the angular difference between successive gabors in the path was less than 30° . They also discovered that the paths become more difficult to identify as the gabors that make up the path go from being collinear to the path to being offset by 90° . This indicates that relative orientation plays a key role in feature integration.



Figure 1. (A) A horizontally-oriented gabor. (B) A vertically-oriented gabor .

Majaj, Raizman, Christian, Kim, Palomares, and Pelli (in prep.) did a series of experiments examining the validity of the Gestalt *Factor of Good Continuation*. Drawing on the Field, Hayes, and Hess (1993) experiments, Majaj et al. used gabor snakes to test threshold contrast for identification as a function of perturbation of grating orientation. Desiring a more ordinary object than an arbitrary string of gabors, Majaj et al. used the string of gabors to draw letters (snake letters). Snake letters are made by placing a string of gabors along the path of the letter. Such letters are characterized by the angle of the offset of the gabor from the path of the original letter. When the gabors are collinear to the path of the letter, the snake letter is “collinear”. When the gabors are offset by 90° , the letter is “orthogonal”. As the letters go from collinear to orthogonal,

they begin to violate the *Factor of Good Continuation*. Majaj et al. showed that efficiency at identifying snake letters decreased in the fovea as the degree tilt of the letters increased from 0° to 90° .

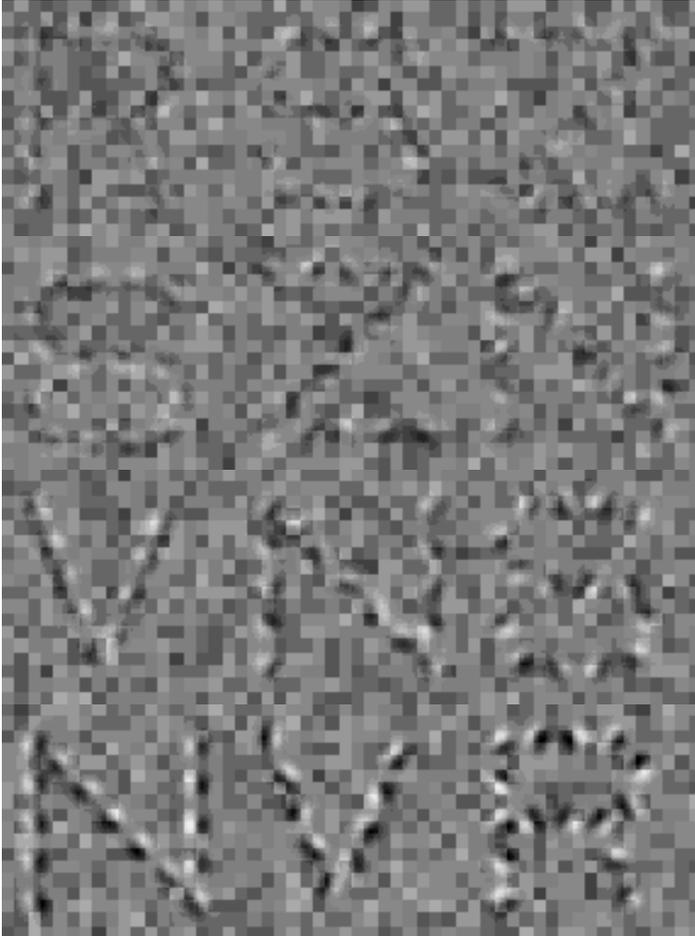


Figure 2. Snake letters in noise, demonstrating that good continuation is important for letter identification. The orientation of the gratings relative to the letter stroke alternates $\pm 0^\circ$, $\pm 30^\circ$, or, $\pm 60^\circ$ (left to right). Starting from the bottom, read up each column as far as you can. The height of the faintest identifiable letter is your contrast sensitivity for such letters. Note that it becomes progressively more difficult to see the 30° and 60° as compared to the 0° condition in equivalent noise (Majaj et al., in prep.).

The demonstration in Fig. 3 shows that it is much harder to see orthogonal than collinear snake letters in the periphery. We measured snake letter identification as a function of deviation from collinearity in the periphery. We wanted to test how the individual Gestalt laws would affect ability to identify stimuli across the visual field. We

picked stimuli that would allow us to compare a stimulus that violated a particular Gestalt principle versus one that did not. Majaj et al. (in prep.) were successful in studying the *Factor of Good Continuation* in the fovea using the snake letters as stimuli. Thus, we chose to use snake letters as the stimuli to test the *Factor of Good Continuation* in the periphery (Fig. 3). In order to study *Prägnanz*, we used Sloan and Künstler letters as stimuli (Fig. 4). Sloan letters are a simple, bold, sans-serif font. A serif is a small line, curve, flourish, or embellishment projecting from the main stroke of a letter. On the other hand, Künstler letters have many features and serifs, to the point where they violate the principle of *Prägnanz*.

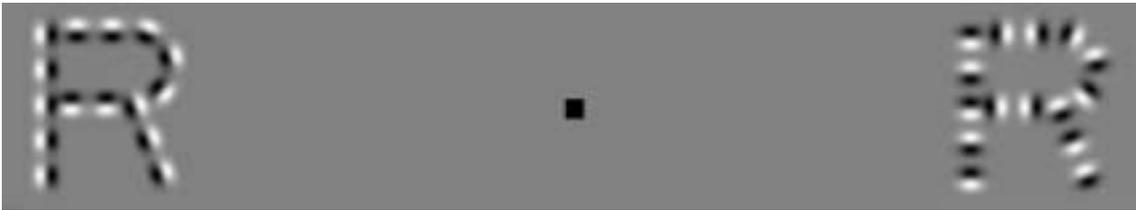


Figure 3. Snake letters. The letter to the left of the fixation point has gabors collinear to the path of the letter; the one to the right has gabors orthogonal to the path of the letter. Fixate on the black dot and observe the collinear and orthogonal letters. Note that it is much harder to identify the orthogonal snake letter than the collinear one. This illustrates that the violation of the *Factor of Good Continuation* affects one's ability to identify a stimulus in the periphery.



Figure 4. Sloan & Künstler letters. The letter to the left is in the Sloan font. The letter to the right of the fixation point is in the Künstler font. Fixate on the black dot and observe the Sloan and Künstler letters. Note that it is much harder to identify the Künstler letter than the Sloan one. This illustrates that the violation of the principle of *Prägnanz* affects one's ability to identify a stimulus in the periphery.

If an observer was less able to identify the Künstler letters than the Sloan letters, it would reveal that the effect was neither an artifact created by the snake letter stimulus nor isolated solely to snake letters. Instead, the implication would be that *Prägnanz* plays an important role in object recognition over the entire range of the visual field.

METHODS

Observers

The observer was a high school student (the author) with normal acuity. For the alphabets used in this task, the observer practiced until well trained in the task (5,000 trials).

Stimuli

The letters and characters were drawn onto the screen using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). PostScript fonts were used. PostScript fonts contain geometric descriptions that allow good rendering of letters at any size. PostScript is rendered by Adobe Type Manager. KuenstlerScriptTwoBold and Sloan are PostScript fonts. The Sloan font was created in the font-drawing program Fontographer. The shapes of the Sloan characters are specified by the NAS-NRC Committee on Vision (1980), based on Louise Sloan's design. In all experiments (i.e. snake letters, Sloan and Künstler letters), the alphabet used was O, D, V, K, S, Z, N, H, and L.

The snake letters were also created using Fontographer. The orientation of the gabors was also varied. *Orientation* describes the tilt of the gabor relative to the path of the letter. The snake letters in the first experiment were tested with orientations of 0°

(collinear), 15°, 22.5°, 30°, 35°, 40°, 45°, 60°, and 90° (orthogonal). When testing object recognition as a function of position in the visual field, only the collinear and orthogonal letters (the extremes) were used.

The observer viewed a gamma-corrected grayscale computer monitor (Pelli and Zhang, 1991) run on a Power Mac 7500, normally from a distance of 60 cm. The background luminance was set to the middle of the monitor's range, about 10 cd/m².

Noise

White noise is visual noise containing all frequencies and orientations (Fig. 2). The white noise was made up of square checks, each a luminance increment sampled independently from a Gaussian distribution truncated at ± 2 standard deviations. At the normal viewing distance of 60 cm, a side of a check subtends 0.063°.

Procedure – Identification of snake letters at 20° eccentricity, as a function of orientation.

Each trial began with the appearance of a small black square that the observer fixates on, along with a message instructing the observer to click the mouse in order to initiate the trial. Once the observer clicked the mouse, a signal letter would appear, for 200 ms, at 20° of eccentricity from the fixation point.

After a 200 ms blank screen, the response screen appeared, containing all of the possible answers. The observer then selected what he thought had been the signal on the response screen with the mouse. If the signal was correctly identified, this was acknowledged with a short beep. After running a specific signal, the program randomly selected the next condition (a snake letter with a different orientation) and initiated

another run. Each run was made up of 40 trials. Calculation of the observer's threshold was made through the QUEST program (Watson and Pelli, 1983). The QUEST program estimates threshold by interpolating data from all of the observer's trials so far in that run. If the observer correctly identifies the signal, the QUEST program lowers the contrast of the signal, making the task more difficult. If the observer incorrectly identifies the signal, the QUEST program increases its contrast. The end result is an accurate measurement of the *contrast threshold* at which the observer correctly identifies the signal 82% of the time. Each orientation (i.e. 0°, 15°, 22.5°, 30°, 35°, 40°, 45°, 60°, or 90°) was run several times allowing for an accurate determination of threshold. Each condition was run with and without a white noise background.

Procedure – Identification of Sloan and Künstler; Collinear and Orthogonal snake letters as a function of eccentricity.

The same experimental procedure was used. However, the signal letter was displayed at a distance from fixation which varied across runs: 0°, 4°, 8°, 12°, 16°, or 20°.

The signals varied in each series of experiments. In the first experiment, where we sought to study *Factor of Good Continuation* in the periphery, the signal varied between snake letters of varying orientation. In the conditions where we tested the principle of *Prägnanz* across the visual field, the signal alternated in each run between collinear and orthogonal snake letters or Sloan and Künstler letters.

RESULTS

Efficiency and the Ideal Observer

The ideal observer is the algorithm that achieves the best possible performance given the same task provided to the human observer. We implement the ideal observer as

a computer program which contains templates of all possible stimuli. The ideal observer is used as a standard against which the results of the human observer are compared. The ideal observer is presented the same task as the human observer. Since the ideal observer has templates to identify the signal with, it will only begin to misidentify the signal when the noise masking the signal generates enough erroneous information that the ideal observer begins to incorrectly correlate stimulus and template. The thresholds of the ideal observer and human observer were measured in white noise in all conditions in order to calculate efficiency.

Efficiency was calculated by taking the energy threshold of the ideal observer and dividing it by the threshold elevation of the human observer. The threshold elevation, the threshold in white noise minus the threshold on a blank field, discounts the effects of the noise inherent in the visual system of the observer. All data points represent average efficiency; calculated as being the geometric mean of several different efficiencies for the same condition. The vertical bars represent ± 1 standard error.

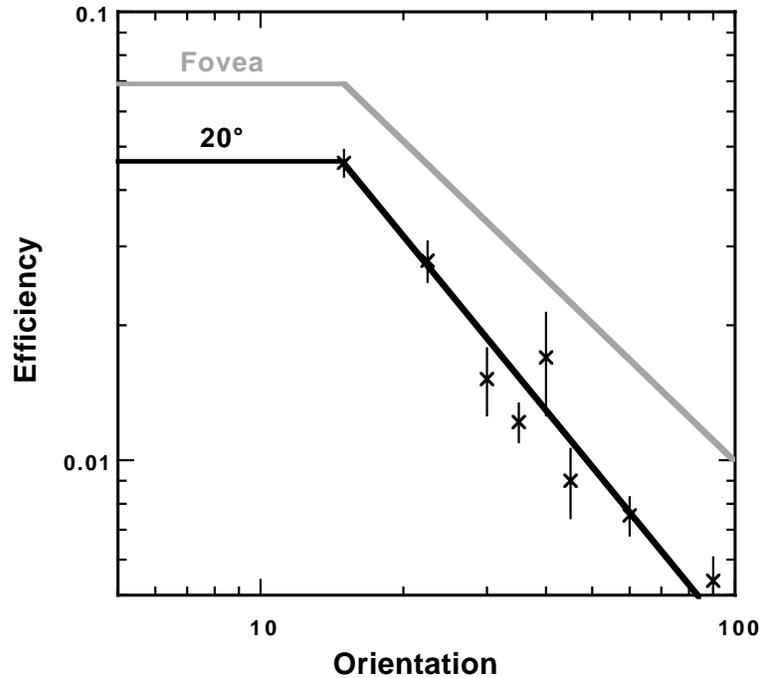


Figure 5. Efficiency as a function of orientation of the gabors relative to the path of the letter. The data in the fovea is from Majaj et al. (in prep.). The vertical bars indicate ± 1 standard error.

The vertical scale is efficiency, plotted on a logarithmic scale. Comparing with the foveal data (from Majaj et al.), our peripheral data have lower efficiency but are otherwise stable. The data in Figure 5 indicates a drop in efficiency, beginning at approximately 15° of orientation at 20° eccentricity, when testing snake letters of various degrees of perturbation. I found that efficiency, moving from the fovea to 20° in the periphery, dropped by a factor of 10, from 5.5% to 0.55%. By 15° of orientation, efficiency has already been cut in half to 2.7%, indicating a mechanism very sensitive to good continuation.

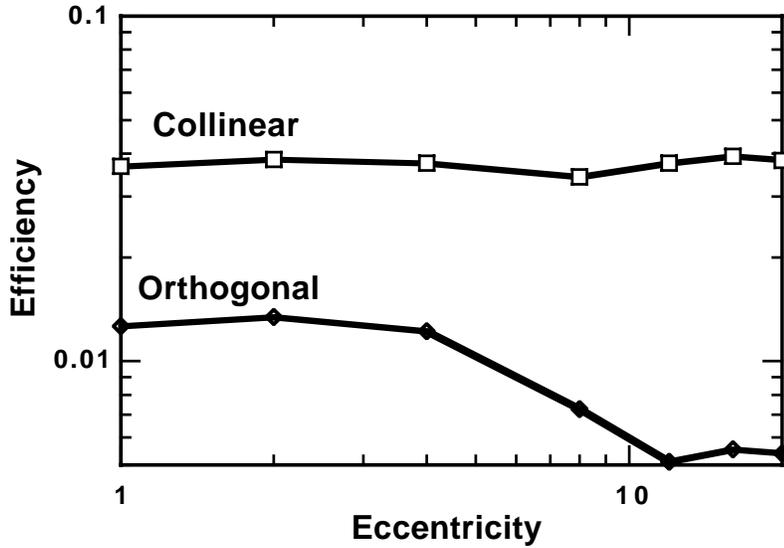


Figure 6. Efficiency as a function of eccentricity for collinear and orthogonal snake letters.

When tested with the collinear snake letters at various eccentricities (Figure 6), the observer maintained the same efficiency for identification of collinear snake letters at all eccentricities. Efficiency for orthogonal letters fell nearly a factor of two, between the fovea and 20° in the periphery, falling from 2.8% to 1.6%.

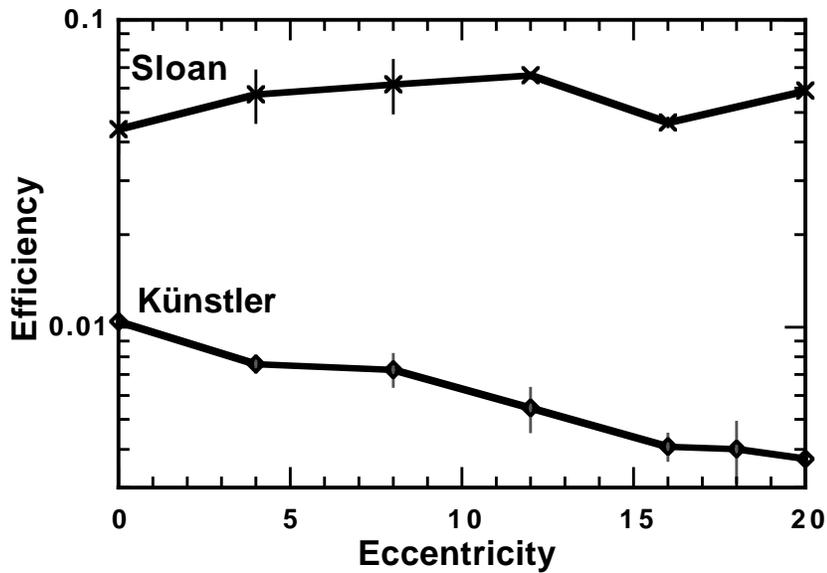


Figure 7. Efficiency as a function of eccentricity for Sloan and Künstler letters. The vertical bars indicate ± 1 standard error.

We also tested Sloan and Künstler letters as a function of eccentricity (Figure 7). For Sloan, efficiency was independent of eccentricity. For Künstler letters, efficiency dropped with eccentricity. Efficiency for identifying Sloan characters began in the fovea at 4.5% and at 20° of eccentricity had risen slightly to 5.9%. Efficiency for identifying Künstler characters began at 1%, and dropped to less than half, or 0.37% at 20° of eccentricity.

DISCUSSION

The collinear and orthogonal snake letters span the full range of orientations. We tested identification efficiency at various eccentricities, from 0 to 20° in the right visual field. Since they gave very different results, and they differ in *Prägnanz*, collinear being simple and orthogonal being complex, we also tried two more conventional fonts, Sloan and Künstler, that are also simple and complex (Figs. 3, 4). Efficiency for simple letters (Sloan and collinear snakes) is independent of eccentricity, while efficiency for complex letters (Künstler and orthogonal snakes) drops with increasing eccentricity. Finding that the periphery copes poorly with objects that violate Gestalt laws suggests that the periphery may have only the Gestalt mechanisms.

Field, Hayes, and Hess (1993) theorized that there are “association fields” that each can group elements together, segregating them from the background. Our results suggest that association fields in the fovea can integrate features regardless of violation of *Prägnanz*, albeit at a slightly lower efficiency. In the periphery, however, the association fields are much less able to group elements that violate *Prägnanz*.

Millions of elderly citizens with central field loss compensate by viewing the world through the periphery of their vision. The effects of crowding, combined with the

lack of feature integration that results from the violation of the *Factor of Good Continuation*, can help explain why reading can be such a difficult task for them.

CONCLUSION

The periphery has historically been neglected in vision research because it was 'bad' at identification and many other tasks. We now see that it provides us with an invaluable tool to isolate the Gestalt mechanisms of feature grouping and object recognition. Thus, when creating fonts to be used in literature, fonts should be created with respect to the Gestalt principles making them easier to identify for all observers thus making an otherwise almost impossible task significantly easier for people with impaired central vision.

ACKNOWLEDGEMENTS

I'd like to thank Denis Pelli for inviting me into his lab, suggesting where each experiment might lead me, and providing me with a background in vision research; Marialuisa Martelli for helping me choose a topic, teaching me how to use MATLAB and guiding me throughout my project; Tracey Berger for helping me put together my demonstrations and guiding me in my last experiments; Najib Majaj for suggestions concerning data interpretation; and Anne de Sostoa Manwell for suggesting Denis Pelli's lab.

REFERENCES

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.

- Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour Integration by the Human Visual System: Evidence for a Local "Association Field". *Vision Research*, 33, 173-93.
- Kovacs, I. (1996). Gestalten of today: early processing of visual contours and surfaces. *Behavioral Brain Research*, 82, 1-11.
- Kovacs, I., Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground discrimination. *Proc. Natl. Acad. Sci. USA*, 90, 7495-7497.
- Majaj, N. J., Raizman, N., Christian, C. J., Kim, E., Palomares, M. C., Pelli, D. G. (in preparation). The Role of Gestalt laws in Letter Identification.
- Pelli, D. G., Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, 31, 1337-1350.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437-442.
- Pelli, D. G., Burns, C. W., Farell, B., and Moore, D. C. (2000). Identifying Letters. *Vision Research*, (in press).
- Watson, A. B., Pelli D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113-120.
- Wertheimer, M. (1923). A Source Book of Gestalt Psychology. *Psychologische Forschung*, 4, 301-350; translated in Ellis, W. D. (1938)