How do we know whether Three Dots form an Equilateral Triangle?

Ernest Greene*
Department of Psychology. University of Southern California, USA

Abstract
A shape can be perceived when a widely spaced pattern of dots is displayed, and the neurons that encode the pattern must be able to register spans and relative orientations of the dots. A number of widely endorsed concepts for how those cues might be registered are not biologically plausible. These include the concepts that spatial locations are specified using coordinate addresses, that cortical orientation-selective neurons can register the spatial relations of widely spaced dots, and that remapping of the pattern to higher visual centers can encode those spatial relations. An alternative retinal mechanism is proposed, along with a hypothesis for how distances from pattern dots to a sensor located at the center of the pattern could encode the shape that is perceived.

“When we speak of the relation of direction of two points towards each other, we mean simply the sensation of the line that joins the two points together …. And the moment you have imagined the line, the relation stands before you in all its completeness… Just so the relation of direction between two lines is identical with the peculiar sensation of shape of the space enclosed between them.”

ILLUSTRATING A MINIMAL SHAPE TASK
Let's do a thought experiment. Assume stimulus configurations that are illustrated in Figure (1). Each panel shows an LED array, designated as the display board, which can show any combination of patterns at arbitrary locations on the board with a brief and simultaneous flash of the pattern dots. Here the pattern dots are highly visible, and they are being displayed for a moment that is too short to allow any eye movement.

Figure (1) is illustrating stimuli for a match-to-sample task. The left panel shows an equilateral 3-dot target pattern. The target is displayed at the same size in successive trials, but with location on the board being varied as well as the amount of rotation. The right panels have examples of choice-options that would be displayed shortly after the target. The upper-right panel shows a “matching” pattern, i.e., it is identical to the target except for being presented at a different location on the display board and at a different rotation. The lower-right panel shows a 3-dot pattern, designated as the non-matching pattern. Positioning of individual dots in non-matching patterns would vary, requiring also that they not be equidistant from one another.

Assume that a fixation dot is provided between display trials, which is not shown in the Figure (1) illustration. For a given display trial, the dots of the target pattern are simultaneously flashed for one microsecond, and three hundred milliseconds later a matching or non-matching pattern is displayed for one microsecond. The flashes are of sufficient intensity that each pattern is quite visible. You, as the observer, are told that your task is to say whether the two patterns are the same or different. For differences of dot location that are as conspicuous as illustrated here, there can be no doubt that you would be able to answer this question correctly.

USING ADDRESS VALUES TO MAKE THE JUDGMENT
How can an observer know whether or not the choice options shown on a given trial are identical to the target? It is not sufficient to answer: “Because I can see that one is an equilateral triangle and the other is not.” This presumes that memory, i.e., knowing what an equilateral triangle looks like, is essential to judgment. Memory might contribute to judgment for this particular set of stimulus alternatives, but we are using a match-to-sample task that would not require memory-at least not long-term memory that included the concept of an equilateral triangle. Using a match-to-sample task helps focus on the specific stimulus attributes that define the shape, for these cues are needed for encoding shapes whether they are stored in memory or not.

You might answer that it is because one of them is symmetrical and the other is not. That answer requires that the visual system has circuits that functionally bisect the space along an axis that
of numerous math modelers and use the coordinate system invented by René Descartes. Each location in the display is provided with an x, y address, with the units in this example being steps between successive dots. The location of each dot in the pattern can be specified using this address. If one of the address values is the same for a pair of dots, they are aligned with either the horizontal or vertical axis. The lower dots in the target pattern of Figure 1 have identical y-values, so here the difference in the size of the x-values specifies the distance (span) between the dots.

Continuing with the target pattern shown in Figure (1), the next task is to determine the span from each lower dot to the dot at the apex of the triangle. For each member of the pair, we can calculate the diagonal span by applying the Pythagorean Theorem. Conceptually the diagonal span would be the square root of \((x^2 + y^2)\), wherein: \(x\) is the distance from a bottom dot to the midpoint between the two dots, and \(y\) is the distance from that midpoint to the top dot. Computationally, having determined the full span between the lower dots, we can simply divide that value in half to provide the \(x\) portion of the Pythagorean formula. These calculations provide the length of the hypotenuse of each hypothetical right-angle triangle, which in this case is the diagonal distance from the apex dot to each of the lower dots. For the target pattern, we affirm that the span is the same for each of the three pairs of dots, which means that they have equilateral positioning.

What about the matching and non-matching choice options? In the examples that are illustrated, none of the dot-pairs are aligned with an axis, so each span must be calculated as a diagonal. Here one can use the difference in the \(x\) components of the two addresses to specify one leg of the right-angle triangle and the difference in the \(y\) components to specify the other leg, and then proceed to calculate the diagonal by applying the Pythagorean concept.

This seems like a great deal of operational complexity for accomplishing what we take to be an easy perceptual judgment. Even if the complexity were acceptable, the use of addressing concepts is not physiologically plausible. We have no evidence that a stimulated photoreceptor knows its position within the photoreceptor array or can convey its location by delivering address values. The same can be said about activated retinal ganglion cells, neurons in primary visual cortex, or for any other neuron population. One might object to this generalization on the basis of the precise anatomical mapping of connections from one population to another. This allows a localized stimulus in the retina to provide location-specific activation of neurons in V1 [1-4], and precise topographic mapping can be demonstrated for a number of neuron populations [5-7]. However, it is not correct to further assume that this provides information that can be used to specify distances among stimulated points.

Some engineers have developed "neuromorphic" retinal circuits, which adopt physiological principles in the processing of images. One approach provides for pixels that register localized changes in light level, for this can reduce the quantity of information that must be delivered to the data bus [8]. For the information to be useful, the pixel must provide not only the amount or intensity or color change, but also the address of each
pixel that registered a change [8]. This is an accommodation to the reality that having precise localization of a stimulus does not, in itself, provide any usable spatial information. Parallel fibers that map activation from one neuron population to another do not generate that information.

One presumes that the visual system has alternative mechanisms for deriving information about spatial relations, but the point here is that we have no basis for invoking a coordinate address system. Given this, it seems especially cavalier that math modelers have assumed addresses, and the ability to calculate distance, alignment, or other spatial relationships. Psychologists and neuroscientists alike appear to have been uncritical in accepting computational concepts.

**AN ALTERNATIVE MEANS FOR NEURONS TO REGISTER DOT SPANS**

Could we encode the spatial relations of these dots using orientation-selective, end-stopped neurons? [9-11] Their physiology is most often discussed as being specific for contours, but one might argue that they could register the relative orientation and span of two dots. To answer this possibility, it is necessary to provide additional details about the Figure (1) stimuli, as they would normally be displayed and judged in my laboratory [12-15]. Room illumination would be dim, so the observer would be able to see the outlines of the display board and could easily center his or her gaze on the board. But to assure central fixation of gaze, we could add four dim dots that appear at the center of the board between trials, disappearing just before the stimuli are flashed. At the standard viewing distance the individual LEDs would subtend about 5 minutes of visual angle, roughly the diameter of the dendritic field of a parasol ganglion cell [16,17]. Given center-surround mechanisms in this population of neurons [18,19], one might argue that a flashed dot would selectively activate a single ganglion cell, though it does not seem terribly material whether the retinal signal is from just one cell or a small cluster at the stimulated location.

The pertinent point is that under these viewing conditions the span between dots in the target pattern would be about 5.5 degrees of visual angle, roughly twice the span of the longest receptive fields of orientation selective neurons in V1 [20]. Given that the span between the dots greatly exceeds the length of orientation-selective neurons, they would not be able to register alignment (orientation) or distances between dots. And there is no doubt that an observer could judge an even larger pattern and assess the spans among its dots.

There is an additional reason to reject the hypothesis that spans are being registered by orientation-selective neurons. The projections to V1 split the visual field down the middle, with stimulation of each retinal hemifield being sent to opposite hemispheres [21-23]. Therefore, viewed with only one eye and with fixation centered on the board, almost every display of a target or matching pattern would send two dots to one hemisphere, and send the third dot to the other hemisphere.

I know of no concept that specifies how spans and orientation of stimulus locations in separate hemispheres would be assessed. It is not plausible that callosal fibers provide information that allows the alignments and spans to be registered. There is no evidence of a covert mapping of locations from V1 in one hemisphere to V1 in the other. This would allow a stimulus that was restricted to one visual field to activate neurons in both hemispheres, one coming directly from lateral geniculate nucleus and the other by means of callosal fibers. Such bi-hemispheric stimulation has not been reported. One might think that horizontal fibers that link orientation-selective neurons [24,25] could provide the assessment of orientation and span between the dots. But that presumes that a single dot displayed to one hemisphere can activate an orientation-selective neuron, and further, link to a horizontal fiber that connects to the margin of V1, then onward through a horizontal fiber in the other hemisphere, finally connecting to the location of the second dot. One would not expect a single dot to activate a V1 neuron that is designed to respond to an elongated bar or grating. Nor would the dot in the opposite hemisphere be able to direct which orientation-selective neuron and which set of horizontal fibers to recruit.

**CAN NEURAL REMAPPING PROVIDE THE BASIS FOR SHAPE RECOGNITION?**

Perhaps we need to look beyond primary visual cortex. Could the assessment of orientation and span between widely separated dots be done by extrastriate neurons that have much larger receptive fields? There is little support for this approach, for research in this area generally uses stimuli having linear extent, e.g., bars and gratings [26-29]. Orientation selectivity of extrastriate neurons is thought to derive from the early processing in V1, but as discussed above the V1 neurons would not be able to register information about the relative orientation or degree of separation of widely spaced dots. Whether the extrastriate neurons are being driven by V1 output, or not, there is at present no evidence that they can register the orientation or degree of separation of widely spaced dots.

Perhaps higher visual centers are directly encoding the triangular dot configurations. Neurons in several extrastriate areas respond selectively to stimuli such as faces and other shapes [30-32], and can do so when the stimulus is placed at various locations in the visual field. This is accomplished by remapping that makes spatial location irrelevant, which is often described as an increase in the size of the receptive field. That certainly appears to be what is needed to deal with the change of location for the tandem displays illustrated in Figure (1). When the target is shown to the lower left of the display board and the observer must judge a choice option that is shown in the upper right, then selective shape responding with a larger receptive field might seem a plausible neuronal mechanism.

However, the remapping concept is fraught with practical problems of implementation. It requires anatomical convergence of fiber connections that can handle the vast number of shapes that we can identify, and do so irrespective of size and location. This is known as the “combinatoric explosion” problem. Here I want to emphasize a different point. A remapping that preserves shape information but makes its location moot assumes that the earlier processing was successful in registering local shape attributes [33-35]. Once registered by the earlier stage, that information can be combined and elaborated upon by a later
stage where the neuronal response does not depend (or depends less) on global positioning of the local cues. But if the earlier stage has not registered the spatial relationships, none of the extant remapping concepts will provide that information.

Further, rather than call for a shape judgment, one can display two widely separated dots and ask only for assessment of span and/or relative orientation. An observer can provide effective discriminations even if one dot is presented to the right visual field and the other dot is presented to the left visual field. Here the entire focus is on the spatial relationship of the two dots, and remapping does not de novo create this information.

**ELEMENTAL SHAPE ENCODING MIGHT BEGIN IN THE RETINA**

The simple fact is that we can register the spans and orientations among widely separated dots, and use that information to perceive shape. If known cortical mechanisms are not able to explain how this is done, we should consider a retinal option. It is possible that these stimulus attributes are derived within the retina, making the hemifield/hemisphere issue moot. A number of retinal cell populations have extended arborization that might be relevant to this process, with horizontal cells and amacrine cells being most conspicuous [36-38]. Polyaxonal amacrine (PA1) cells have highly restricted dendritic fields and a dense arbor of axon branches that extend over a very large area [39-47]. Action potentials of PA1 cells spread like waves through the axonal arbors of these cells. This may play a role for encoding stimulus motion [44-46] and it may not be a great stretch to suggest that distances among stimulated points in the image might be registered by the travel time of these waves.

In macaque the axonal fields of PA1 cells have diameters of roughly 2000 microns in central vision and 6000 microns in the periphery [40]. If one can assume similar dimensions in human retina, a given PA1 cell would cover about 7.5 arc° in central vision and about 40 arc° in the periphery. These neurons tile the retina as a mosaic, with receptive fields that are slightly larger than parasol ganglion cells [47]. The kinetics of their light response is similar to parasols and they are more transient than midget ganglion cells [48]. Conduction velocity of the axons is approximately half a meter per second [47]. If one assumes an axon length of 3000 microns or less, travel time from soma to tip would be 5-10 milliseconds.

A model making use of travel time of PA1 spikes as a method for registering spatial location or distances among stimulus markers would have to assume distance-encoding sensors and methods for synchronizing the timing. The PA1 cells are especially sensitive to sharp transitions of the visual scene, so they might be synchronized by microsaccades of the eye, thus providing signals that radiated from marked locations to distance sensors. This would allow time-of-travel to specify marker locations.

If the encoding system is provided with an array of sensors, do they all respond to passing waves or do we have a method for choosing one of them to deliver a summary of marked locations? The system may be designed to have the centroid of the pattern provide that summary. The amplitude of wave convergence is maximal at the centroid, so the centroid of the equilateral triangle shown in Figure (1) would have waves arriving at the same moment from each of the three points. Assessing distance and angle of three dots in relation to the center of the pattern provides an excellent basis for specifying equilateral positioning and could well be how the nervous system encodes that configuration. A mechanism for registering those positions might be easier to implement than trying to determine distances and orientations among dot pairs. Further, it provides a simple basis for establishing that a non-matching pattern is different than the target, for their distances to the centroid would not be the same.

If the equilateral triangle illustrated in Figure (1) were shown on our LED display board at the viewing distance that I normally use, the span from each flashed dot to the centroid would be a little over 3 degrees of visual angle. That would be about 900 microns as measured on the retina. The wave of spike propagation out the PA1 axonal branches is roughly 500,000 microns/second [47] so it would take about two milliseconds for those distances to be registered at the centroid.

At this point we cannot affirm that PA1 cells are providing the encoding of spatial relations within a pattern. An alternative role for the fast-conducting PA1 spikes might be to coordinate or poll other neuron populations, with registering of distance and alignment being done by another population of wide-field neurons. The key point here is to suggest that encoding of spatial information might be done within the retina, with the travel-time of spreading signals providing a potential basis for delivering that information.

**DISTANCE TO THE CENTROID OF A SHAPE OR PATTERN CAN PROVIDE FOR RECOGNITION**

An optimal encoding of more complex patterns or shapes, the latter being considered as a plethora of stimulus-marked locations, would likely require registering the direction from which waves arrived as well as the distances from those locations. However, it is worth noting that a summary that only records the distances can be exceptionally effective at identifying the shape/pattern [49]. For example, one can create an inventory of arbitrary shapes, then for each shape calculate the location of the centroid and the distance of each dot from that centroid. For a given shape those distances are distributed into a histogram wherein each bin specifies how often a given distance was observed. This essentially converts the 2D information in the pattern into a 1D summary [49].

For discriminating among shapes in the inventory, one then compares the histograms for degree of match. A number of comparison methods are available, one of the simplest being the sum of squared differences. Here one calculates the difference in height for each bin of two histograms, squares that value, and then sums the squares that were derived across all bins. If the 1D histograms differ substantially, the sum of squared differences will be large; if they are similar, that value will be small [49].

As evidence that this method provides good discrimination among shapes, an inventory of 450 shapes like the one shown in Figure (2) was created. A centroid histogram was derived for each shape, and the sum of squared differences was calculated for all

---

Distance alone does not provide an optimal shape or pattern summary, but it is useful to note that summaries based on distance can be very efficient and effective. It is unlikely that the nervous system would derive distributions or compare them in the manner that is described above. Nonetheless, a biologically plausible mechanism that accomplishes something similar is not out of the question.

CONCLUSION

Discrete dots can be used to surround a space and provide the perception of a shape, the minimum being three dots. Further, the spatial positioning of the dots can be critical for what shape is perceived. If the three dots are equidistant from one another, we perceive an equilateral triangle. It is not clear what shape is perceived. If the three dots are equidistant from one another, we perceive an equilateral triangle. It is not clear what shape is perceived. If the three dots are equidistant from one another, we perceive an equilateral triangle. It is not clear what shape is perceived. If the three dots are equidistant from one another, we perceive an equilateral triangle. It is not clear what shape is perceived. If the three dots are equidistant from one another, we perceive an equilateral triangle. It is not clear what shape is perceived. If the three dots are equidistant from one another, we perceive an equilateral triangle. It is not clear what shape is perceived. In the manner that is described above. Nonetheless, a biologically plausible mechanism that accomplishes something similar is not out of the question. encoded in the retina, where there are wide-field neurons that could register the relative positioning of the dots.

ACKNOWLEDGMENT

Research funding for this laboratory was provided by the Neuropsychology Foundation and the Quest for Truth Foundation.

REFERENCES


