Population coding in the retina

Sheila Nirenberg* and Peter E Latham

Recent advances in multi-electrode recording have brought us closer to understanding how visual information is encoded by populations of retinal ganglion cells. By monitoring the visual responses of many ganglion cells at once, it is now possible to examine how ganglion cells act together to encode a visual scene.

Addresses
Department of Neurobiology, University of California at Los Angeles, 10833 Le Conte Avenue, Los Angeles, California 90095-1763, USA
*e-mail: sheilan@UCLA.edu

Current Opinion in Neurobiology 1998, **8**:488–493

http://biomednet.com/elecref/0959438800800488

Introduction

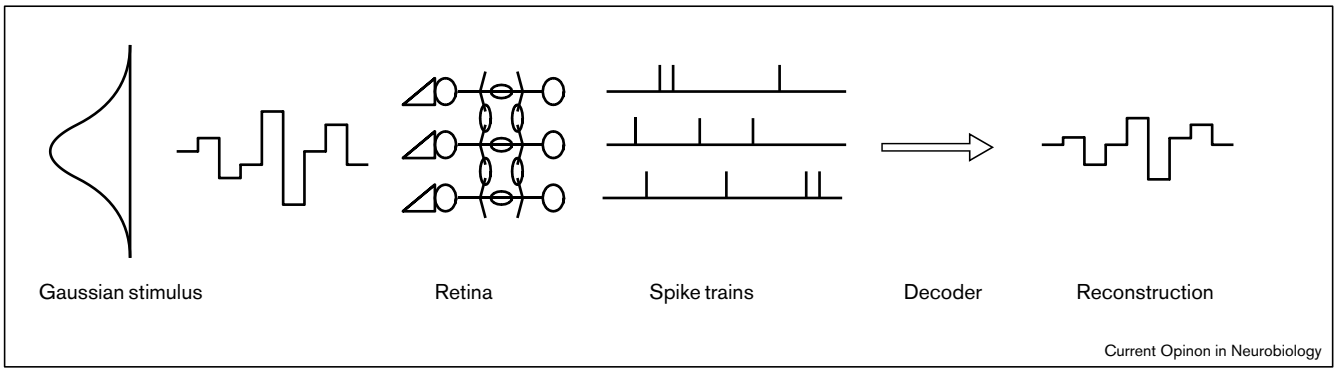
The first stage of visual processing occurs in the retina. Visual stimuli excite the retinal input neurons (the photoreceptors), which convert the visual stimuli into electrical signals. These signals are then propagated though the retinal circuitry to the output neurons (the ganglion cells) and are relayed to the brain in the form of action potentials.

A primary goal of retinal research is to determine how visual information is transformed by the retina; more specifically, to determine how a visual stimulus is transformed into a pattern of action potentials in the ganglion cells. The experimental strategy for addressing this question is now standard: present the retina with visual stimuli and then record ganglion cell responses. The working hypothesis is that if we knew the input/output relationship of every ganglion cell, we would know the code used by the retina to represent visual information. If we knew the code, then we would be able to reconstruct a visual scene (to the same extent as the brain does) given the spike trains of ganglion cells. (We are not implying that the brain actually reconstructs a stimulus; however, it does extract information from it. Performing a reconstruction is just a way of estimating how much information the brain could extract.)

Recent advances in multi-electrode recording techniques have taken us a significant step closer to achieving this goal. These techniques make it possible to, first, examine the input/output relationships of many ganglion cells rapidly and efficiently (see [1•–4•]), and, second, examine the input/output relationships of many ganglion cells at the same time. The latter is necessary for determining whether or not ganglion cells act as independent encoders. This is important because it bears on the experimental strategy one might use to deduce the retinal code. If the ganglion cells act independently, we can, in principle, find the code by determining the input/output relationships of all the ganglion cells separately, and then use these individual relationships to piece together the whole code [5•]. If the ganglion cells do not act independently, then we can only determine the code by examining the behavior of the ganglion cells as a population.

Here, we focus on two topics. First, we review recent experiments using multi-electrode recording to examine population codes used by ganglion cells. Second, we review evidence that ganglion cells do not act independently

Figure 1



Schematic overview of a reconstruction experiment. A spatially uniform stimulus was projected onto the retina. The intensity of the stimulus changed every 15 ms to new one, which was drawn at random from a Gaussian distribution of intensities. Spike trains from ganglion cells were recorded using a multi-electrode array. The spike trains were then presented to a decoding algorithm, which constructed an estimate of the stimulus. Adapted with permission from [6••].

and discuss how this affects strategies one might use to characterize the retinal code.

How does a population of ganglion cells encode a visual stimulus?

A recent paper by Warland *et al.* [6••] describes how a population of ganglion cells collectively encodes a visual stimulus. It presents one of the first attempts at decoding a population signal. Using a multi-electrode array, the authors recorded the spike trains of numerous ganglion cells from the tiger salamander retina while presenting a time-varying stimulus. The stimulus was a spatially uniform field whose intensity varied randomly over time. (Specifically, the intensity changed every 15 ms to a new one, which was drawn at random from a Gaussian distribution.) The authors then applied various decoding strategies to the spike trains to try to reconstruct the time course and intensity of the stimulus (Figure 1).

Warland *et al.* [6••] approached this reconstruction problem as an optimization problem. They modeled the transformation from spike trains to stimulus using many parameters and then adjusted those parameters to minimize the mean square error between the actual stimulus and the reconstructed stimulus.

Two kinds of reconstruction methods were used. The first was a linear reconstruction method based on methods used for single neurons by Bialek *et al.* [7], and works in the following way: Consider spike trains from p neurons, and let t_i^{ν} denote the i^{th} spike of neuron ν. The linear reconstruction attempts to express the stimulus, $S(t)$, as the sum of temporal filters, $f^{\nu}(t)$,

$$S^{\text{estimate}}(t) = a + \sum_{\nu=1}^{p} \sum_{i} f^{\nu}(t - t_i^{\nu}). \qquad (1)$$

The shapes of the filters, $f^{\nu}(t)$, plus the constant a are adjusted to minimize the mean square error between the estimated stimulus, $S^{\text{estimate}}(t)$, and the actual stimulus, $S(t)$,

$$\text{mean square error} = \int dt \left[S(t) - \mathrm{S}^{\text{estimate}}(t) \right]^2. \qquad (2)$$

The second method used an artificial neural network trained by back-propagation to reconstruct the stimulus. In its most basic form, this method starts with a linear reconstruction as in equation 1, and then passes the output of the linear reconstruction through a nonlinearity. Again, the optimization parameters are the shapes of the filters and the constant term (a in equation 1). See Box 1 for a more detailed description of the two reconstruction methods.

From these reconstruction experiments, we learned three things about how ganglion cells encode the visual stimulus:

1. The quality of the reconstruction of the stimulus depended on the number and the type of ganglion cells used to carry out the reconstruction. Generally, different types of cells carried non-overlapping information, so the quality of the reconstruction increased rapidly as more cells of different type were included (Figure 2). The quality of the reconstruction hardly improved at all when cells of the same type were added.

2. The quality of the reconstruction depended on the frequency spectrum of the stimulus: fast changes in intensity (above 10 Hz) and very slow changes (below 1 Hz) were not well reconstructed. Performance was optimal at about 2.5 Hz. Warland *et al.* [6••] suggest that the optimal bandwidth represents a tradeoff between the lowpass properties of the photoreceptors and the bandpass nature of retinal processing. A candidate mechanism for the low frequency cutoff is differentiation in the inner retina, possibly at the bipolar cell synapse.

3. The linear reconstruction was as effective in reconstructing the stimulus as the neural network, suggesting that most of the information about this stimulus was captured by linear operations on the spike trains. The implications of this finding for single-cell processing is beyond the scope of this review, but a discussion can be found in [8]. Its implications for multi-cell processing and population coding is discussed in the next section.

Do ganglion cells act as independent encoders?

Numerous experiments have shown that the activity of retinal ganglion cells is strikingly correlated. Evidence for such correlations was first found in cat retina over 30 years ago [9]. Subsequent experiments in goldfish [10,11], rabbit [12], cat [13–15] and salamander [16] have yielded essentially the same result — that is, that neighboring ganglion cells tend to fire together (or to be silent together) more often than would be expected by chance.

The correlated behavior of two cells can be readily displayed in a cross-correlogram, which gives the probability, $P(\tau)$, of two cells firing a time τ apart. Typical cross-correlograms show either a peak or a dip near zero delay: ganglion cells of the same type (e.g. two ON-center cells) show a peak, whereas cells of opposite type (e.g. one ON-center cell and one OFF-center cell) show a dip [17] (Figure 3a).

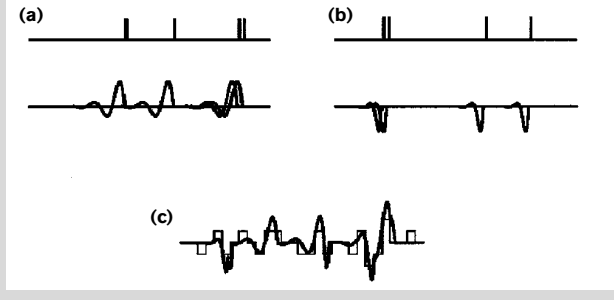
Most of the studies mentioned above examined correlated activity when the retina was in the presence of a constant stimulus, that is, darkness or constant illumination. Meister *et al.* [16] tested whether such correlated activity could be found when the retina was presented with a time-varying stimulus. The stimulus used in this

Box 1

1. Linear reconstruction

Reconstructing a stimulus from a set of spike trains requires a transformation from a discrete set of events (the spikes) to a continuous function of time (the stimulus). One way to do this is by linear reconstruction. With this method, each spike is replaced by a filter – a continuous, but sharply peaked, function of time. The filters from both spike trains are then summed to produce a reconstruction of the stimulus (see equation 1 in the text). The shapes of the filters are chosen so that the mean square error between the reconstruction and the original stimulus, averaged over the whole experiment, is minimized. (If one thinks of the spike train as a series of delta-functions, then the linear reconstruction is simply the convolution of those delta-functions with the optimized filters.)

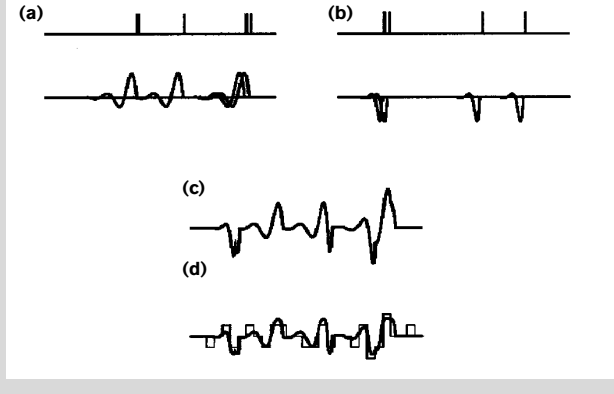
This method is illustrated below for two cells viewing the same stimulus. Spike trains from the two cells are labeled **(a)** and **(b)**. Below each spike train is a set of filters. To produce a reconstruction of the stimulus, the filters from both spike trains are summed, producing the thick trace in **(c)**. The thin trace in that panel shows the original stimulus.



2. Neural network reconstruction

Warland *et al.* [6••] used a number of different neural network architectures to perform nonlinear stimulus reconstruction. They all start out in the same way as the linear reconstructions: spike trains – **(a)** and **(b)** – are replaced by a set of filters, which are then summed to produce **(c)**. In the simplest neural network architecture, which has no hidden units, the summed filters, **(c)**, are then passed through a sigmoidal nonlinearity to produce the reconstruction, shown as a thick trace in **(d)**. The thin trace in that panel shows the original stimulus. The optimization parameters are again the shapes of the filters.

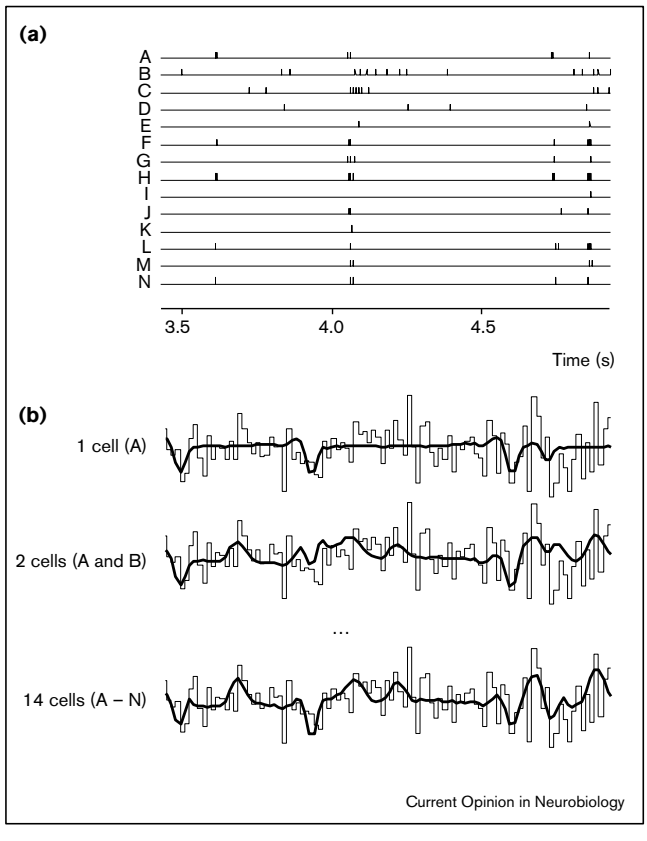
In architectures with hidden units, an additional layer is added to the neural network. For these architectures, the thick trace in **(d)** represents the output of a hidden unit, rather than the final reconstruction. To reconstruct the stimulus, the output of each hidden unit undergoes the same class of transformations as was performed on the input layer: they are combined linearly and then passed through a sigmoidal nonlinearity. The output of that final nonlinearity is the reconstructed stimulus. (The additional layer is not shown, because it is identical in structure to the input layer.) Each hidden unit has associated with it a different set of linear filters; therefore, more free parameters are added as more hidden units are used. Although increasing the number of free parameters might be expected to improve the quality of an estimate, this was not the case in the Warland *et al.* [6••] experiment – the quality of the reconstruction was insensitive to the number of hidden units.



Current Opinion in Neurobiology

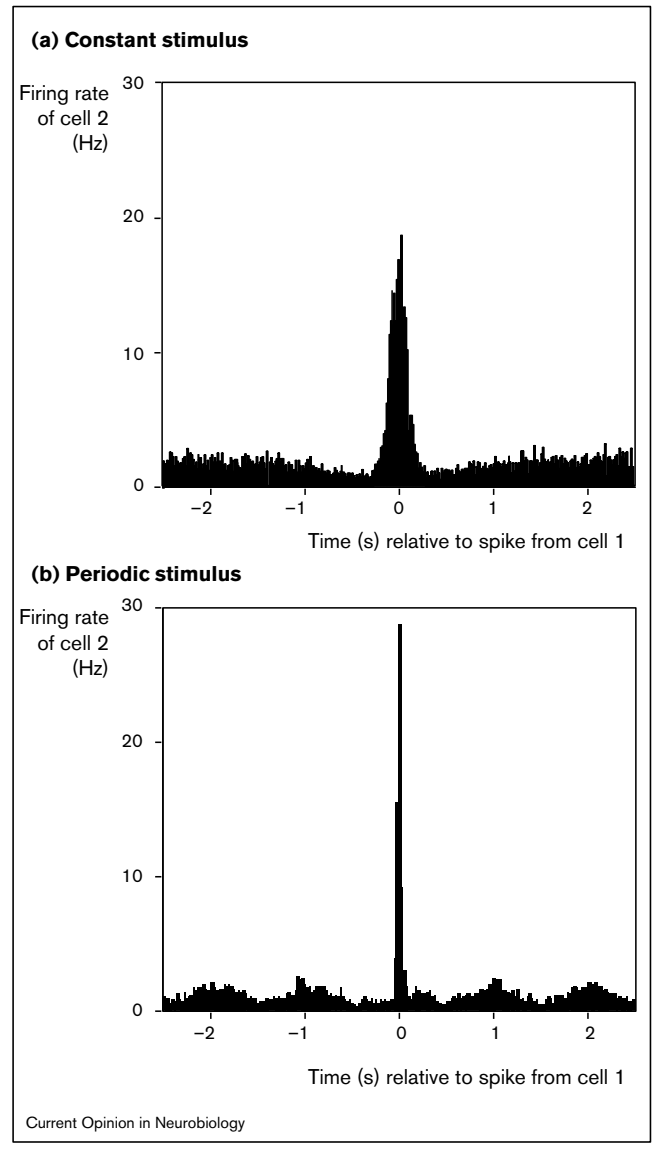
experiment was a spatially uniform illumination field that switched on and off at regular intervals. Consistent with the previous studies using constant illumination, the authors found that the responses of neighboring cells were tightly locked in time (Figure 3b). Note that the stimulus itself causes the responses of the cells to be correlated. For example, the two cells in Figure 3b are OFF-center cells; thus, they both fire whenever the light goes off. These stimulus-induced correlations produce periodic peaks that are small and wide compared to the large peak that occurs near zero delay. The large, central peak reflects correlations above and beyond those produced by the stimulus and indicates that the two cells did not respond to the stimulus independently. Presumably, some common

Figure 2



Visual information from a population of retinal ganglion cells. The retina was viewing the stimulus used in Warland *et al.* [6••], which was a spatially uniform field whose intensity varied randomly over time. **(a)** Spike trains from 14 ganglion cells. **(b)** Reconstructions of the stimulus using various combinations of spike trains. The thin traces represent a segment of the stimulus. The thick traces represent the stimulus reconstruction. The reconstruction improved as more cells were used. The improvement was large when cells of different type were combined, and small when cells of the same type were combined. Adapted with permission from [6••].

Figure 3



Correlated activity in response to a constant stimulus versus correlated activity in response to a time-varying stimulus. **(a)** Cross-correlograms of two retinal ganglion cells presented with constant illumination. Adapted from [17]. **(b)** Cross-correlograms of two cells presented with periodic (1 Hz) illumination: 500 ms at high intensity followed by 500 ms at low intensity. Adapted with permission from [16].

input triggered them both to fire, or one cell triggered the other [18•]. Meister *et al.* [16] found that approximately 50% of all ganglion cell spikes fell within the central peak of the cross-correlogram of at least one cell pair.

One would expect such correlations to have important implications for how the brain interprets the signals of ganglion cells. (Here, and in the remainder of this review, 'correlations' refers to correlations above and beyond those produced by the stimulus; this is the more standard definition.) One would expect the implications to become clear when one tries to reconstruct the stimulus from ganglion cell spike trains. Reconstruction methods that treat the cells as independent encoders should produce different results from reconstructions that take correlated activity into account.

Interestingly, the results of the Warland *et al.* [6••] experiment suggest that correlated activity plays a minimal role in encoding visual stimuli — at least the stimuli used in those experiments. We infer this because the linear reconstruction method, which treats the cells as independent encoders, and the neural network, which can utilize correlations, performed equally well at reconstructing the stimulus. As described in Box 1, the linear reconstruction method operates by computing filters for ganglion cells and then summing them to produce an estimate of the stimulus. Once the filters have been chosen, the contribution of each spike to the stimulus estimate is explicitly independent of the other spikes. The

neural network method, on the other hand, allows spikes from other cells to influence a given spike's contribution to the stimulus estimate. Warland *et al.* [6••] even demonstrated that the neural network can use correlated activity by providing both reconstruction methods with artificial data in which correlations among spike trains contained information about the stimulus. They found that the neural network method did significantly better than the linear reconstruction method — by a factor of about two. Thus, the fact that the neural network method could have picked up information lurking in the correlated activity, but did not, suggests to us that ganglion cells were acting as independent encoders.

Warland *et al.* [6••] interpret their experiments differently and suggest that ganglion cells do in fact influence each other. Their argument is based on the fact that the optimal filter chosen for each cell changes as more cells are added to the decoder.

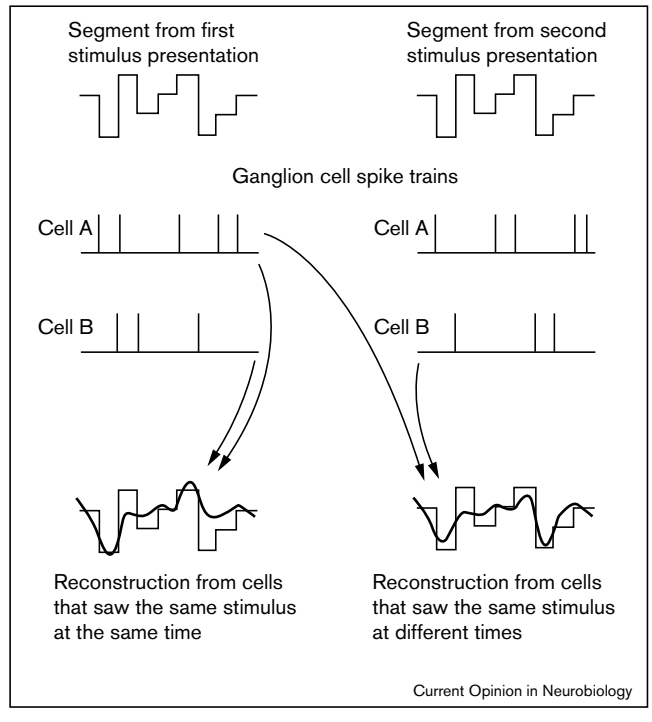
Fortunately, a simple experiment can resolve this difference in interpretation by testing directly whether or not correlations are important. This experiment would compare the quality of a reconstruction when ganglion cells are explicitly independent to the quality of a reconstruction when they are not. To force the cells to be independent, one could use the following trick. Let's say one is recording from N cells. Present the retina with the same stimulus N times (e.g. the time-varying stimulus used in Warland *et al.* [6••]). On each presentation, record from a different cell. Use these N recordings to reconstruct the stimulus. Call this reconstruction 1. Then present the retina with the same stimulus once more, but this time record from all N cells at the same time. Now perform the reconstruction with these simultaneous recordings. Call this reconstruction 2. Any differences between reconstructions 1 and 2 can be attributed directly to correlations. This experimental procedure is illustrated schematically in Figure 4.

So, what is the answer? Does a population of cells encode a visual stimulus as a team of interdependent encoders or a collection of independent encoders? This is still an open question. Experiments such as the one described above would help resolve it. In addition, as Warland *et al.* [6••] point out, one needs to examine decoding strategies used on more complex stimuli, such as spatially varying images. It may well be that correlated activity does not play a major role in encoding spatially uniform stimuli, but does contribute to coding more natural images.

Conclusions

A fundamental issue in neuroscience is how information is encoded in the activity of a population of neurons. Recently, Warland *et al.* [6••] addressed this issue in studies of the retina by asking how visual information is encoded by a population of retinal ganglion cells. Their approach was to record the spike trains from multiple retinal ganglion cells while displaying a visual stimulus, and then use those spike trains to reconstruct the stimulus.

Figure 4



Outline of an experiment designed to address whether correlations convey information about the stimulus. First, the retina would be presented with two identical visual stimuli, one after the other, then the spike trains from two cells, labeled A and B, would be recorded. Next, the reconstruction when cells A and B both saw the first presentation (left panel) would be compared to the reconstruction when cell A saw the first presentation and cell B saw the second (right panel). In the first reconstruction, the spike trains may (and undoubtedly will) contain correlations. In the second reconstruction, the cells are forced to be independent, because they are monitored at different times. Any differences in the reconstructions can be attributed directly to correlations.

Two main points emerged from this work. First, as one might expect, the quality of the reconstruction depended on the number of ganglion cells used to carry out the reconstruction. Interestingly, though, the quality increased only slightly when cells of the same type were used, but increased strikingly when cells of different types were used. For example, two ON-center cells provided more information than one ON-center cell, but much less than one ON-center and one OFF-center cell.

The second point was that a linear reconstruction method worked as well as a nonlinear one (a neural network), indicating that nonlinear operations were not playing a major role in the reconstruction. This provides evidence that the ganglion cells were acting as independent encoders, which is surprising in light of numerous reports that retinal ganglion cells show a high degree of correlated activity. If ganglion cells are independent encoders, then

what is correlated activity for? One possibility is that it is used for encoding more complex images than the spatially uniform ones used in Warland *et al.*'s [6••] experiment. Future experiments involving multiple ganglion cell reconstructions of spatially and temporally varying images should be able to address this issue, and, at the same time, provide valuable information on how visual scenes are encoded in populations of retinal ganglion cells.

Acknowledgements

We thank John Assad, Steve DeVries, Markus Meister and Pam Reinagel for helpful discussions.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. DeVries SH, Baylor DA: **Mosaic arrangement of ganglion cell**
• **receptive fields in rabbit retina.** *J Neurophysiol* 1997, **78**:2048-2060.

A multi-electrode recording study of the ganglion cell types of the rabbit retina. The authors classified eleven different physiological types and determined their receptive fields and spatial arrangement in the retina. This provides important, detailed information about how the different cell types are spatially arranged to encode the visual world.

2. Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M:
• **Adaptation of retinal processing to image contrast and spatial scale.** *Nature* 1997, **386**:69-73.

A multi-electrode recording study that examined network adaptation. The authors showed that the retina adapts not only to light, but also to the statistics of the light pattern. This latter adaptation was found to occur on a scale of seconds, about a hundred times more slowly than the light response.

3. Berry MJ, Warland DK, Meister M: **The structure and precision**
• **of retinal spike trains.** *Proc Natl Acad Sci USA* 1997, **94**:5411-5416.

A multi-electrode recording study that examined the reproducibility of ganglion cell responses. The authors presented the retina with a random flicker stimulus and measured the trial-to-trial variation in responses. They found that ganglion cells responded with brief periods of firing, called 'firing events', and that these events were highly reproducible: the jitter from trial to trial was as low as 1 ms. In addition, they found that the information conveyed by the timing of the firing events was much greater than that conveyed by the number of spikes per event.

4. Nirenberg S, Meister M: **The light response of retinal ganglion**
• **cells is truncated by a displaced amacrine circuit.** *Neuron* 1997, **18**:637-650.

A multi-electrode recording study that examined how retinal ganglion cell responses are generated. The authors investigated how a set of amacrine cells contributes to the formation of transient ganglion cell responses. They ablated the amacrine cells and found that transient responses became prolonged. This indicated that these responses are shaped, at least in part, by a circuit mechanism, rather than a cellular mechanism. In addition, the circuit must involve delayed inhibition.

5. Jabobs A, Werblin F: **Spatiotemporal patterns at the retinal**
• **output.** *J Neurophysiol* 1998, **80**:447-451.

The authors used an innovative technique to estimate multiple cell responses from single-cell recordings: instead of showing the same stimulus while recording from different cells, one cell was monitored while a stimulus was presented sequentially at different spatial locations. By assuming translation invariant responses and disregarding correlations, the authors were able to use their single-cell recordings to construct the response of a population of cells.

6. Warland DK, Reinagel P, Meister M: **Decoding visual information**
•• **from a population of retinal ganglion cells.** *J Neurophysiol* 1997, **78**:2336-2350.

The authors investigated how a time-varying stimulus is encoded by a population of retinal ganglion cells. They attempted to reconstruct the stimulus from the spike trains of ganglion cells using two reconstruction strategies: a linear reconstruction method and a neural network (nonlinear) method. The linear reconstruction performed as well as the nonlinear method, indicating that most of the information about the stimulus was contained in linear operations on the spike trains.

7. Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D: **Reading a neural code.** *Science* 1991, **252**:1854-1857.

8. Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W: *Spikes: Exploring the Neural Code*. Cambridge, Massachusetts: MIT Press; 1997:76-98.

9. Rodieck RW: **Maintained activity of cat retinal ganglion cells.** *J Neurophysiol* 1967, **30**:1043-1071.

10. Arnett DW: **Statistical dependence between neighboring retinal ganglion cells in goldfish.** *Exp Brain Res* 1978, **12**:49-53.

11. Johnsen JA, Levine MW: **Correlation of activity in neighbouring goldfish ganglion cells: relationship between latency and lag.** *J Physiol* 1983, **345**:439-449.

12. Arnett D, Spraker TE: **Cross-correlation analysis of the maintained discharge of rabbit retinal ganglion cells.** *J Physiol* 1981, **317**:29-47.

13. Mastronarde DN: **Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X- and Y-cells.** *J Neurophysiol* 1983, **49**:303-324.

14. Mastronarde DN: **Correlated firing of cat retinal ganglion cells. II. Responses of X- and Y-cells to single quantal events.** *J Neurophysiol* 1983, **49**:325-349.

15. Mastronarde DN: **Interactions between ganglion cells in cat retina.** *J Neurophysiol* 1983, **49**:350-365.

16. Meister M, Lagnado L, Baylor DA: **Concerted signaling by retinal ganglion cells.** *Science* 1995, **270**:1207-1210.

17. Mastronarde M: **Correlated firing of retinal ganglion cells.** *Trends Neurosci* 1989, **12**:75-80.

18. Brivanlou IH, Warland DK, Meister M: **Mechanisms of concerted**
• **firing among ganglion cells.** *Neuron* 1998, **20**:1-20.

The authors characterized correlated activity in the retina and investigated the underlying circuit mechanisms. Three types of correlated activity were found: broad (firing within 40–100 ms), medium (firing within 10–50 ms) and narrow (firing within 1 ms). Disruption of chemical synaptic transmission affected the broad correlations, but not the medium or narrow ones. This result, along with analysis of the strength and time course of the correlations, suggested that broad correlations are probably caused by shared photoreceptor input, that medium correlations are caused by shared input from an interneuron such as an amacrine cell, and that narrow correlations are caused by gap junctions among ganglion cells.