Ruling out and ruling in neural codes

Adam L. Jacobs[a], Gene Fridman[b], Robert M. Douglas[c], Nazia M. Alam[a], Peter. E. Latham[d], Glen T. Prusky[a], and Sheila Nirenberg[a,1]

[a]Department of Physiology and Biophysics, Weill Medical College of Cornell University, New York, NY 10065; [b]Department of Neurobiology, University of California, Los Angeles, CA 90095; [c]Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, BC, Canada V5Z 3N9; and [d]Gatsby Computational Neuroscience Unit, University College of London, London WC1N 3AR, United Kingdom

Communicated by David W. McLaughlin, New York University, New York, NY, January 16, 2009 (received for review September 2, 2008)

The subject of neural coding has generated much debate. A key issue is whether the nervous system uses coarse or fine coding. Each has different strengths and weaknesses and, therefore, different implications for how the brain computes. For example, the strength of coarse coding is that it is robust to fluctuations in spike arrival times; downstream neurons do not have to keep track of the details of the spike train. The weakness, though, is that individual cells cannot carry much information, so downstream neurons have to pool signals across cells and/or time to obtain enough information to represent the sensory world and guide behavior. In contrast, with fine coding, individual cells can carry much more information, but downstream neurons have to resolve spike train structure to obtain it. Here, we set up a strategy to determine which codes are viable, and we apply it to the retina as a model system. We recorded from all the retinal output cells an animal uses to solve a task, evaluated the cells' spike trains for as long as the animal evaluates them, and used optimal, i.e., Bayesian, decoding. This approach makes it possible to obtain an upper bound on the performance of codes and thus eliminate those that are insufficient, that is, those that cannot account for behavioral performance. Our results show that standard coarse coding (spike count coding) is insufficient; finer, more information-rich codes are necessary.

Bayesian | ganglion cells | ideal observer | neural coding | population coding

One of the most pressing problems we face in neuroscience is determining what the neural code is. We know that neural signals come in the form of trains of action potentials, but we do not know what the unit of information is. Is it the number of spikes produced over some behaviorally relevant time interval (e.g., the length of a saccade), or is it the individual spike or some pattern of spikes? Several variations of the latter have been proposed, including patterns across spike trains and within them. (For a discussion of proposed codes and their relative merits, see refs. 1–10.) Getting a clear answer as to what the unit of information is affects essentially all work in systems neuroscience. For experimental work, it tells us what resolution we should use for analyzing data, and for theoretical work, it tells us what quantity we need for performing neural computations.

At first glance, it might seem that determining what the unit of information is—at least for a given brain area—is a straightforward problem, one that could be addressed as follows: Give an animal a task to perform and measure its performance; then take the spike trains the animal uses to perform the task and decode them several times, each time making a different assumption about what the unit of information is. For example, first assume it is the number of spikes in a relatively long time interval, such as the time of a stimulus presentation or a saccade, then assume it is the number of spikes in a shorter time interval, etc. Then, with each assumption, measure how well the task was performed and compare it to the performance of the animal.

Although this approach is straightforward in principle, it is not so straightforward in practice. To get a definitive answer, several conditions have to be met. First, the number and distribution of cells used for the decoding has to be the same as the number and distribution of cells the animal uses. To give a simple example for why this matters—the underlying idea behind coarse coding is that individual cells by themselves do not carry much information, but, together, as a population, they do and could be sufficient. Unless one records from all the cells the animal uses to solve a task, one cannot reject this notion and assert with any certainty that a finer code (i.e., one that carries more information per cell) is needed. Second, the length of time over which the spike trains are evaluated has to be the same as the length of time the animal uses. When an animal examines a stimulus, it typically looks multiple times. Unless one accumulates data from multiple looks and multiple cells, again, one cannot rule codes out. Finally, the last condition is that the decoding algorithm used to test codes has to be at least as good as the one the animal uses. Because no one knows the algorithm the animal uses, the only option is to use optimal (i.e., Bayesian) decoding—a strategy that extracts as much information from the spike trains as can be extracted (ref. 11; see also refs. 12–15).

If these conditions are met, one gets an upper bound on the performance of a code. One is basically giving the code its best chance—assessing it using the same number and distribution of cells the animal uses, accumulating data for as long as the animal does, and using a decoding strategy that's as good or better than the one the animal uses. If an upper bound on a code's performance is obtained, and that upper bound falls short of the animal's performance, then that code can be ruled out.

Our aim here was to set up a scenario where these conditions could be met. One of the few places this can be done is the retina. Here's the reasoning: First, the retinal output cells, the ganglion cells, form a bottleneck in the nervous system—they are the sole source of visual information to the brain. This means that all of the cells needed for the recording are confined to a small, well-defined location. Second, the transfer of signals from the retina to the brain is feedforward. Because there is no feedback, the retina can be removed from the animal and recorded from in vitro. This offers a significant advantage, because it allows multielectrode arrays to be used, making it possible to obtain a dataset large enough to match the number and distribution of cells the animal uses (16–18). Third, the number of cells needed can be controlled by regulating the size of the stimulus. A stimulus subtends a certain number of degrees of visual angle, which corresponds to a known area on the retina and, therefore, a known number of ganglion cells (19). Finally, the length of time over which data needs to be collected can be controlled by regulating the duration of the stimulus and the number of times it is presented.

We thus set out to measure the performance of a set of widely proposed codes using a combined in vivo and in vitro experi-

Author contributions: S.N. designed research with additions from A.J.L. and P.E.L.; A.L.J., G.F., N.M.A., and G.T.P. performed research; R.M.D. contributed new reagents/analytic tools; A.L.J., P.E.L., and S.N. analyzed data; and S.N. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: shn2010@med.cornell.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0900573106/DCSupplemental.

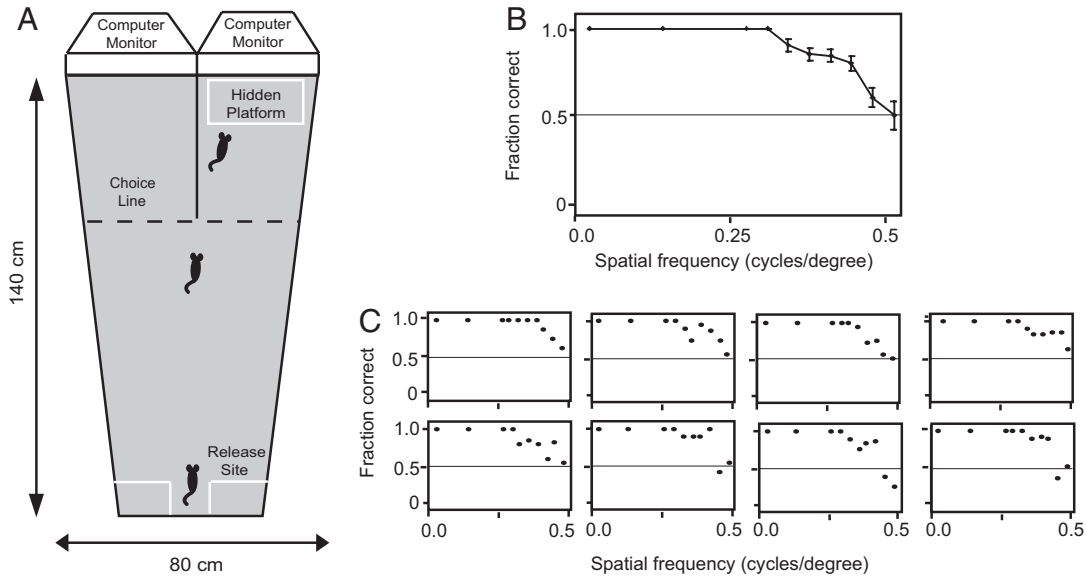

Fig. 1. Performance of the animal on a 2-alternative, forced choice visual discrimination task. (*A*) Schematic of the task. The animal is dropped into a pool, where the only escape is a hidden platform. The clue to the location of the platform comes from 2 computer monitors. One shows a grating, the other a uniform gray field, and the location of the platform is associated with the grating. If the animal can distinguish the grating from the gray image, it can find the platform and escape. Each image subtended 10 × 15° of visual angle at the choice line (see figure), was presented 8 times, and each presentation lasted for 300 ms. A barrier was set up so that the animal could see only 1 image at a time. (*B*) Mean performance for 8 animals, plotted as the fraction of times the animal chose the grating as a function of spatial frequency. Error bars were computed using binomial statistics. The standard deviation was $[P(1\text{-}P)/n]^{1/2}$, where P is the probability of choosing the grating and n is the number of trials (n ranged from 10–115 trials/spatial frequency). (*C*) Performance of each animal shown separately.

ment. We started in vivo: We first measured the performance of an animal on a visual task (Fig. 1). The animal was the mouse, and the task was a 2-alternative, forced choice visual discrimination task developed for rodents (20). On each trial, the animal was presented with 2 stimuli, a uniform gray field and a sinusoidal grating, with the spatial frequency of the grating chosen from a discrete set of uniformly spaced frequencies. The animal's assignment was to determine which of the 2 was the grating. Fig. 1*A* shows a schematic of the task, and Fig. 1*B* shows the behavioral performance for a set of 8 animals. As shown in the figure, performance was high when the spatial frequency of the grating was low and dropped to chance when the spatial frequency was high [≈0.5 cycles per degree (cpd)].

We then presented the same task to the retina in vitro, matching, of course, all stimulus conditions (image size, duration, intensities, spatial frequencies, phases and contrasts; see *Methods*) and decoded the ganglion cell spike trains. We decoded them several times, each time making a different assumption about what the unit of information is—that is, each time parameterizing the spike trains in a different way. We started by parameterizing the spike trains with just a single number, the number of spikes in a stimulus presentation. This is a simple coarse code, a spike count code. We then parameterized the spike trains finely, such that the unit of information was the individual spike. We referred to this as a spike timing code (also commonly called an instantaneous rate code). For this, the underlying assumption is 2-fold—that the unit of information is the single spike, and that the occurrence of a spike does not depend on the occurrence of other spikes. Last, we took into account temporal correlations and referred to this as a temporal correlation code. For this, the underlying assumption is also that the unit of information is the individual spike, but this time the occurrence of a spike does depend on the occurrence of other spikes—in this case, we assumed a dependence on the time of the previous spike (See *Methods* for a mathematical description of each code). In sum, then, we decoded ganglion cell spike trains using 3 widely proposed codes, ones that follow a natural progression from simple to more complex, and measured each one's performance against the performance of the animal.

The results are shown in Fig. 2. The spike count code (blue trace) performed much worse than the animal (black trace) ($P \ll 0.0001$, psignifit test (21), see *Methods*). Although the animal performed the task well (>70% correct up to 0.4 cpd), the spike count code maintained this level only up to ≈0.1 cpd. This argues that the animal cannot be using this code. The argument is strong because, as mentioned above, the code was given its best chance, that is, it was assessed using the same number and distribution of cells the animal uses (pooling the cells optimally), taking data for as long as the animal does (again, pooling optimally) and, finally, using optimal decoding (optimal for that code). This means that for the animal to reach its observed behavioral performance level, it has to be using a more information-rich code. This latter assertion follows because there is no other way for the animal to get the information: (*i*) as mentioned above, the retinal output cells are the sole source

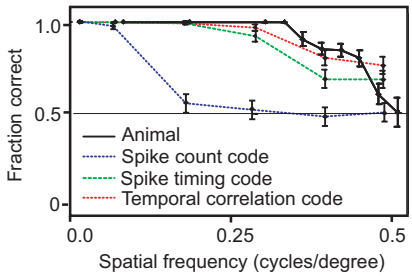

Fig. 2. The spike count code and spike timing (instantaneous rate) code performed worse than the animal, whereas the temporal correlation code (within spike train temporal correlations) did not (spike count code: $P \ll 0.0001$; spike timing code: $P < 0.02$; temporal correlation code: $P > 0.3$). As in Fig. 1, error bars were computed using binomial statistics. The standard deviation was $[P(1 - P)/n]^{1/2}$, where P is the probability of choosing the grating and n is the number of trials (for the codes, $n = 98$ trials/spatial frequency; for the animals, n ranged from 10–115 trials/spatial frequency).

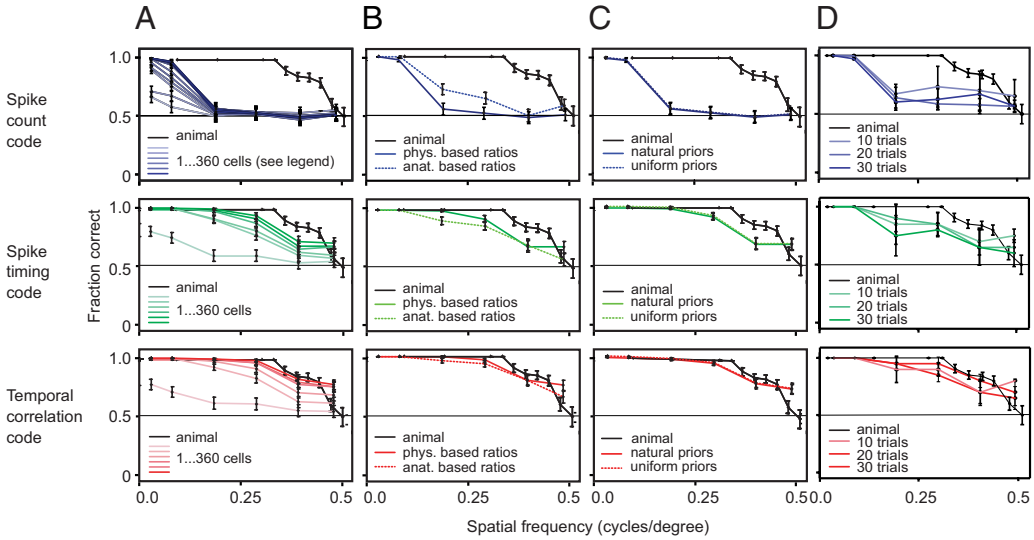

Fig. 3. The failure of the spike count code and the spike timing (instantaneous rate) code was robust to potential errors in the estimation of the critical parameters: cell number, cell distribution, priors on the stimulus, and shapes of the response distributions. (*A*) Performance of spike count and spike timing codes remained worse than the animal with cell numbers up to 360 (the upper bound estimate on cell number) (spike count: $P \ll 0.0001$; spike timing: $P < 0.02$); whereas performance of the temporal correlation code reached that of the animal, i.e., was not significantly different ($P > 0.3$). For spike count, the traces indicate 1, 2, 4, 8, 16, 32, 64, 128, 256, 300, and 360 cells. For the spike timing and temporal correlation codes, the traces indicate 1, 64, 128, 256, 300, and 360 cells. Note that for all 3 codes, performance became very slowing growing at numbers well below 360. Error bars were computed as in Fig. 2. (*B*) Performance for all codes shifted slightly when the distribution of cell classes was drawn from anatomical versus physiological estimations (see *SI Appendix*, section D for the distributions), but the conclusions remained the same: The spike count and spike timing codes still performed worse than the animal; the temporal correlation code did not. (*C*) Performance of the codes remained essentially the same whether uniform or natural priors were used; again, all conclusions remained the same. (*D*) Conclusions were not changed when the number of trials used to build the response distribution was systematically varied; error bars here indicate the standard deviation for 3 cross validations).

of visual input to the brain, and (*ii*) the brain can not create information de novo. As indicated by the data processing inequality, a well-known theorem in signal processing (22), information cannot be generated after the fact by postprocessing; a system can manipulate the information it receives, perform computations on it, etc., but it can not create new information.

The second result is that the spike timing (or instantaneous rate) code (green trace) also performed worse than the animal ($P < 0.02$, psignifit test). Note, though, that this code carries much more information than the spike count code (>70% correct at 0.3 cpd versus 0.4 cpd for the animal). Put in practical terms, if one were to build a retinal prosthetic with a spike count code, it would fall substantially short, but if one were to build a prosthetic with a spike timing code, it would put the animal within striking distance of normal acuity.

Finally, the last result is that the temporal correlation code (red trace) did perform the task as well as the animal ($P > 0.3$, psignifit test). As mentioned above, this is a simple within-spike train temporal correlation code, essentially a spike timing code with a soft refractory period. Although this result does not prove that this is the code the animal uses, it does show that it carries sufficient information and constitutes a viable candidate code. (See *SI Appendix*, Fig. S5 for a complete set of rasters for several cells; the figure shows, at the raw data level, how different codes perform.)

How robust are these results? The answer depends on how well they stand up to potential errors in the estimates of the critical parameters, specifically, the estimates of cell number, cell distribution, priors on the stimulus, and shapes of the response distributions. The first source of potential error is in the estimate of the number of ganglion cells the animal uses. The stimulus covers 0.144 mm^2 of retina. Two recent electron microscopic estimates of ganglion cell number (23, 24) indicate a range of 300 to 360 cells for this area (see *SI Appendix*, section D. We measured the performance of each code using both numbers, and there was essentially no difference (Fig. 3*A*) (the spike count and spike timing codes still performed worse than the animal, $P \ll 0.0001$ and $P < 0.02$, respectively; the temporal correlation code did not ($P > 0.3$). The figure shows performance as a function of cell number; as indicated in the figure, performance growth slows down at numbers much lower than these.

The second source of potential error is in the estimation of the ganglion cell distribution. Physiological studies in mouse show that the distribution is skewed toward ON-type ganglion cells (16–18). Anatomical studies in mouse, though, suggest a more even representation (particularly, equal numbers of ON and OFF cells (25), with deviations reflecting differences in coverage, i.e., cell types with larger dendritic trees occur proportionally less often than those with smaller ones. We measured the performance of each code, building the distribution of cell types both ways (Fig. 3*B*) (see *SI Appendix*, section D for the distributions). Although the choice of distribution shifted the performance slightly, the conclusions remained the same: The spike count and spike timing codes still performed worse than the animal (spike count, $P \ll 0.0001$; spike timing code, $P < 0.02$); the temporal correlation code did not ($P > 0.3$). (For decoding results with other distributions and performance curves for all individual cells in the dataset, sorted by cell class, see *SI Appendix*, section D.)

The third issue is the estimation of stimulus priors. We measured performance using both uniform priors (P(gray) = 1/2; $P(k)$ is a constant, where k is spatial frequency) and natural priors (P(gray) = 1/2; $P(k) \propto 1/k^2$), the latter following from ref. 26. As shown in Fig. 3*C*, column 3, there was essentially no effect: The spike count and timing codes still performed worse than the animal (spike count, $P \ll 0.0001$, spike timing, $P < 0.02$); temporal correlation code did not ($P > 0.3$). (This is not surprising because the priors become relevant only if very little data are used to decode the spike trains, which was not the case here.)

Finally, the last issue concerns the estimation of the response distributions. Because a Bayesian (probabilistic) decoding

method is being used, the response distribution for each stimulus must be estimated, and the quality of the estimate depends on the number of responses. An insufficient number could lead to a mis-estimate of code performance; both an underestimate and an overestimate are possible (see *SI Appendix*, section E for discussion). To address this, we ran the analysis such that the response distributions were built with different numbers of stimulus repeats (Fig. 3*D*). The results show that for the spike count and spike timing codes there was no significant trend as the number of repeats was increased, that is, the performance of these 2 codes did not significantly change, and both remained below the performance curve of the animal. For the temporal code, there was also no significant trend as the number of repeats increased, but, here, nearly all points in the performance curves came into contact with the animals' behavior curve. This supports the notion that this code cannot be ruled out and stands as a viable candidate code. Note that as an additional check we performed this analysis with multiple cross validations; this is represented by the error bars in panel d. This further demonstrates the robustness of the results: Even when the variance that occurs with different cross validations in taken into account, the differences in the performances of the codes is clear: The spike count code performs considerably worse than the animal, the spike timing code performs slightly worse, and the temporal correlation code reaches the animals' performance.

Discussion

A critical problem in systems neuroscience is determining what the neural code is. Many codes have been proposed—coarse, fine, temporal correlation codes, synchronous firing codes, among others. The space of candidates has grown as more and more studies have shown that different aspects of spike trains can carry information (see refs. 1–10) Our aim here was to shrink the space of possibilities, to set up a rigorous strategy for eliminating codes so we can close in on those that are truly viable. The strategy was to obtain an upper bound on the performance of each code and compare it to the performance of the animal. The upper bound was obtained by measuring code performance using the same number and distribution of cells the animal uses, the same amount of data the animal uses, and a decoding strategy that is as good or better than the one the animal uses. If the upper bound performance falls short of the animal's performance, the code can be eliminated, because this indicates very strongly that the animal cannot be using it.

We tested 3 widely proposed codes, referred to as a spike count code, a spike timing code, and a temporal correlation code. They followed a natural progression, summarized as follows. For the spike count code, the relevant quantity in the spike train was assumed to be spike number. For the spike timing code, also referred to as an instantaneous rate code, the relevant quantity was assumed to be the timing of the spikes, and the occurrence of a spike did not depend on the occurrence of other spikes. Last, for the temporal correlation code, the relevant quantity was assumed to be the timing of the spikes, but, here, the occurrence of a spike was assumed to depend on the occurrence of other spikes; the specific dependence was on the time of the previous spike.

Our results showed that 2 of the codes, the spike count and spike timing codes, did, in fact, fall short. The performance of the spike count code fell substantially short, as shown in Figs. 2 and 3. This result also held when spikes were counted in windows smaller than the length of the stimulus presentation, indicating that the failure of this code was not being exaggerated by counting spikes in the full 300-ms window. Even when spikes were counted only in 100 ms and 50 windows, the spike count code performed substantially worse than the animal (see *SI Appendix*, section I). The second result was that the spike timing code also fell short. Note, though, that the failure of this code was much less than that of the spike count code (see Figs. 2 and 3). Finally, the last result was that the temporal correlation code did perform as well as the animal. Although this does not demonstrate that this is the code the animal uses, it does show that it carries sufficient information and constitutes a viable candidate code.

These findings have significant implications for how downstream neurons must perform their computations—they argue that simple coarse coding algorithms built around spike counting, pooling, etc, are not realistic, at least at the retina/brain interface, and new models—those that take into account additional features of the spike train—need to take their place, because these additional features carry essential meaning (i.e., the codes that do not have them do not pass).

The findings also raise the intriguing issue of generalization. The problem of finding the neural code has often been likened to the problem of finding the genetic code, but although there is one genetic code (the relevant quantity or "unit of information" is always 3 nucleotides = 1 codon), it is not clear that there will be 1 neural code. The results here apply to the transfer of information from the periphery to the brain, a transfer that may require a particularly information-rich code. There is the intriguing possibility that the brain switches coding strategies when faced with problems with different constraints (e.g., high level visual processing, perception).

We conclude with 2 caveats. First, because we use electrophysiological recording, which has inherent limitations, we cannot completely eliminate the possibility that the animal has access to a high performing cell class that we can not detect. However, the dataset was large, it contained all of the known cell classes reported for mouse (16–18) and even extreme perturbations to the distributions of the cell classes, perturbations that exaggerated the performance of the highest performing classes (see *SI Appendix*) did not overturn the conclusions, providing strong evidence that this possibility is small.

Second, although we followed a progression from simple to complex codes—from spike count to spike timing to spike timing with temporal correlation—we were not able to test all possible codes in between, because the scope is too large; therefore, we indicate that other spike pattern permutations (e.g., a coarse code with temporal correlations, a code with multicell noise correlation that cannot be accounted for by pairwise noise correlations, etc.) remain candidates for testing. Our aim was to present a generalizable framework for testing neural codes and to test those most widely proposed. Our results showed that 2 of these codes failed—that is, they failed to account for behavioral performance—but a third code succeeded and stands as a viable candidate code.

Methods

Matching Stimulus Parameters in Vitro to Those in the Behavioral Task. ***Stimulus size.*** In the mouse, 1° of visual angle corresponds to 31 μm on the retina (19); thus, each stimulus was 310 μm $\times$ 465 μm = 0.144 mm^2 on retina both in vitro and in vivo.

Stimulus presentation. Stimuli were presented in flashes, both in vitro and in the behavioral task (flashed on for 300 ms, then off for 1,200 ms, for 8 repeats). The stimuli were presented in flashes to circumvent the problem of producing saccadic eye movements in the in vitro condition. Briefly, when an animal views a stimulus for more than a few hundred ms, it makes saccades (saccade frequency ranges from $\approx$3 times per second in primates (27, 28) to about once every 2–3 s in rodents (29, 30). By presenting the stimuli in flashes both in the behavioral task and in vitro—flashes that are shorter than the time between saccades—we equalize the 2 conditions, that is, we force both the retina in the dish and the retina in the animal to perform the task using short stimulus snapshots. Note that finer eye movements, such as microsaccades, have been found not to exist in nonfoveate retinas, including mouse and rabbit (31–33).

Stimulus intensity. Mean stimulus intensity in vivo was 3.76 lumens/mm^2 at the retina; intensity when the stimulus was off was 7.52 $\times$ 10^{-3} lumens/mm^2.

Mean stimulus intensity in vitro was 6.85 lumens/mm^2 at the retina; intensity when the stimulus was off was 1.37×10^{-2} lumens/mm^2. For the calculations used to match the intensity in vitro to that in vivo, and for a measure of the robustness of the conclusions of this article to the accuracy of this matching (see *SI Appendix*, section F).

Light adaptation conditions. To control light adaptation conditions, the pool in the behavioral task was covered with dark, 1-way-visible drapes. If the animal's eyes shifted away from the stimulus to the walls or water and then back again, it was shifting from a dark to a light condition, similar to the condition produced by the stimulus, which oscillated from dark to light. The conditions in vitro were matched to this by placing the retina in a chamber surrounded by a dark curtain.

Spatial frequencies, contrast and phase. In the behavioral task, the spatial frequencies ranged from 0.035 cpd to 0.5 cpd. The same range was used in vitro, but with fewer intervals ($n = 6$ spatial frequencies in vitro versus $n = 10$ in vivo). Grating contrast, both in the behavioral task and in vitro, measured as $L_{max} - L_{min}/L_{max} + L_{min}$, where L_{max} and L_{min} were the maximum and minimum intensities, was 99.6%. With respect to phase: in the behavioral task, each spatial frequency was presented at one phase; however, because eye movements likely occur between stimulus flashes (29, 30) >1 phase was likely received. To accommodate this in vitro, each spatial frequency was presented at multiple (3) phases (e.g., see *SI Appendix*, Fig. S5). All procedures were in accordance with National Institutes of Health guidelines.

Decoding Ganglion Cell Spike Trains. As mentioned in the main text, we decode ganglion cell spike trains using a Bayesian approach because it allows us to extract as much from the spike trains as can be extracted (i.e., it maximizes the fraction of correct predictions one can extract from the spike trains) (11). It is a direct approach, with no intervening steps, such as stimulus reconstruction, etc., that can lead to a loss of information. We begin by estimating the probability that a particular stimulus was presented, given that a particular set of ganglion cell responses occurred. This probability is denoted $P(s|r)$, where s is the stimulus and r is the set of ganglion cell responses. We find $P(s|r)$ by presenting each stimulus repeatedly, recording the resulting ganglion cell responses, and estimating the conditional response distribution, $P(r|s)$. We then use Bayes theorem, $P(s|r) = P(r|s)P(s)/P(r)$, to determine $P(s|r)$ from $P(r|s)$.

Given $P(s|r)$, we then use it to perform the task. In the task, 2 stimuli, a grating and a gray screen, are presented. Each produces a set of responses. The question we ask in the task is, Which of the 2 sets of responses corresponds to the grating? Letting r_1 and r_2 be the responses to the 2 stimuli, s_1 and s_2 respectively, we answer this by comparing $P(s_1 = \text{grating}|r_1)$ to $P(s_2 = \text{grating}|r_2)$. If the first quantity is larger, we say r_1 corresponds to the grating; otherwise, we say r_2 does. (As is standard, we use half the responses, chosen at random, to generate $P(r|s)$, and the other half to perform the task.)

With the Bayes' formalism, we have a natural way to test different codes, because different codes correspond to different treatments of r. To test the spike count code, we treat r as spike count; to test the spike timing code, we treat r as a set of spike arrival times and assume that the occurrence of a spike is independent of the occurrences of other spikes; and to test the temporal correlation code, we also treat r as a set of spike arrival times, but, this time, assume that the occurrence of a spike is not independent of the occurrences of other spikes; the specific dependence we assumed was a dependence on the time of the previous spike on the same spike train. (Note that in all cases, the same length of response is used; the difference is in the treatment of the responses.)

Formally, we construct the response distribution, $P(r|s)$, for each code as follows: For the spike count code, $r = n =$ {the number of spikes in a stimulus presentation}. Because each trial of the task involves multiple cells and multiple stimulus presentations (i.e., multiple looks), we use n_{il} to denote spike count from cell i on look l, and write

$$p_{\text{spike count code}}(r|s) = \Pi p(n_{il}|s) \qquad [1]$$

For the spike timing code (also referred to as an instantaneous rate code), $r =$ {a list of spike times at resolution dt}, denoted t_{ijl}, where t_{ijl} is the jth spike on the lth look of neuron i. In this case,

$$p_{\text{spike timing code}}(r|s) = \prod_l \left[\prod_{ij} v_i(t_{ijl}|s)dt \right] \exp\left[-\sum_i \int_0^T dt v_i(t|s) \right] \qquad [2]$$

where $v_i(t|s)$ is the firing rate of cell i at time t, given that stimulus s was presented, and the upper and lower limits (0 and T) correspond to the start and end of each trial. $v_i(t|s)$ is found using a cubic spline parameterization following (34, 35) as this captures firing rate accurately using a small number of parameters; see *SI Appendix*, section A, this also includes a description of the terms in Eq. **2**.

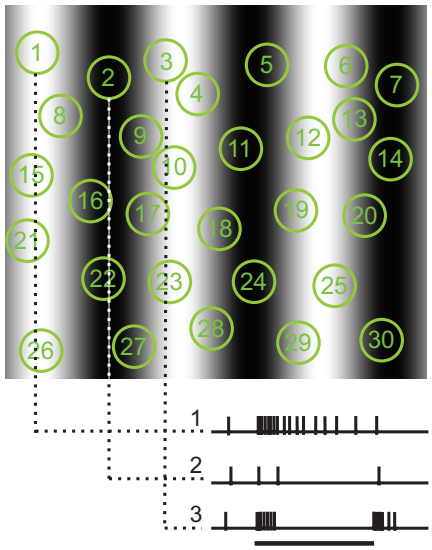

Fig. 4. Decoding ganglion cell responses, an illustration. The grating indicates a stimulus, and the circles indicate the receptive field positions of an array of cells relative to it. The traces below show the responses of several of the cells to one presentation of the stimulus. When we decode, we decode the joint response for the population—that is, we decode a vector, whose components are the responses of the cells on that stimulus presentation. When we assume the spike count code, the components are single numbers, spike counts. When we assume the spike timing or temporal correlation codes, the components are sets of spike arrival times. For the formal description of each code, see *Methods*.

For the temporal correlation code, again $r =$ {a list of spike times at resolution dt}, denoted t_{ijl}, but this time the firing rate has an additional dependence on the time of the previous spike on the same spike train,

$$p_{\text{temporal correlation code}}(r|s) = \prod_l \left[\prod_{ij} v_i(t_{ijl}, \tau(t_{ijl})|s) \right] \exp\left[-\sum_i \int_0^T dt v_i(t, \tau(t)|s) \right] \qquad [3]$$

where $\tau(t)$ is the time interval between t and the spike that preceded t on the same neuron. $v_i(t, \tau(t_{ijl})|s)$ is found using cubic splines (34, 35); see *SI Appendix*, section A.

Note that although we compute the likelihood from the marginals as indicated in the equations, when we decode, we only decode true joint responses, that is, we only decode joint responses that actually occur, as illustrated in Fig. 4. This approach has been tried and tested: Several studies have shown that decoding true joint responses using a decoder constructed this way (where the likelihood was computed from the marginals) has little or no effect on the estimation of the stimulus; see refs. 10, 36–39 and *SI Appendix*, section C.

Note also that Eqs. **1–3** do not imply that spatial relations are disrupted. (The subscripts on the equations indicate this.) Within each retina, each cell retains its spatial position relative to the stimulus and relative to the other cells (again, see Fig. 4).

Comparing Behavior and Decoder Performance. To compare behavior and decoder performance we used the standard Wichmann and Hill maximum-likelihood method (21) (see *SI Appendix*, section B).

Electrophysiological Recording. Recordings were made from isolated mouse retinas using a multielectrode array as described in ref. 10. Each recording contained 20–30 cells.

ACKNOWLEDGMENTS. We thank Peter Dayan, Horace Barlow, and Benoit Roux for helpful discussion. This work was supported by grants from National Institutes of Health and the Beckman Foundation (to S.N.).

1. Oram MW, Xiao D, Dritschel B, Payne KR (2002) The temporal resolution of neural codes: Does response latency have a unique role? *Philos Trans R Soc Lond Ser B* 357:987–1001.
2. Victor JD (1999) Temporal aspects of neural coding in the retina and lateral geniculate. *Network* 10:R1–66.
3. Borst A, Theunissen F (1999) Information theory and neural coding. *Nat Neuro* 2:947–957.
4. Johnson DH, Ray W (2004) Optimal stimulus coding by neural populations using rate codes. *J Comput Neurosci* 16:129–138.
5. Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: A rigorous definition. *J Comput Neurosci* 2:149–162.
6. Nirenberg S, Latham PE (2003) Decoding neuronal spike trains: How important are correlations? *Proc Natl Acad Sci USA* 100:7348–7353.
7. Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. *Curr Opin Neurobiol* 4:569–579.
8. MacLeod K, Backer A, Laurent G (1998) Who reads temporal information contained across synchronized and oscillatory spike trains? *Nature* 395:693–698.
9. Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. *Science* 252:1854–1857.
10. Nirenberg S, Carcieri SM, Jacobs AL, Latham PE (2001) Retinal ganglion cells act largely as independent encoders. *Nature* 411:698–701.
11. Ripley BD (1996) *Pattern recognition and neural networks* (Cambridge Univ Press, Cambridge, MA), pp 19–20.
12. Barlow HB (1977) in *Vertebrate Photoreception*, eds Barlow HB, Fatt P (Academic), pp 337–358.
13. Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: A comparison of neuronal and phychophysical performance. *J Neurosci* 12:4745–4765.
14. Geisler WS (2003) in *The Visual Neurosciences*, eds Chalupa L, Werner J (MIT Press, Cambridge, MA), pp 825–837.
15. Simoncelli E, Olshausen B (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
16. Balkema GW Jr., Pinto LH (1982) Electrophysiology of retinal ganglion cells in the mouse: A study of a normally pigmented mouse and a congenic hypopigmentation mutant, pearl. *J Neurophysiol* 48:968–980.
17. Carcieri SM, Jacobs AL, Nirenberg S (2003) Classification of retinal ganglion cells: A statistical approach. *J Neurophysiol* 90:1704–1713.
18. Stone C, Pinto LH (1993) Response properties of ganglion cells in the isolated mouse retina. *Vis Neurosci* 10:31–39.
19. Remtulla S, Hallett PE (1985) A schematic eye for the mouse, and comparisons with the rat. *Vision Res* 25:21–31.
20. Prusky GT, West PW, Douglas RM (2000) Behavioral assessment of visual acuity in mice and rats. *Vision Res* 40:2201–2209.
21. Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 63:1293–1313.
22. Cover TM, Thomas JA (1991) *Elements of Information Theory* (Wiley, New York), pp 576.
23. Jeon CJ, Strettoi E, Masland RH (1998) The major cell populations of the mouse retina. *J Neurosci* 18:8936–8946.
24. Williams RW, Strom RC, Rice DS, Goldowitz D (1996) Genetic and environmental control of variation in retinal ganglion cell number in mice. *J Neurosci* 16:7193–7205.
25. Doi M, Uji Y, Yamamura H (1995) Morphological classification of retinal ganglion cells in mice. *J Comp Neurol* 356:368–386.
26. Ruderman DL, Bialek W (1994) Statistics of natural images: Scaling in the woods. *Phys Rev Lett* 73:814–817.
27. Gawne TJ, Richmond BJ, Optican LM (1991) Interactive effects among several stimulus parameters on the responses of striate cortical complex cells. *J Neurophysiol* 66:379–389.
28. Fuchs AF (1967) Saccadic and smooth pursuit eye movements in the monkey. *J Physiol* 191:609–631.
29. Bruckner R (1951) Slit-lamp microscopy and ophthalmoscopy in rat and mouse. *Doc Ophthalmol* 5–6:452–554.
30. Fuller JH (1985) Eye and head movements in the pigmented rat. *Vision Res* 25:1121–1128.
31. Collewijn H (1977) Eye- and head movements in freely moving rabbits. *J Physiol* 266:471–498.
32. Martinez-Conde S, Macknik SL, Hubel DH (2004) The role of fixational eye movements in visual perception. *Nat Rev Neurosci* 5:229–240.
33. Chalupa LM (2008) *Eye, Retina and Visual System of the Mouse* (MIT Press, Cambridge, MA).
34. DiMatteo I, Genovese CR, Kass RE (2001) Bayesian curve-fitting with free-knot splines. *Biometrika* 88:1055–1071.
35. Kass RE, Ventura V (2001) A spike-train probability model. *Neural Comput* 13:1713–1720.
36. Golledge HD, et al. (2003) Correlations, feature-binding and population coding in primary visual cortex. *NeuroReport* 14:1045–1050.
37. Petersen RS, Panzeri S, Diamond ME (2001) Population coding of stimulus location in rat somatosensory cortex. *Neuron* 32:503–514.
38. Averbeck BB, Lee D (2003) Neural noise and movement-related codes in the macaque supplementary motor area. *J Neurosci* 23:7630–7641.
39. Oram MW, Hatsopoulos NG, Richmond BJ, Donoghue JP (2001) Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J Neurophysiol* 86:1700–1716.