14. A. G. Knudson, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 820 (1971).
15. D. Mirchandani *et al.*, *Am. J. Pathol.* **147**, 92 (1995).
16. P. H. Gumerlock *et al.*, *J. Natl. Cancer. Inst.* **89**, 66 (1997).
17. C. J. Sherr, J. M. Roberts, *Genes Dev.* **13**, 1501 (1999).
18. D. M. Prowse, L. Bolgan, A. Molnar, G. P. Dotto, *J. Biol. Chem.* **272**, 1308 (1997).
19. C. A. Kim, J. M. Berg, *Nature Struct. Biol.* **3**, 940 (1996).
20. Results are expressed as mean ± standard error. Statistical analysis was performed using a two-way analysis of variance (ANOVA). For direct comparisons between groups, the Bonferonni correction has been applied.

21. We thank J. Friedman, M. Oren, N. Mohandas, and Z. Ronai for helpful advice and P. Dotto, B. Li, M. Mari, D. Musat, T. Ouchi, M. Ramirez, and P. Unger for technical assistance. Supported by grants from the NIH (DK37340 to S.L.F, CA79918 to A.C.F., and CA78207 to A.C.K.), Mount Sinai School of Medicine Dean's Research Incentive Fund (S.L.F.) and Child Health Research Center (5 P30 HD28822 to J.A.M.), Charles H. Revson Foundation (K.E.H.), Howard Hughes Medical Institute (G.N.), American Liver Foundation (G.N.), University of Newcastle upon Tyne Department of Medicine (H.L.R.), and the Peel Medical Research Trust, UK (H.L.R.).

18 September 2001; accepted 6 November 2001

Independent and Redundant Information in Nearby Cortical Neurons

Daniel S. Reich,[1,2]* Ferenc Mechler,[2] Jonathan D. Victor[2]

In the primary visual cortex (V1), nearby neurons are tuned to similar stimulus features, and, depending on the manner and time scale over which neuronal signals are analyzed, the resulting redundancy may mitigate deleterious effects of response variability. We estimated information rates in the short–time scale responses of clusters of up to six simultaneously recorded nearby neurons in monkey V1. Responses were almost independent if we kept track of which neuron fired each spike but were redundant if we summed responses over the cluster. Redundancy was independent of cluster size. Summing neuronal responses to reduce variability discards potentially useful information, and the discarded information increases with cluster size.

How do neurons in the sensory cortex work together to represent a stimulus? Cortical neurons with similar stimulus selectivities are found in close proximity to one another (*1*–*3*). This might reflect a mechanism of coping with large trial-to-trial variability in the responses of individual neurons: Downstream neurons could simply sum the activities of many neurons with similar sensitivities. However, because response variability is correlated across neurons (*4*, *5*), the ability of averaging to increase the signal to noise ratio is limited (*6*). Also, the fact that responses are variable does not imply that the cortex averages signals from multiple neurons, because averaging would ignore stimulus-related information encoded into which neuron fires each spike. Theoretical (*7*) and experimental (*8*, *9*) work has shown that neurons tuned to similar stimuli can convey largely independent information, especially when their responses are noisy.

Figure 1 describes two pairs of nearby V1 neurons that illustrate the range of behavior we encountered (*10*). The first pair (Fig. 1, A to D) responded robustly to the stimulus, as shown by the sharp and reliable firing rate fluctuations during typical 1-s segments (Fig. 1A). Although the neurons occasionally responded to similar stimulus features (double-headed arrow), their responses were usually distinct (single-headed arrow). Overlapping "off" subregions were evident in snapshots of the spatial receptive fields (Fig. 1B), obtained by cross-correlating the stimulus with the spike train (*10*). The "on" subregions, however, were on opposite flanks, indicating that the neurons conveyed at least some distinct spatiotemporal information. The second pair (Fig. 1, E to H) responded less robustly and more variably, and the spatial overlap was more complete.

The Pearson correlation coefficient is a measure of similarity between paired responses. We distinguished between signal correlations, which compare bin-by-bin average spike counts across trials, and noise correlations, which compare trial-by-trial deviations from the average response in each bin (*8*). The signals were essentially uncorrelated for the first pair [correlation coefficient (r) = −0.021] but highly correlated for the second (r = 0.52). The noise was uncorrelated for both pairs (r = −0.039 and −0.015, respectively).

Information rates were substantially higher for the first pair than for the second (Fig. 1, C and G) (*11*). To assess population coding, we compared two schemes: the summed-population code, which did not consider which neuron fired each spike, and the labeled-line code, which did (*12*). The summed-population information rate for the first pair was slightly higher than the information rate for the first cell alone, but the labeled-line information rate was 49% higher than the summed-population information rate. For the second pair, the difference between the summed-population and labeled-line information rates was only 5.6%. Ignoring which neuron fired each spike blurred distinctions between the responses of the first pair, resulting in a reduced multineuronal information rate. For the second pair, responses were more similar, and neuronal identity mattered less.

For each pair of neurons and code, we calculated a redundancy index as a normalized measure for comparing recording sites with vastly different information rates (*13*). The redundancy index was 0 when neurons carried independent information and 1 when the information was completely redundant. Figure 1, D and H, shows that, for both pairs, the redundancy index was higher for the summed-population code than for the labeled-line code. The difference was greater for the first pair, in which the responses were robust and distinct, so that combining spikes from those two neurons blurred distinctions and emphasized redundancy. The labeled-line code revealed the underlying independence of the responses and yielded a redundancy index of essentially 0. For the second pair, redundancy indices for both codes were near 0, meaning that the responses were almost independent, even for the summed-population code. This surprised us at first because the neurons responded to similar stimulus elements, and we gained little information when we paid attention to which neuron fired each spike. However, both signal and noise contributed to information rates, and when signals were small and correlated, and noise large and uncorrelated, information in the summed-population code was nonredundant.

We evaluated signal and noise correlations at a series of bin sizes for all pairs of neurons in our database (Fig. 2). Correlations were generally positive, but signal correlations were higher on all time scales. On short time scales, noise correlations clustered tightly about 0, whereas signal correlations were more widely distributed with positive median values <0.25. On longer time scales, median correlations reached higher levels for signal (~0.5) than for noise (~0.25). On long time scales, our results correspond approximately to other monkey visual cortex studies that used different stimuli and analyses (*3*, *8*, *14*).

Information rates and redundancy indices can be calculated for any cluster size, not just

---

[1] Laboratory of Biophysics, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA. [2] Department of Neurology and Neuroscience, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA.

*To whom correspondence should be addressed. E-mail: reichd@rockefeller.edu

pairs, and characterize the whole population. Figure 3 presents results from clusters of up to six simultaneously recorded nearby neurons. Scatter plots of information rates (top row) reveal that higher information rates were always obtained with the labeled-line code than with the summed-population code. The information rates for both codes monotonically increased with cluster size, without saturating, and were strongly correlated at all cluster sizes. Ignoring neuron identity led to a progressively larger loss of information with larger clusters. Although clusters with more than six neurons were not considered here, the trend suggests that the modest difference observed for pairs becomes much larger even for small populations.

The labeled-line redundancy indices were always smaller (bottom row) than were the summed-population indices, meaning that under this code, nearby neurons were nearly independent. The redundancy index comparison highlights the difference between the two codes. However, even with the summed-population code, redundancy indices were substantially smaller than 1, because noise tended to be independent across neurons (as exemplified by the second pair in Fig. 1). We found no synergistic coding in the responses to checkerboard stimuli.

Figure 4A shows the number of distinct sites at which clusters of each size were recorded; within each group, all subsets (pairs, triplets, etc.) were analyzed. Median information rates for the labeled-line code approximated those in the separate sum at all cluster sizes (Fig. 4B). The summed-population code transmitted progressively less information, which was noticeable even with triplets. The median redundancy was nearly independent of cluster size for both codes, and the labeled-line code was significantly less redundant than the summed-population code (redundancy index ~0.1 versus ~0.4) (Fig. 4C). Because of the positive redundancy indices (even for the labeled-line code), treating large clusters of neurons as if they were fully independent would overstate the amount of information conveyed.

Because information is logarithmic, and because the redundancy index was independent of cluster size, the difference in the number of stimuli that could be discriminated with each code increased exponentially with cluster size. Our rich dynamic stimulus made information saturation due to limited stimulus sets less likely than in previous studies (*15*), and we found no evidence of saturation with up to six neurons.

In the retina, pairs of nearby ganglion cells convey independent information (*9*), which suggests that the retinal code can be understood by treating each cell as an independent encoder. This seems advantageous, because ganglion cells are a wiring bottleneck in the visual system. V1 may not be under a similar constraint, and the near independence of nearby V1 neurons most likely arises from other principles. Because complex wiring induces considerable signal mixing, independence in the retina need not persist in V1.

In V1, nearby neurons share many but not all response properties (*3*). Our experiments did not identify the ways in which particular stimulus attributes are encoded by nearby neurons but did reveal that, whereas information conveyed on time scales <15 ms is largely independent, responses are 30 to 50% correlated on

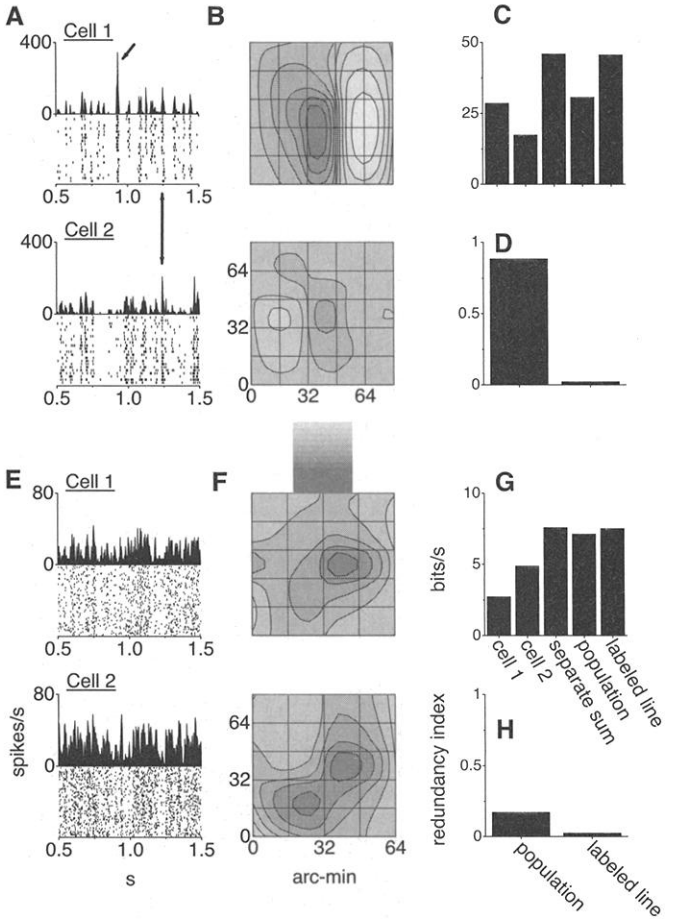

Fig. 1. Responses of two pairs of neurons in macaque V1. (**A** and **E**) Firing-rate histograms (3.7-ms bins) aligned with raster diagrams depicting spike times fired during 1-s segments of each trial. (**B** and **F**) Contour-map snapshots of each neuron's receptive field averaged over a selected 14.8-ms time window (*18*). Dark regions signify areas in which dark stimuli were excitatory and bright stimuli were inhibitory, and light regions signify the reverse. Time windows (chosen near the peak response) were as follows: (**B**) 44 to 59 ms, (**F**) 74 to 89 ms. Grayscale represents the change from background firing rate: (**B**) ±5 spikes/s, (**F**) ±0.4 spikes/s. (**C** and **G**) Information rates for each neuron alone, the separate sum, and each multineuron code (summed population and labeled line). (**D** and **H**) Redundancy indices for summed-population and labeled-line codes. Vertical axis scales for the firing- and information-rate plots are fivefold greater for the first pair than for the second pair.

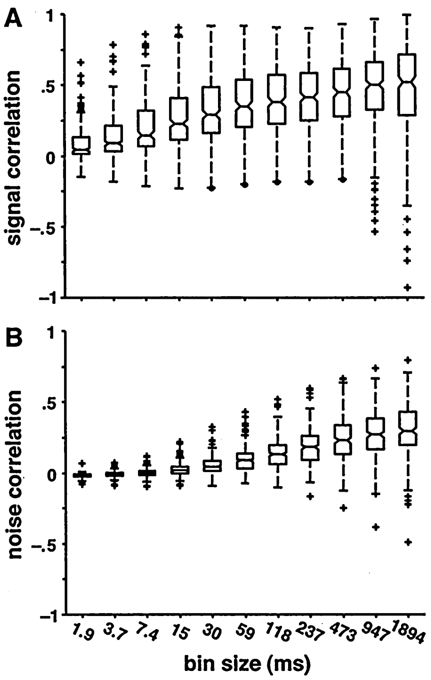

Fig. 2. Distribution of (**A**) signal and (**B**) noise correlations across 274 neuron pairs, parametric in bin size (ms). Box plots show medians and quartiles, notches are estimates of uncertainty in the medians, whiskers delineate 1.5 times the interquartile range, and + signs represent outliers.

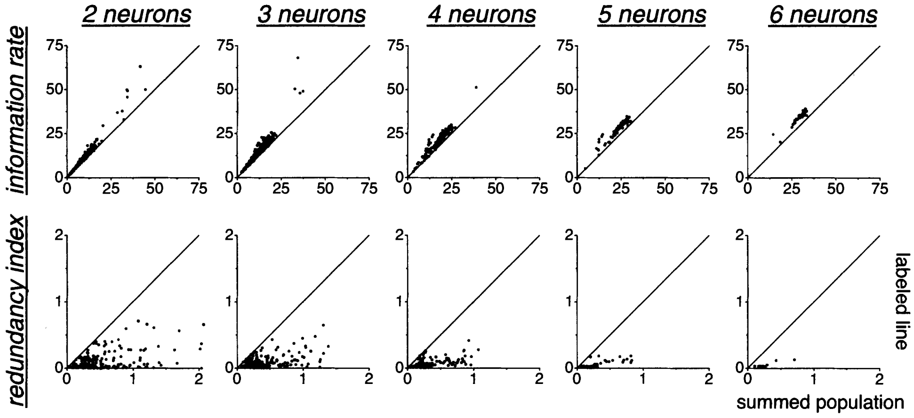

Fig. 3. Information rates (**top**) and redundancy indices (**bottom**) for the summed-population (horizontal axis) and labeled-line (vertical axis) codes for clusters of two to six neurons. Each point represents a different neuron cluster. Neurons could participate in several clusters of identical and different sizes, depending on the number of neurons isolated at that site.

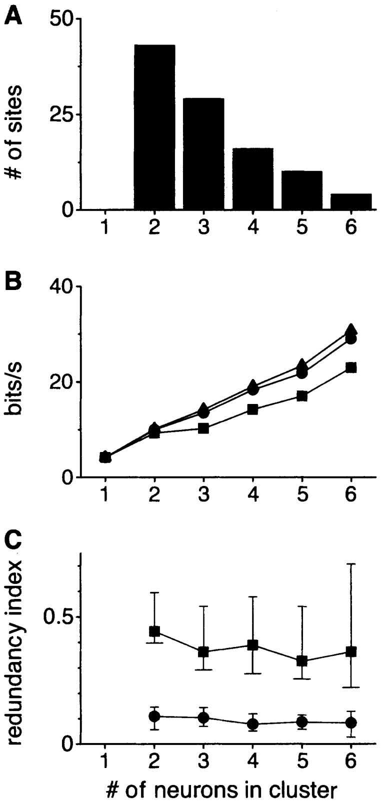

Fig. 4. (**A**) Distribution of the number of unique sites for clusters of different sizes. (**B**) Median information rates versus cluster size for the summed-population (square) and labeled-line (circle) codes and for the separate sum (triangle). Medians were first calculated at each recording site, and then grand medians were calculated across recording sites. (**C**) Median redundancy indices versus cluster size for the summed-population (square) and labeled-line (circle) codes. Error bars are 95% confidence limits on the medians and were derived from 1000 bootstrap resamplings.

time scales >60 ms. Whether the short-time scale information concerns stimulus attributes (such as spatial phase) for which preferences are not shared by nearby neurons remains to be determined.

Our results suggest that keeping track of which neuron fires each spike preserves a considerable amount of information already present in the responses. This strategy removes some redundancy across neurons with similar tuning properties, and it prevents an even greater information loss from summing responses of neurons with different selectivities. It also requires resources devoted to keeping track of disparate inputs, a tradeoff that might be costly. Although our experiments did not address mechanisms of information decoding in cortical networks, the existence of a complex dendritic machinery for processing synaptic inputs (*16, 17*) suggests that codes that use information available in the labeled-line code are plausible.

References and Notes

1. D. H. Hubel, T. N. Wiesel, *J. Physiol.* (*London*) **160**, 106 (1962).
2. V. B. Mountcastle, *J. Neurophysiol.* **20**, 408 (1957).
3. G. C. DeAngelis, G. M. Ghose, I. Ohzawa, R. D. Freeman, *J. Neurosci.* **19**, 4046 (1999).
4. E. Zohary, M. N. Shadlen, W. T. Newsome, *Nature* **370**, 140 (1994).
5. I. Lampl, I. Reichova, D. Ferster, *Neuron* **22**, 361 (1999).
6. M. N. Shadlen, W. T. Newsome, *J. Neurosci.* **18**, 3870 (1998).
7. L. F. Abbott, P. Dayan, *Neural Comput.* **11**, 91 (1998).
8. T. J. Gawne, T. W. Kjaer, J. A. Hertz, B. J. Richmond, *Cereb. Cortex* **6**, 482 (1996).
9. S. Nirenberg, S. M. Carcieri, A. L. Jacobs, P. E. Latham, *Nature* **411**, 698 (2001).
10. We made simultaneous tetrode recordings from clusters of individually isolated, nearby V1 neurons in sufentanil-anesthetized paralyzed monkeys (*18*). Stimuli were checkerboards in which the luminance in each check was modulated by a binary m-sequence (*19*) (mean luminance = 150 cd/m^2, contrast = 100%, check size = 16 × 16 arc-min, durations were either 60.6 or 7.6 s). Spike sorting (*20*), based on principal components of the waveforms, was semiautomatic and slightly modified from (*21*). Erroneous spike assignment could only reduce information transmitted through the labeled-line code and did not affect information transmitted through the summed-population code.
11. We estimated time-averaged instantaneous information rates by assessing variability across responses to multiple stimulus repetitions (*22*). See (*23*) for methodological details, including bias correction and time scale choice (typically 3.7 to 14.8 ms). We ignored response correlations on time scales longer than one bin, which were not present in the stimulus but could arise from intrinsic neuronal dynamics. Information rates at different instances are not additive, and the total transmitted information does not equal the product of information rate and time interval.
12. The summed-population code takes the response in each time bin to be the total number of spikes emitted across all neurons. The labeled-line code uses a vector of spike counts, with each dimension corresponding to a different neuron, and preserves stimulus-related information encoded into the neuron that fires each spike. The labeled-line information rates can be no lower than the summed-population rates, because the summed population can be recovered by summing across neurons in the labeled line. The rate at which information would have been conveyed if the neurons were completely independent—the separate sum—is obtained by adding together information rates for each neuron alone. This quantity is not an upper bound for the labeled-line code, because coding across neurons might be synergistic. Note that both codes consider responses in short time bins, rather than time-averaged firing rates, so that both are examples of temporal codes. Bias-correction methods are detailed in (*24*).
13. The redundancy index is defined as $(1 - I_c/I_s)/(1 - \max\{I_i\}/I_s)$, where I_i is the information rate of neuron i, I_c is the information rate in the code being considered, and I_s is the separate sum of information rates for each cell alone. The redundancy index can be >1 if stimulus-related information in jointly analyzed responses is contradictory or confusing, and it can be <0 if the neurons code synergistically. Other reports define similar quantities (*9, 25*).
14. W. Bair, E. Zohary, W. T. Newsome, *J. Neurosci.* **21**, 1676 (2001).
15. E. T. Rolls, A. Treves, M. J. Tovee, *Exp. Brain Res.* **114**, 149 (1997).
16. C. Koch, I. Segev, *Nature Neurosci.* **3** (suppl.), 1171 (2000).
17. S. R. Williams, G. J. Stuart, *J. Neurosci.* **20**, 8238 (2000).
18. D. S. Reich, F. Mechler, K. P. Purpura, J. D. Victor, *J. Neurosci.* **20**, 1964 (2000).
19. E. E. Sutter, in *Nonlinear Vision: Determination of Neural Receptive Fields, Function, and Networks*, R. B. Pinter, B. Nabet, Eds. (CRC Press, Boca Raton, FL, 1992), pp. 171–220.
20. D. S. Reich, thesis, The Rockefeller University, New York, NY (2001).
21. M. S. Fee, P. P. Mitra, D. Kleinfeld, *J. Neurosci. Methods* **69**, 175 (1996).
22. S. P. Strong, R. Koberle, R. R. de Ruyter van Steveninck, W. Bialek, *Phys. Rev. Lett.* **80**, 197 (1998).
23. D. S. Reich, F. Mechler, J. D. Victor, *J. Neurophysiol.* **85**, 305 (2001).
24. The multidimensionality of the labeled-line code increases calculation bias. Simulations revealed that this bias can substantially affect the noise entropy (*22*), a measure of trial-by-trial variability. To address this, we assumed that noise entropy (but not total entropy) is additive across neurons, which is equivalent to assuming that noise is uncorrelated on short time scales (justified empirically in Fig. 2); the resulting systematic error is likely to be small (*26*). This yielded an overestimate of noise entropy and an underestimate of information rate. However, because we used a data-dependent method to determine time-bin duration, our bias correction may have led to a more severe underestimate of information rate for the labeled-line code than for the summed-population code. Indeed, we occasionally found neuron clusters for which the calculated information rates were artifactually higher for the summed-population code; these neuron clusters were not considered.
25. C. K. Machens *et al.*, *J. Neurosci.* **21**, 3215 (2001).
26. S. Panzeri, S. R. Schultz, *Neural Comput.* **13**, 1311 (2001).
27. We thank B. Knight, A. Herz, C. Machens, M. Meister, S. Nirenberg, K. Purpura, and R. de Ruyter van Steveninck for advice. Supported by NIH grants GM07739 and EY07138 (D.S.R.) and EY9314 (J.D.V.).

29 August 2001; accepted 18 October 2001