

Gauging sensory representations in the brain

Giedrius T. Buračas and Thomas D. Albright

Howard Hughes Medical Institute,
Salk Institute for Biological Studies, La Jolla

Running title: Gauging neuronal representations

Address correspondence to:
Giedrius T. Buračas
Systems Neurobiology Lab – VCL
Salk Institute
10010 North Torrey Pines Rd.
La Jolla, CA 92037
Phone: (619) 453-4100 ext. 1471
Email: giedrius@salk.edu

The stream of information that enters a sensory system is a product of the ecological niche and the way in which the information is sampled. The most salient feature of this sensory stream is rich temporal structure caused by changes in the environment and self-motion of sensors (e.g. rapid eye or whisker movements). Recent years have seen substantial progress in understanding how such rapidly varying stimuli are represented in responses of sensory neurons. The critical observation that emerged from these studies, is that in a large variety of sensory systems individual action potentials convey substantial amounts of information, which permits the discrimination of rapidly varying stimuli with high temporal precision.

Cognitive neuroscience hinges on the doctrine that the brain represents the world in patterns of neuronal activity. Recently this "representationalist" *credo* was forcefully restated by Blakemore and Movshon (1995): "The task of sensory systems is to provide a faithful representation of biologically relevant events in the external environment. <...> These representations are <...> (rich) because they contain representations of objects, states, and events that are abstracted from the primitive sensory signals; they are (simple) because they represent the distillation of the vast quantities of raw measurement information offered to the central nervous system by each sensory surface." In more technical terms, the raw information that enters sensory systems is processed so that most of the biologically irrelevant information is discarded and the representations of relevant information are endowed by invariances (e.g. Albright, 1992). As a consequence, neuronal responses in higher sensory areas are tuned to abstract features. In the present review we offer a cursory update on the advances that have taken place since the last TINS review of related issues (Bialek and Rieke, 1992). These advances were stimulated by novel information-theoretic approaches that can be used to gauge the "richness" and "simplicity" of neuronal representations of stimuli that vary on ecologically relevant time scales (DeWeese, 1996; Stevens and Zador, 1996; de Ruyter van Steveninck et al., 1997; Strong et al., 1998; Zador, 1998) and by accumulation of neurophysiological data from a wide variety of species.

Neuronal encoding vs. decoding of ecologically-relevant stimuli

The most frequently asked questions about neuronal representations are: What is represented? How is it represented? The first question has inspired efforts to make sense of a neuronal message by decoding it. The decoding step has conventionally been preceded by an answer to the second question, which addresses the encryption of sensory messages in neuronal responses. Until recently, beliefs about how information in a neuronal response is encoded were influenced by an early observation that multiple presentations of identical stimuli yield highly variable numbers of action potentials (e.g., Adrian, 1926). In order to get a good estimate of average firing rates in response to a particular stimulus the duration of stimulus presentation had to be extended to lengths on the order one second. This manipulation implanted a tacit assumption that neurons encode stimulus values in average firing rates; according to this assumption information about exact spike arrival times is largely irrelevant.

A decade ago this assumption was challenged by a series of experiments that exploited a newly developed method to estimate of how well sensory systems can discriminate among stimuli that vary at rates commonly found under natural conditions (e.g. Bialek, and Rieke, 1992). This method took the "organism's point of view," in which

an observer must reconstruct the stimulus solely from the observed spike train, without assumptions about the encoding scheme (for explanation of how discrimination performance can be evaluated by means of information theory see Box 1 and Fig 1a,b). The logic of this approach thus defies the conventional wisdom whereby the question "What is encoded?" is addressed only after the code is cracked. The results of stimulus reconstruction experiments conducted with invertebrates and amphibians demonstrated that sensory neurons of these organisms could faithfully track stimuli that varied on an ecologically-relevant time scale (e.g. ~30 msec) with high efficiency. Surprisingly, the stimulus could be estimated with reliability approaching physical limits, despite the fact that its rate of change exceeded that of spike generation. Since these neurons did not produce enough spikes in sufficiently short intervals for firing rate to be judged, a code based on mean firing rate could be ruled out for these particular nervous systems (see Table I). More recently experiments with rapidly varying stimuli were performed in tiger salamander and rabbit retinas. The results were comparable: In all of these systems individual spikes were strikingly informative (conveying in the range of 1-3 bits of information about the stimulus), suggesting that this might be a property shared by all sensory systems (see Table I, middle; Berry et al., 1997; Warland et al., 1998). It remained to be seen, however, whether spikes in mammalian neocortex would be capable of matching the discrimination power of retinal spikes upstream.

Studies of the informativeness of neocortical neurons that were conducted using constant stimuli routinely yielded about 0.5-2 bits/sec, suggesting that these neurons could discriminate among only 2-4 stimulus values (Table I, top). In effect, an average neocortical spike, in contrast to its non-cortical relatives, was deemed to be worth very little – only about 0.01-0.2 bits/spike. Inspired by the results from non-mammalian neural systems, two laboratories recently set out to test the limits of the informativeness of neocortical neurons (Bair et al., 1997; Buračas et al., 1996;1998). When motion sensitive neurons of the middle temporal area (area MT) of primate (rhesus monkey) visual cortex were presented with a rapidly varying stimulus and the stimulus reconstruction approach was applied, the net worth of a neocortical spike jumped by nearly two orders of magnitude to ~0.6 bits/spike. This was true despite the fact that one laboratory used anesthetized monkeys and the linear stimulus reconstruction method (Bair et al., 1997), and another used alert monkeys and a nonlinear reconstruction method (Buračas et.al., 1998). A casual comparison of panels b, c and d in Fig 2 reveals that spikes recorded from area MT become very informative when the neuron is challenged by a rapidly varying stimulus: Although spike count variability is equally high when the stimulus is moving at a constant rate (Fig 2,d; variance/mean=1.3) and when speed of motion is highly variable (Fig 2,c; variance/mean=1.4), in the latter case (Fig 2c) each spike conveys much more information (nearly one bit) about the stimulus.

Neurosemantics: content in spikes

What is the content of the information conveyed by these informative spikes? When asked to make fine discrimination about the exact stimulus value – the "what" of the stimulus – the spikes of neocortical neurons failed miserably, but became very informative when challenged with time-varying stimuli. Indeed, the spikes produced in response to varying stimuli were generated with high temporal precision (0.9-5 msec; cf. Timing precisions of other neural systems in Table I), which supports the idea that most of the information carried by spikes was about stimulus timing. The same preference of "when"

over “what” was reported for salamander retinal ganglion cells (Berry et al., 1997) and motion sensitive neurons of the fly (de Ruyter van Steveninck et al., 1997), suggesting that this preference might reflect a fundamental principle underlying the design of sensory systems. Indeed, this preference for “when” over “what” is a strategy that maximizes transmitted information (Softky, 1996). For example, if a neuron doubles the precision with which it represents stimulus orientation (stimulus “what”), information increases by only one bit (i.e. information grows logarithmically with message precision). By contrast, if the same neuron doubles its timing precision, effectively halving its transmission time, information rate doubles (information grows linearly with transmission speed)! This biased allocation of channel bandwidth in favor of timing is especially useful for sensory systems that rely on large populations of neurons to convey the signal: One can readily increase the precision of the stimulus estimate by simply pooling more neurons (for a review of population coding see Oram et al., 1998). By contrast, the temporal precision of responses can not be increased easily in this way.

In conclusion, when sensory systems confront stimuli that vary on ecologically-relevant time scales, individual neuronal spikes convey in the range of 1-3 bits of information. It appears that the information conveyed is primarily about the timing of abrupt variations in the stimulus.

Gauging efficiency of neuronal representations

The high information rates in most of the tested neural systems could have resulted solely from the fact that transmission speed was being maximized. The other factor that may have contributed to high information rates is use of an efficient code to represent information (see Box II; Rieke et al., 1993;1997). Indeed, the coding efficiencies as estimated from information rates obtained by reconstruction method were rather high for all tested systems (11-60%), meaning that a large percent of variability in the spike trains was caused by stimulus rather than noise. This was also true for neocortical neurons (up to 45%), despite their comparatively high trial-to-trial variability.

The reconstruction method discussed above addresses information about a specific stimulus parameter. Recently a number of researchers have developed algorithms for calculating the “total” information available in a spike train (DeWeese, 1996; Stevens and Zador, 1996; Strong et al., 1998). This “direct” method has two advantages over the reconstruction method: (1) This method does not suffer from information loss that might be caused by an imperfect reconstruction algorithm and is capable of providing a rather accurate estimate of information rate (provided that there is enough data). (2) It makes no assumptions about what stimulus parameter is encoded in the spike train. When applied to the fly (de Ruyter van Steveninck et al., 1997) and monkey data (Buračas et al., 1998) these two properties conspired to yield information rates that substantially exceeded the rates obtained using the reconstruction method (see Table I).

It seems unlikely that the two fold increase in information rate (Buračas et al., 1998) reflects only the recovery of information lost by imperfect stimulus reconstruction, since the stimulus employed was simple (could take on only one of two motion values) and readily reconstructed. A more likely possibility is that the two-fold improvement in the information rate was at least partially caused by the fact that the activity of these neurons conveys information not only about direction of motion but also other stimulus properties, such as spatial phase.

To quantify the degree to which a neuronal response captures a particular stimulus parameter, we introduce a metric for representational specificity S_R , which is a ratio of the information about the stimulus parameter in question (obtained by reconstruction method) relative to total information (obtained by the direct method; see Box II). According to this metric, at least 45% of the total information conveyed by spike trains recorded from individual MT neurons reflected stimulus direction. The lower bound of the representational specificity was even higher in the fly H1 neuron ($\geq 80\%$), which is understandable in view of the fact that, in contrast to area MT, the fly has very few neurons to estimate direction of motion.

Progression of information abstraction in sensory systems

These newly developed information theoretic methods allow one to precisely quantify the degree to which neuronal representations of sensory information become more abstract – by selective loss of irrelevant information – as one proceeds from periphery to central structures. This approach has been taken in recent studies of the electric fish *Eigenmannia* (Gabbiani et al., 1996). Peripheral electroreceptors, which sense the structure of the environment reflected in the electric field, were found able to convey large quantities of “raw” information (in terms of signal-to-noise ratio) about the rapid variation of the electric field. By contrast, pyramidal neurons of the first electroreceptive central nucleus (lateral-line lobe) were able to discriminate high-level features with high efficiency. According to the thesis presented, if explicitly computed, representational specificity (S_R) should grow for abstract, and not for low-level features as more and more low-level information is discarded and abstract features are computed.

Mechanisms supporting high information rates

The high information rates observed in a wide spectrum of species testify that encoding of sensory information may be very rapid and is characterized by a fast modulation of firing rate. König *et al.* (1996) argued, based on computer simulations, that this fast modulation is possible if stimuli cause synchronous synaptic inputs and the *modus operandi* of the neurons involved is coincidence detection. This hypothesis is supported by recent experiments that used synthetic synaptic input currents injected into cortical sensory neurons in brain slices (Stevens and Zador, 1998; c.f. Shadlen and Newsome, 1998). This synchronous mode of signal encoding might explain high timing precision observed in primate area MT (Bair and Koch, 1996; Buračas et al., 1998), which is as far as five synapses away from photoreceptors. Indeed, local field potentials (LFPs), which index local population activity, appear to covary with single unit responses in area MT, in a fashion consistent with the synchronous stimulus-locked activity needed to propagate precise timing information to downstream neurons (Buračas et al., 1998).

Open questions

A parallel line of research on temporal aspects of neuronal coding has explored the hypothesis that static stimuli may be encoded in the temporal distribution of neuronal responses (e.g. Richmond et al., 1987; Optican & Richmond, 1987). Recently McClurkin and colleagues (e.g. McClurkin et al., 1996) reported that different temporal neuronal response waveforms are used to encode color and shape in areas V1, V2, and V4. Indeed, afferents to these areas possess a wide range of response latencies, which could in principle

support the encoding of static features in time (Golomb et al., 1994). In view of the data reviewed herein, however, it will be important to determine whether this hypothetical temporal encoding remains invariant for time-varying stimuli.

Most of the reviewed work addressed the information conveyed by single neurons. How will information rates be affected by pooling responses from many neurons? The results of one study that addressed this question for retinal ganglion cells (Warland et al., 1998; see also Meister et al., 1995) suggest that response pooling may have nontrivial effects: As the pool of neurons increases, the contribution of each neuron to the stimulus reconstruction changes in a complicated way. Furthermore, response correlations between neurons (Oram et al., 1998), counter to conventional wisdom, might increase signal-to-noise ratio and, therefore, information rates.

Most importantly, the reviewed measurements of information rates only address how much of information is available at a given processing level. It remains to be seen how this information about rapidly varying stimuli is utilized and how it relates to phenomenal perceptual experience. Thus while many important questions confront us still, the approach reviewed herein frames the discourse and provides a novel means for investigation of sensory processing under conditions that approximate the richness of normal sensory experience.

Acknowledgements

We are grateful to D. Chklovskii, M. DeWeese, A. Messinger, O. Rukšėnas, G. Stoner, A. Thiele, and A. Zador for useful comments on the manuscript, and S. Janušonis and D. Kirvelis for stimulating discussions. TDA is an Investigator of the Howard Hughes Medical Institute. Correspondence should be addressed to GTB (e-mail: giedrius@salk.edu).

References

- Adrian, ED (1926) *The Basis of Sensation: The Action of the Sense Organs*. W.W.Norton, New York.
- Albright, TD (1992) Form-cue invariant motion processing in primate visual cortex. *Science*, **255**, 1141-1143.
- Bair, W, Cavanaugh, JR, & Movshon, JA (1997) Reconstructing stimulus velocity from neuronal responses in area MT. In: *Advances in Information Processing Systems, 9*. Proceedings of Conference on Advances in Neural Information Processing Systems, Denver, CO, USA, 1-4 Dec. 1996. Eds. MIT Press, 34-40.
- Bair W, & Koch, C (1996) Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comp.*, **8**, 44-66.
- Berry, MJ, and Meister, M (1998) Refractoriness and neural precision. *J. Neurosci.*, **18**, 2200-2211.
- Berry, MJ, Warland, DK, & Meister, M (1997) The structure and precision of retinal spike trains. *Proc. Natl. Acad. Sci. USA*, **94**, 5411-5416.
- Bialek, W, and Rieke, F (1992) Reliability and information transmission in spiking neurons. *TINS*, **15**, 428-433.
- Bialek, W, Rieke, F, de Ruyter van Steveninck, RR, and Warland, D (1991) Reading a neural code. *Science*, **252**, 1854-1857.

- Blakemore and Movshon (1995) In: Ed. Gazzaniga, *Cognitive Neuroscience*, MIT Press, 235-236.
- Buračas, GT, Zador, A, DeWeese, M, & Albright, TD (1996) Measurements of information rates in monkey MT neurons in response to time-varying stimuli. *Soc. Neurosci. Abs.*, **22-1**, 717.
- Buračas, GT, Zador, A, DeWeese, M, & Albright, TD (1997). Synchronous locking of area MT neurons to rapidly changing stimuli. *Soc. Neurosci. Abs.*, **23-1**, 457.
- Cover, TM, & Thomas, JA (1991) *Elements of information theory*. New York : Wiley.
- DeWeese, M (1996) Optimization principles for the neural code. In: *Advances in Neural Information Processing*, **8**. Proceedings of Conference on Advances in Neural Information Processing Systems, Denver, CO, USA, 27-30 Nov. 1995. Eds.: Touretzky, D.S., Mozer, M.C. & Hasselmo, M.E. MIT Press, 281-7.
- FitzHugh, R (1957) The statistical detection of threshold signals in the retina. *J. Gen. Physiol.*, **40**, 925-948.
- Gabbiani, F, Metzner, W, Wessel, R, & Koch, C (1996) From stimulus encoding to feature extraction in weakly electric fish. *Nature* **384**, 564-567.
- Golomb, D, Kleinfeld, D, Reid RC, Shapley RM (1994) On temporal codes and the spatiotemporal response of neurons in the lateral geniculate nucleus. *J. Neurophysiol.*, **72**, 2990-3003.
- Newsome, WT, Britten, KH, and Movshon, JA (1989) Neuronal correlates of a perceptual decision. *Nature* **341**, 52-54.
- McClurkin JW, Zarbock JW, & Optican LW. (1996) Primate striate and prestriate cortical neurons during discrimination .2. Separable temporal codes for color and pattern. *J. Neurophysiol.*, **75**, 496-507.
- Meister, M, Lagnado, L, Baylor, DA (1995) Concerted signaling by retinal ganglion cells. *Science*, **270**, 1207-1210.
- Optican, LM, & Richmond, BJ (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J. Neurophysiol.*, **57**, 162-178.
- Oram MW, Foldiak, P, Perrett, and Sengpiel, F (1998) The “ideal homunculus”: decoding neural population signals. *TINS*, **21**, 259-265.
- Reich, DS, Victor JD, Knight BW, Ozaki T, & Kaplan E (1997) Response variability and timing precision of neuronal spike trains in vivo. *J. Neurophysiol.*, **77**, 2836-41.
- Richmond, BJ, and Optican, LM (1990) Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. II. Information transmissions. *J. Neurophysiol.*, **64**, 370-380.
- Rieke, F, Warland, D, and Bialek, W (1993) Coding efficiency and information rates in sensory neurons. *Europhys. Lett.*, **22**, 151-156.
- Rieke, F, Warland, D, de Ruyter van Steveninck, RR, Bialek, W (1997) *Spikes*. A Bradford Book. MIT Press.
- Roddey, JC, & Jacobs, GA (1996) Information theoretic analysis of dynamical encoding by filiform mechanoreceptors in the cricket cercal system. *J. Neurophysiol.*, **75**, 1365-1376.
- Rolls, ET, Treves, A, Robertson, RG, Georges-Francois, and P, Panzeri, S (1998) Information about spatial view in an ensemble of primate hippocampal cells. *J. Neurophys.*, **79**, 1797-1813.

- Rolls, ET, Treves, A, Tovee, MJ, and Panzeri, S (1997) Information in the neuronal representation of individual stimuli in the primate temporal visual cortex. *J. Comp. Neurosci.*, **4**, 309-333.
- de Ruyter van Steveninck, RR, Lewen, GD, Strong, SP, Koberle, R, Bialek, W (1997) Reproducibility and variability in neural spike trains. *Science*, **21**, 1805-8.
- de Ruyter van Steveninck, RR, & Bialek, W (1988) Real-time performance of a movement sensitive neuron in the blowfly visual system: Coding and information transfer in short spike sequences. *Proc. R. Soc. London Ser. B*, **234**, 379-414.
- Shadlen, MN, Newsome, WT (1998) The variable discharge of cortical neurons: Implications for connectivity, computations, and information coding. *J. Neurosci.*, **18**, 3870-3896.
- Shannon, CE (1948) A mathematical theory of communication. *Bell Sys. Tech. J.* **27**, 379-423.
- Softky W (1996) Fine analog coding minimizes information transmission. *Neural Networks*, **9**, 15-24.
- Stevens, CF & Zador, AM (1998) Input synchrony and the irregular firing of cortical neurons. *Nature Neurosci.*, July, in press.
- Stevens, CF & Zador, AM (1996) Information through a Spiking Neuron. In: *Advances in Neural Information Processing Systems*. Proceedings of Conference on Advances in Neural Information Processing Systems, Denver, CO, USA, 27-30 Nov. 1995. D.S. Touretzky, M.C. Mozer and M.E. Hasselmo, eds., MIT Press **8**:75-81.
- Strong, SP, Koberle, RR, de Ruyter van Steveninck, RR, & Bialek, W (1998) Entropy and information in neural spike trains, *Phys. Rev. Lett.*, **80**, 197.
- Tolhurst, DJ, Movshon, JA, & Dean, AF (1986) The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* **23**, 775-785.
- Tovee, MJ, Rolls, ET, Treves, A, & Bellis, RP (1993) Information encoding and the responses of single neurons in the primate temporal visual cortex. *J. Neurophysiol.* **70**, 640-654.
- Warland, DK, Reinagel, P, and Meister, M (1998) Decoding visual information from a population of retinal ganglion cells. *J. Neurophysiol.*, **78**, 2336.
- Zador A (1998) Impact of synaptic unreliability on the information transmitted by spiking neurons. *J. Neurophysiol.*, **79**, 1219-1229.

Box 1. Prolegomena of Information Theory.

Mutual information can be defined intuitively by the following thought experiment. Imagine an observer relying on responses of two perfect neurons, the visual neuron being capable of perfect discrimination between red and green, and the tactile capable of perfect discrimination between apple and cherry shapes. Our stimulus set consists of four possible combinations of fruit and color (Fig. 1a). If all stimuli have equal probability of appearance ($=1/4$), then this stimulus set offers $H(\text{stimulus})=H(\text{color,shape})=\log_2 4=2$ bits of potential information, or in Shannon's terms – uncertainty (also called “entropy”). If the observer obtains only a response r_{shape} of the “shape neuron” to a randomly drawn stimulus, all uncertainty associated with shape is resolved, but uncertainty about color remains: $H(\text{color} | r_{\text{shape}})=\log_2 2=1$ bit. The gained total mutual information is $I(\text{stimulus}; r_{\text{shape}})=H(\text{stimulus})-H(\text{color} | r_{\text{shape}})=1$. However, if the observer obtains responses of both neurons, the stimulus can be determined exactly – the remaining uncertainty is zero – and we gain two bits of information. Hence, the mutual information is equal to the amount by which uncertainty about the stimulus is reduced after observing a response (Shannon, 1949).

In general, we can write the mutual information as a function of both the conditional and marginal stimulus and response probability distributions:

$$I(r; x) = H(x) - H(x | r) = -\sum_x p(x) \log_2 p(x) - \sum_i p(r_i) \sum_x p(x | r_i) \log_2 p(x | r_i).$$

Here $p(x)$ is the *a priori* distribution of the stimulus parameter x , and $p(x|r)$ is the *a posteriori* conditional distribution of the parameter x given response r . This equation implies that in order to evaluate $H(x|r)$, one first has to convert back from the neuronal response distribution, to stimulus probability distribution by means of Bayes rule (see Fig 1b):

$$p(x | r) = c \cdot p(r | x) p(x),$$

where $c=1/p(r)$. Estimation of the above equation is a simple task for constant-stimulus sets since their size only rarely exceeds 2^7 stimuli (for an equiprobable stimulus set of this size the uncertainty is 7 bits). By contrast, time varying stimuli may offer uncertainty rates of 100 bits/sec, which implies that the stimulus ensemble may be composed of $2^{100} \approx 10^{30}$ different stimuli of one second duration! These rich stimulus ensembles are much closer to the real world stimuli, but they require a special approach, such as optimal linear reconstruction (“reconstruction method”), which was developed by Bialek and colleagues (Bialek and Rieke, 1991; Rieke et al., 1997; for a nonlinear version see Buračas et al., 1998). This method can be used to estimate a lower bound on information about the stimulus parameter x .

Recently a number of researchers came up with an alternative method for estimation of the information rates, which is based on the fact that mutual information is symmetric with respect to stimulus and response uncertainties (Fig.1c):

$$I(r; s) = H(s) - H(s | r) = H(r) - H(r | s) = H(r) - \sum_i p(s_i) H(r | s = s_i).$$

The rightmost expression suggests that one can evaluate information rates by using two types of stimuli: (1) For the evaluation of the first term – total spike train uncertainty – one needs a large range of stimuli drawn randomly from a given stimulus ensemble, because these stimuli have to test the limits of response variability. (2) The second term

reflects the variability of the response when the stimulus is fixed, and can be evaluated from responses to a typical stimulus instance repeated many times (for details see Strong et al., 1998; Buračas et al., 1998). Since this method ("direct method") is not limited by the quality of stimulus reconstruction, it can yield, given sufficient amounts of data, an accurate estimate of information that is associated with the stimulus ensemble. In addition, since this method does not use information about stimulus parameters, it evaluates the total information that a neuron carries about all aspects of the stimulus.

Box 2. Information transfer vs. computation

It is common to hear in neuroscientific gatherings that information theory, which was designed for communication problems, is irrelevant for neuroscience, since the brain is not designed to simply transfer information, but rather to perform some sort of computations that discard irrelevant information. Every computation, however, including neural, uses some information that is transferred to the site of computation. In addition, information theory can be used to evaluate the efficiency of neuronal computations. For example, a computational system that extracts a parameter x (say direction of motion) from the raw data $s(x)$, can be evaluated in terms of its coding efficiency E_c , which is a ratio of information about the parameter x , $I(x;r)$, and the total uncertainty in the neuronal response $H(r)$: $E_c = I(x;r)/H(r)$. Here r is the signal produced by the computational system (e.g. Rieke et al., 1997; Warland et al., 1998). In addition, we propose that the efficiency with which a certain parameter x is computed from the stimulus $s(x)$ can be measured by representational specificity S_x in the following way. If we calculate the information about parameter x , $I(x;r)$ (e.g. by means of the reconstruction method), and then the total information (using the direct method) about the stimulus $I(s;r)$ then the representational specificity for the parameter x is: $S_x = I(x;r)/I(s;r)$ (Fig 3). Note, that since the reconstruction method estimates the lower bound for information, when using this method we only obtain a lower bound for S_x . This quantity may also be used in evaluating invariance of neuronal representations: When the stimulus is allowed to vary along an additional dimension y , then $S_x = I[x;r]/I[s(x,y);r]$ can be used to quantify the degree to which the representation of the parameter x is invariant with respect to parameter y .

Table I. Information rates and coding efficiency in a number of species

Neural system and species	Information rate (bits/spike)	Information rate (bits/sec)	Coding efficiency ¹	Best timing precision and variance/mean (Fano factor)	Reference
Constant stimulus²					
Cat retinal ganglion cells	~0.04-0.1	0.4 - 0.8 ³	–	–	FitzHugh, 1957
Primary visual cortex (V1) of rhesus monkeys	0.62	0.62 ⁴			Richmond, and Optican, 1990
Middle Temporal (MT) area of rhesus monkeys	0.025	0.89 ± 0.29	–	~1.3	Buračas et al., 1998
Inferior Temporal (IT) area of rhesus monkeys	~0.13	0.9 ⁴	–	–	Optican and Richmond, 1987; Rolls et al., 1997
Hippocampus of rhesus monkeys	~0.18 ⁵	0.32 (max = 1.2) ⁶	–	–	Rolls et al., 1998
Variable stimulus – Reconstruction method					
H1 motion-sensitive neuron of a fly	0.75	64	30%		Bialek et al., 1991
Frog auditory afferents	0.66	23	11%	~2 msec	Rieke et al., 1993
Vibratory receptors of the bullfrog sacculus	2.6	155	50-60%	–	Rieke et al., 1993
Cricket mechanoreceptors	0.6-3.2	75-294	50-60%	~0.4 msec	Rieke et al., 1993; Roddey and Jacobs, 1996
Salamander retinal ganglion cells	1.9	3.7 (up to 10 for a population of >10 cells)	26% (>79% for >10 cells)	–	Warland et al., 1998
Salamander retinal ganglion cells	3.6	–	–	<5 msec; 0.34	Berry et al., 1997
Area MT of anesthetized rhesus monkeys	~0.65	6.7 (max = 12.3)	–	2-4 msec	Bair et al., 1997; Bair and Koch, 1996
Area MT of alert rhesus monkeys	0.6	5.5	<30%	–	Buračas et al., 1998
Variable stimulus – Direct method					
H1 motion-sensitive neuron of a fly	2.43	80	50%	1.5-3 msec; <0.1	de Ruyter van Steveninck et al., 1997
Area MT of alert rhesus monkeys	1.5	12 (max=29)	Up to 45%	<2 msec; ~1.4	Buračas et al., 1998

¹ Evaluated for the bin size used in estimating the information rate.

² Information rates are based on spike counts in time windows of duration ≥ 400 msec.

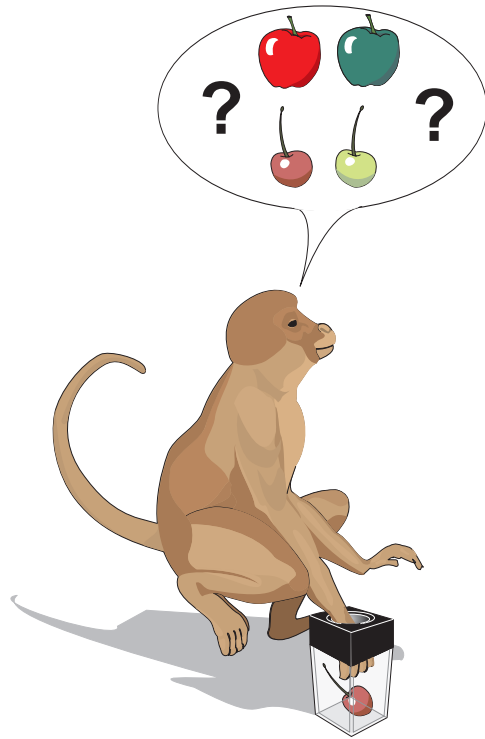
³ The information rate is given in bits/stimulus. The stimulus was 5 msec flash, but the integration time was a few hundred msec.

⁴ Estimated for the stimulus presentation period of 400 msec and the first three principal components of the average firing rate.

⁵ Obtained by dividing 1.2 by the reported firing rate of 6.5 spikes in 0.5 sec (Rolls et al., 1998).

⁶ Estimated in a period of 500 msec.

a) $H(x,y) = -\log 1/n = 2 \text{ bits}$

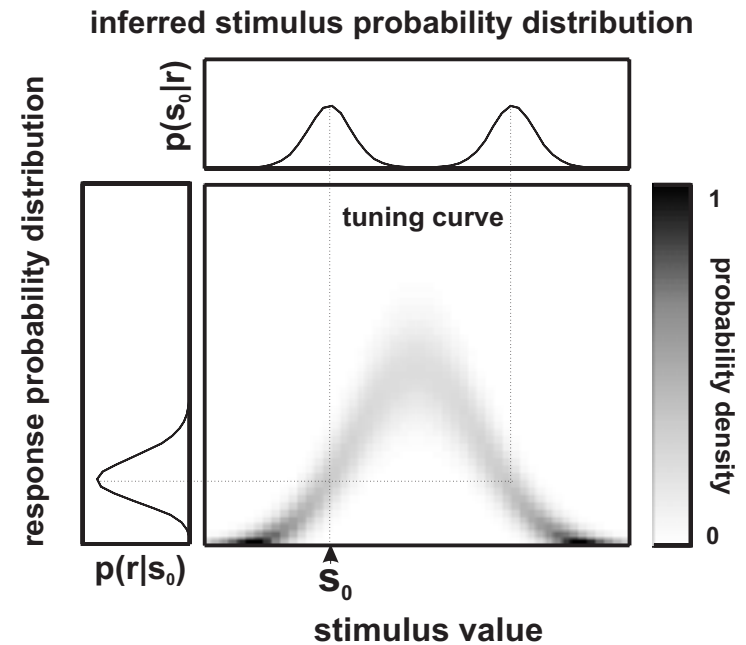


$H(x|r)=0 \text{ bits}$

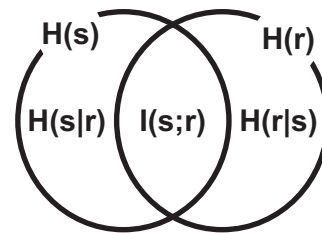
$H(y|r)=1 \text{ bit}$

$I(\{x,y\};r)=H(x,y)-H(\{x,y\}|r)=1 \text{ bit}$

b)



c)



$$I(s;r) = H(r) - H(r|s)$$

$$= H(s) - H(s|r)$$

Figure 1. Mutual information measures the reduction of uncertainty after observing a signal. (a) The stimulus set consists of four different stimuli. Each stimulus appears with equal probability of $1/4$. Hence, the initial stimulus uncertainty $H(x,y)$ is 2 bits. The subject possesses neurons that convey the signal r – about the color (x) and about the shape (y) – which permits a perfect discrimination of the feature. Bottom: after observing (feeling) the stimulus shape the observer is left with uncertainty about the unobserved feature. The gained information $I(s;r)$ is the difference between the initial uncertainty, and uncertainty after the signal was observed. Observing responses of both neurons conveys 2 bits about the stimulus s . This is equivalent to a perfect discrimination among all $2^{I(s;r)} = 4$ stimuli. (b) Calculation of mutual information from stimulus uncertainties involves reconstruction of the stimulus probability distribution. x_0 is the parameter of a presented stimulus. The probability distribution of neuronal responses associated with this stimulus $p(r|x_0)$ is plotted on the left. It can be obtained from the joint stimulus and response probability $p(x,r)$ depicted in the center. Top: The stimulus probability distribution $p(x_0|r)$ inferred from the response to stimulus x_0 . (c) Mutual information is symmetric with respect to stimulus and response uncertainties. The circles are diagrams depicting uncertainty of the stimulus and response sets. The area of their intersection corresponds to the mutual information. Hence, the mutual information can be estimated from the total response uncertainty (i.e. range of possible neuronal responses) and the uncertainty that remains after presenting a specific stimulus (noise).

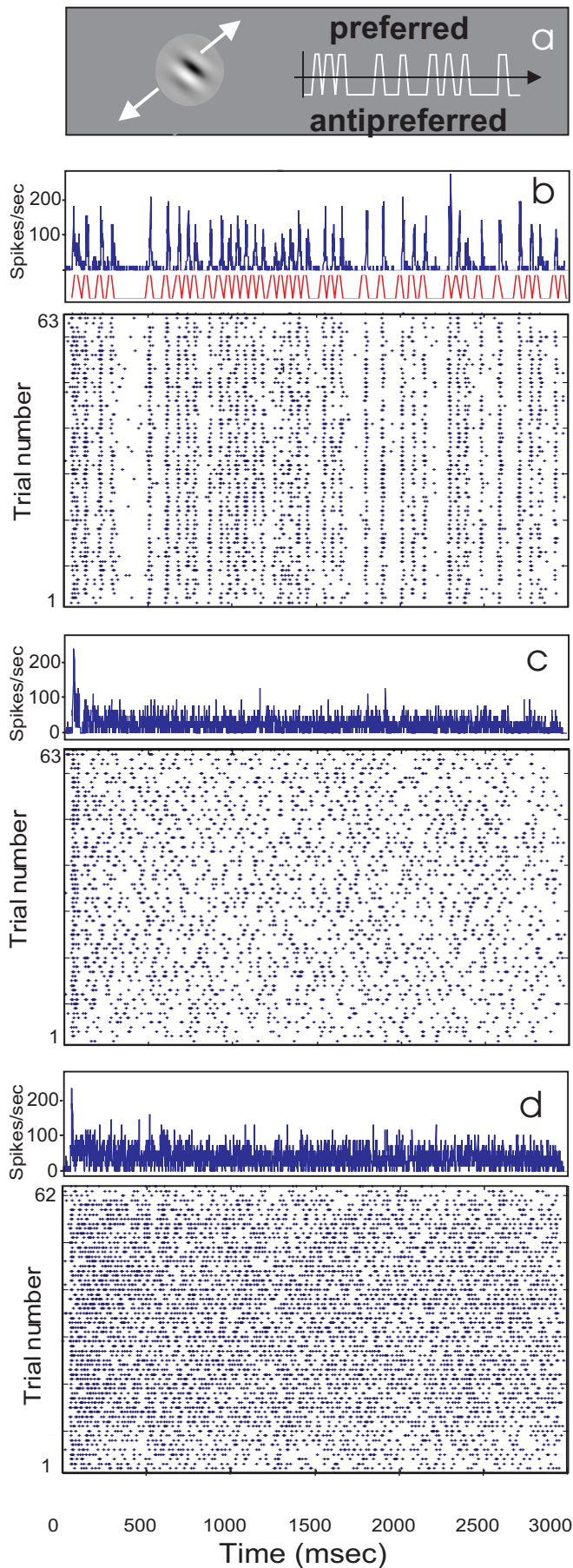
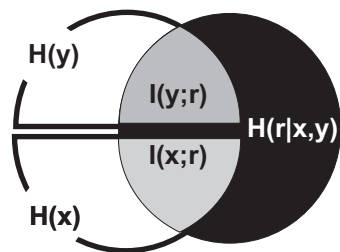


Figure 2. The variability of the response may be due to either noise or the signal. (a) rapidly varying stimulus that induced maximum information rates in area MT neurons (Bura et al., 1998). This stimulus was composed of a moving sinusoidal grating windowed with a stationary gaussian window. The direction of motion reversed repeatedly after short unpredictable periods of time. (b) typical response of a MT neuron to one such stimulus repeated many times. Top: a 1 msec bin histogram. Below is shown a raster of responses to individual stimulus presentations. Neuronal responses to direction reversals was timed with high precision. (c) responses to the stimuli drawn randomly from the same stimulus set. Each spike carries over a bit of information about stimulus timing, regardless of the fact that the raster looks as noisy as that produced in response to stimulus moving with constant speed in the preferred direction (d). In the latter raster indeed most of the variability reflects noise which is not correlated with the signal.



$$I(\{x,y\};r) = I(x;r) + I(y;r)$$

$$R(x) = I(x;r) / I(\{x,y\};r)$$

Figure 1. The quality with which a neuron represents a particular parameter can be estimated by calculating representational specificity $R(x)$, which is a fraction of total information dedicated to a specific parameter x . The split circle on the left depicts stimulus entropies associated with two independently varying parameters x and y . The uncertainty $H(r|x,y)$, shaded in black, measures variability in response r , which is not correlated with stimulus parameters. Information about x and y (the two gray regions) in this case are independent, and hence, additive.