

Statistical and information properties of head direction cells

JOHN C. BAIRD, JEFFREY S. TAUBE, and DAMEN V. PETERSON
Dartmouth College, Hanover, New Hampshire

The human channel capacity for identifying sensory stimuli is compared with channel capacities based on neurophysiological findings. Studies have shown that cells in the postsubiculum (PoS) and the anterior dorsal thalamus (ADN) of the rat discharge as a function of the animal's head direction in the horizontal plane. We compute the statistical properties of the firing rates of head direction (HD) cells and the potential amount of information transmitted by these cells according to two theoretical models. The *cell response model* for single cells indicates that information transmitted is much less than 0.5 bits. The *population response model* developed for cell ensembles generates values in the range of 1–3.2 bits, suggesting that a cell population can distinguish between two and nine head directions, depending on the value used for the standard deviation of directions over which a cell fires. These values are similar to those found in human psychophysical studies of the channel capacity for unidimensional sensory attributes.

George Miller (1956) summarized a large array of psychophysical studies showing that the human channel capacity for unidimensional stimulus attributes ranges between 2 and 3 bits of information. Recent efforts have attempted to determine the neural basis of this limit (e.g., Norwich, 1993), and empirical channel capacities have been measured in a number of different species (reviewed by Rieke, Warland, de Ruyter van Stevenink, & Bialek, 1997). However, the possible connections between the animal and human findings regarding channel capacity typically are not the main focus of either psychophysical or neurophysiological research. In interpreting either source of results, a distinction should be made between information theory as a mathematical tool and the substantive assumptions guiding its use in the analysis of empirical data. Depending on the underlying assumptions, the mathematics may be applied in alternative ways to the same data set, and hence, different answers can arise from different assumptions about how the formal structure of the theory is mapped onto psychophysical or neurophysiological parameters (e.g., Baird, 1997; Norwich, 1993; Rieke et al., 1997; Skaggs, McNaughton, Gothard, & Markus, 1993). Most importantly, for the present paper, the response options for a single cell can be considered as either a binary event (*spike or no spike*), as a continuum

(or discrete set) of firing rates, or as a set of different types of spike trains. Attention must also be paid to the units for expressing the quantity of mutual information (information transmitted), which can be in bits/stimulus event, bits/sec, or bits/spike. The differences in measurement units may explain why the psychophysical and neurophysiological findings are so seldom compared.

In the present work, we introduce two models to calculate the amount of information transmitted by single neurons and by populations of neurons responsible for the behavioral ability of a rat to distinguish among different head directions. The dependent measures are expressed as information transmitted per behavioral head direction. Our long-term goal is to identify links between the animal and human domains in hopes of gaining some clues as to the neurophysiological basis of human channel capacities for sensory stimuli.

Spatial Cells

Previous studies have identified two types of allocentric spatial cells in the rat brain: place cells and head direction (HD) cells. Place cells discharge as a function of the animal's location in the environment and are found throughout the hippocampus (O'Keefe & Dostrovsky, 1971; for reviews, see Muller 1996; O'Mara, 1995). Each place cell encodes a different location within the environment, and the population of place cells is believed to represent a cognitive map for the animal (O'Keefe & Nadel, 1978). In contrast, HD cells discharge as a function of the rat's head direction in the horizontal plane and were initially identified in the postsubiculum (PoS; Taube, Muller, & Ranck, 1990a, 1990b; for review, see Taube, 1998). For example, an HD cell might fire whenever the animal pointed its head northeast. HD cell discharge is independent of the animal's location, trunk position, and on-going behavior in

This work was supported by NIMH Grants MH48924 and MH01286 to J.S.T. We thank Heather Burton, Paul Dudchenko, Edward Golob, Jeremy Goodridge, Robert Stackman, Matthew Tullman, David Volk, and Kimberly Worboys for contributing head-direction cells to the data pool. We also are grateful for constructive comments by Kenneth Norwich, Shuji Mori, Lawrence Ward, and Bruce Schneider. Correspondence concerning this article should be addressed to J. C. Baird, Psychological Applications, 456 Cornish Turnpike, Newport, NH 03773 (e-mail: john.c.baird@dartmouth.edu).

the environment. Thus, in situations where the head is pointing in one direction and the torso is pointing in another direction, cell firing will only depend on the direction the head is pointing. HD cell firing is largely movement independent; cells discharge whether the rat is moving or remains still, as long as the animal is pointing its head in the proper direction. Little, if any, adaptation of firing rate is present when the head remains fixed at the preferred direction. Response characteristics of HD cells are usually stable across recording sessions many days apart. In sum, HD cell firing most likely represents an animal's perceived orientation of its head with respect to its environment.

Each cell contains a single head direction where cell firing is maximal (peak firing rate); this head angle is referred to as the cell's preferred direction. Cell firing decreases as the head moves away from the cell's preferred direction, and the range of head angles where cell firing is above that of background level is approximately 90°. Individual cells have different preferred firing directions, and a population of these cells within a brain area have preferred directions that are uniformly distributed over 360°. During the past 5 years, HD cells have been identified in several areas of the brain, including the anterior dorsal nucleus of the anterior thalamus (ADN; Taube, 1995a), lateral dorsal thalamic nucleus (Mizumori & Williams, 1993), retrosplenial and medial prepirate cortices (Chen, Lin, Green, Barnes, & McNaughton, 1994), lateral mammillary nuclei (Stackman & Taube, 1998), and dorsal striatum (Wiener, 1993).

Previous studies have clearly demonstrated multisensory convergence onto HD cells in experiments that manipulated visual cues and *idiothetic* (vestibular, proprioceptive, motor efference copy) cues. The preferred direction of all HD cells is maintained for several minutes when the animal moves about, even in darkness, but can be simultaneously rotated about an earth vertical axis by appropriate manipulation of surrounding visual landmarks. For example, repositioning a large salient landmark mounted on the wall of a cylindrical test chamber by 90°, 180°, or 270°, produces a similar shift in the preferred direction of all HD cells monitored concurrently (Taube et al., 1990b). Landmark removal does *not* lead to a cessation of firing, but only to an equal angular shift in the preferred direction of all cells by a random amount (Taube et al., 1990b). This finding indicates that afferent input driving one HD cell similarly influences other HD cells and that HD cells within a particular brain area behave as a network and their preferred directions always remain at a fixed angle apart (in register) from one another. It is also clear that vestibular input onto HD cells is critical for the generation of the directional signal, as neurotoxic lesions of the vestibular apparatus abolishes the direction-specific firing of ADN HD cells (Stackman & Taube, 1997).

In order to understand how animals use HD cells for navigation, HD cell activity has been monitored as an animal entered either a novel environment or an environ-

ment in which a conflict situation was set up in relation to the established orientation cues (Taube & Burton, 1995). For cell firing to continue when the animal moves into a novel environment, the animal must use a path integration approach and rely on idiothetic cues, since there are no familiar azimuth cues for orientation. When animals locomoted into a novel environment, HD cells continued to discharge in a similar direction as they did in a familiar environment. In contrast, cells from animals transported passively on a cart into the novel environment did *not* maintain the same preferred direction (Taube, Stackman, & Dudchenko, 1996), which again emphasizes the importance of motor and kinesthetic signals in continually updating the animal's perceived orientation. In situations where the spatial information from visual and idiothetic cues were in conflict, HD cells usually relied on the visual spatial information. Thus, HD cells integrate information from both landmark and path integration navigational systems and respond predictably when azimuth direction cues are ambiguous or conflict.

The properties of HD cells in the PoS and ADN have received the most detailed attention. Specifically, in our laboratory, the cell's firing rate is recorded while simultaneously monitoring the animal's head direction using a video tracking system. From these data, tuning curves are constructed that show a cell's firing rate as a function of head direction. Several parameters are extracted from these plots, including peak firing rate, directional firing range, background firing rate, and asymmetry (the extent to which cell discharge decreases equally in both directions when moving away from the cell's preferred firing direction; Taube et al., 1990a). These tuning curves show there is a graded continuum of peak firing rates across different HD cells, with some cells having low peak firing rates (~5 spikes/sec) and others having high peak firing rates (>100 spikes/sec); intermediate values are also present. The peak rate usually remains stable across different recording sessions and days. Each HD cell can be characterized by one of a number of descriptive measures. One of the purposes of the present paper is to describe in statistical terms how much information an HD cell conveys about the animal's head direction and to relate this measure to the statistical parameters of the tuning curves for individual cells.

Despite the efforts of neuroscientists (e.g., Barlow, 1995) to attribute all psychophysical judgments to the operation of a handful of neurons (or even a single neuron), it remains unclear how effective a single cell can be in discriminating among a set of stimulus features (e.g., head directions). Lately, neural explanations of perceptual and psychophysical results have turned to models based on the firing rates of cell populations, rather than on the rates of individual neurons (Arbib, Érdi, & Szentágothai, 1998; Baird, 1997). Of special relevance, Georgopoulos and colleagues (Georgopoulos, 1995; Georgopoulos, Caminiti, Kalaska, & Massey, 1983; Georgopoulos & Massey, 1988) have postulated that the summed activity across a population of cells (tuned to different directions

and weighted by firing rate) in the primate motor cortex more accurately reflects the animal's future directed movements than the activity observed for individual neurons. Population models have also described the ability of cell ensembles in the cricket cercal sensory system to code environmental directions based on air current direction (J. P. Miller, Jacobs, & Theunissen, 1991; Theunissen & Miller, 1991).

Analogous arguments can be drawn for place and HD cells. For example, it has long been known that place fields occupy a significant portion of the environment (>10%) and that different place cells contain overlapping fields within the same environment. These considerations along with empirical data from multiple simultaneous recordings led Wilson and McNaughton (1993) to propose that the activity across the population of hippocampal place cells represents the neurophysiological basis for the animal's perceived location (population place vector). Similarly, the activity across a population of HD cells in a given brain area may represent the neurophysiological basis for an animal's perceived head direction (population HD vector).

In sum, an ongoing concern is the extent to which an animal can distinguish between different directions (or places) on the basis of the discharge rates of single cells, as well as of cells functioning together in a population network. In attempting to clarify this issue, information theory has been applied to the activity of single cells, and equations have been derived to specify the amount of information transmitted by a single cell, as well as the amount transmitted by ensembles (Georgopoulos, 1995; Georgopoulos & Massey, 1988; Theunissen & Miller, 1991). Information theory also has been used to determine the information transmitted by neurons in the primary visual cortex and inferior temporal cortex (Gershon, Wiener, Latham, & Richmond, 1998; Optican & Richmond, 1987; Richmond & Optican, 1990; Tovée, Rolls, Treves, & Bellis, 1993).

Skaggs et al. (1993) used information theory to characterize how much information (bits/spike) is encoded by individual hippocampal place cells, and they interpreted this measure as an indicator of how well a single neuron can distinguish among locations in the environment. Investigators have employed this measure in several studies to compare properties across place cells in different brain areas (e.g., Jung & McNaughton, 1993; Poucet, Thinus-Blanc, & Muller, 1994; Taube, 1995b), as well as under different environmental conditions (e.g., Gothard, Skaggs, Moore, & McNaughton, 1996; Markus, Barnes, McNaughton, Gladden, & Skaggs, 1994). More recently, Taube and Muller (1998) reported on the information transmitted (bits/spike) by a HD cell concerning the animal's head direction.

General Procedure

In the context of our attempts to understand the function of HD cells and their relation to human psychophysics, we conducted analyses at two levels: behavioral

and physiological. At the behavioral level, we wished to determine how many directions a rat can identify without error, and therefore, information/stimulus was the most appropriate measure. On the other hand, it was equally important to obtain an indication of how information transmitted relates to physiological properties, such as mean firing rate of a single neuron. For this purpose, a rate measure was appropriate. A viable theory of how animals code head direction will probably consider the relevant behavioral data (information/stimulus) and show how it arises from physiological mechanisms whose transmission properties are assessed in rate units (information/sec). Although a measure expressed in units of bits/spike may also be useful in neurophysiological modeling (Rieke et al., 1997), in our opinion *it is an inappropriate indicator of the relative ability of different cells to distinguish among a set of head directions.*

In light of previous research, the purpose of the present work was fourfold. (1) Describe the statistical properties of HD cells in terms that permit the fitting of analytic distributions to cell populations. (2) Calculate the amount of information transmitted (bits/direction) by cells acting alone and together. (3) Link information measures with other statistical properties of HD cells. (4) Compare channel capacities obtained for head direction cells with those obtained in human sensory psychophysics.

METHOD

Cells were obtained from the data pools of several previous studies conducted on HD cells in the PoS and ADN of female Long-Evans rats. Only HD cells that were well isolated from background activity and from other possible HD cells recorded on the same electrode wire simultaneously were used in the final analyses. The methods for training animals, electrode construction, surgical procedures, recording HD cells, and data analysis are described in previous papers and are only summarized briefly here (Taube, 1995a, 1995b; Taube et al., 1990a).

All cells were recorded from food-restricted animals as they foraged for food pellets (20 mg) thrown randomly into a cylindrical apparatus that was placed on end. The cylinder was 76 cm in diameter, 51 cm high, colored gray, and was surrounded by a black floor-to-ceiling curtain that was positioned about 0.7 m from the cylinder wall. Four overhead lights arranged uniformly at the ceiling provided a low level of illumination. The primary orienting cue was a white sheet of cardboard attached to the inside of the cylinder wall that occupied approximately 100° of arc. A vertically oriented color video camera with a zoom lens was suspended over the cylinder 206 cm above the floor. Well-trained rats spent most of their time moving about the apparatus searching and consuming the food pellets. They usually visited all parts of the cylinder floor and spend approximately equal amounts of time at each head direction.

HD cells were monitored in 8-min recording sessions as the animal's head direction was tracked at 60 Hz using two differently colored light-emitting diodes (LEDs), spaced about 10 cm apart along the rostral caudal axis and secured to the animal's head. The resolution of the tracking system was approximately 2°. For each sample, the animal's head direction in the horizontal plane was calculated from the relative positions of the two LEDs and sorted into one of 60 6° bins along with the number of spike discharges. The total time and the number of spikes in each bin for a particular session were summed from the collected samples. The cell's firing rate was calculated as the total number of spikes in each head-direction

bin divided by the time spent at that head direction. The direction at which maximum firing occurred is referred to as the cell's preferred firing direction, and the rate of maximum firing is referred to as the cell's peak firing rate.

Results

Preliminary analysis indicated no important differences between the statistical properties of cells in the PoS and in the ADN, and hence, the data from these two locations are combined in the following analyses (with the exception of Figure 4). The initial pool contained 191 cells, 107 from the ADN, and 84 from the PoS. This pool was further reduced in the manner described below.

Statistical properties. Our initial analysis considers the tuning-curve characteristics of cells over changes in head direction. The firing rates for each cell were standardized so that the peak rate for a single direction was equal to the mean of a three-parameter Gaussian distribution fit to the firing rate as a function of head direction. This function is described by Equation 1:

$$f = Ke^{-(1/2)[(d - \mu) / \sigma]^2}, \tag{1}$$

where f is firing rate, d is head direction (1 to 60 bins), μ is the mean head direction, σ is the standard deviation, and K is a scale factor. The fits of Equation 1 were $r \geq .95$ for 174 out of 191 cells (91%). The Gaussian is an adequate model for this application because the random variable (angle) folds back on itself at substantial distances from the mean, well beyond the point where it could effect the fit.¹ The model provides excellent fits within its range of application, and is well suited to our purposes of obtaining measures of a cell's tuning characteristics over head direction.

Figure 1 shows results for three examples: one cell (1A) yielding an excellent fit of the Gaussian ($r = 1$); a second cell (1B) yielding a fit of $r = .95$; and a third cell (1C) that shows double peaks (suggesting two cells were recorded), leading to a relatively poor fit ($r = .90$). For the analyses described below, only those cells with a fit of $r \geq .95$ were included, though the pattern of results obtained for the full data set is not substantially different from those obtained for the reduced set. As would be expected, the variability among cells was greater in the full data set than in the reduced set. The 95% cutoff eliminated cells because of high background firing rates, considerable noise in the tuning function, or the existence of a multimodal distribution (indicating that several cells were being recorded on the same electrode).

Figure 2 shows a histogram of the standard deviations (SDs in degrees) of the Gaussian distributions for the reduced pool. The solid curve represents the best-fitting gamma distribution, used to summarize the trend in the data. The distribution of SDs is itself relatively normal (Gaussian), except that there is a somewhat longer right-hand tail, indicating a positive skew toward the larger SDs.

Figure 3 shows the distribution of cells as a function of peak firing rate. There is a preponderance of cells at the lower end of the scale, with a long tail stretching across

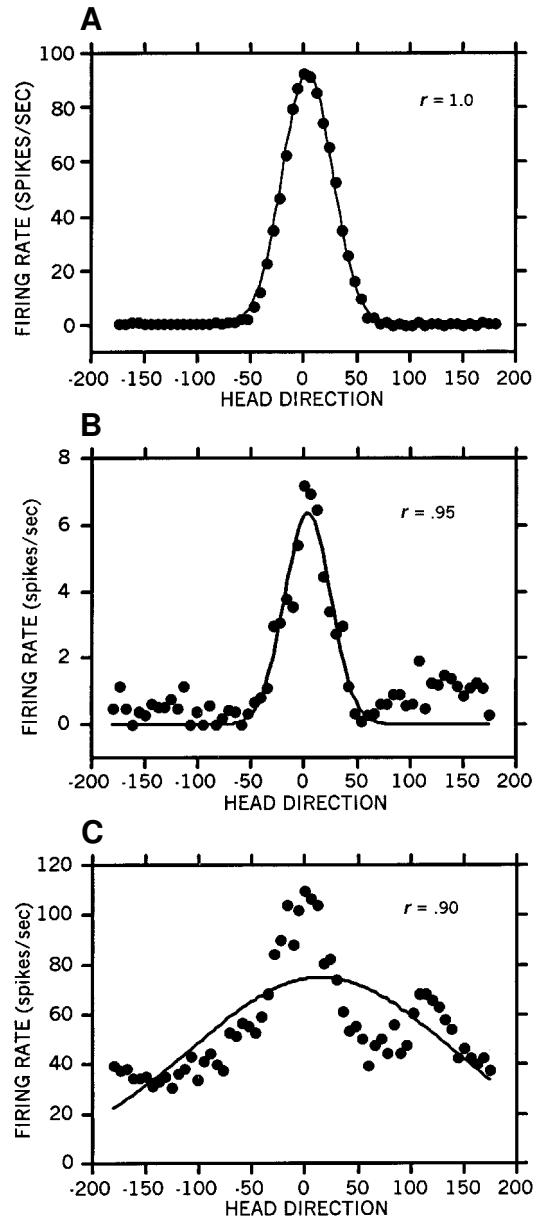


Figure 1. Firing rate as a function of head direction (arbitrary degrees) for three cells. (A) A single cell yielding an excellent fit ($r = 1$) of a Gaussian distribution (Equation 1). (B) A single cell yielding the least acceptable fit ($r = .95$) of a Gaussian in order to be included in further analysis. (C) A single cell yielding a relatively poor fit ($r = .90$) of the Gaussian. This cell was not included in further analysis.

higher values (positive skew). A gamma distribution was fit in order to emphasize this trend. Best-fitting parameter values are indicated on the graph. There is a high positive correlation between the peak firing rate and both the mean firing rate ($r = .90$) and the SD of the firing rate ($r = .99$), as well as between the mean firing rate and the SD of the firing rate ($r = .94$). In contrast, the correlation between mean firing rate and the SD of the Gaussian is

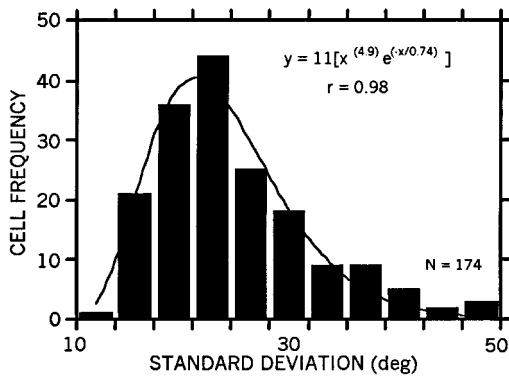


Figure 2. Histogram of cell frequency as a function of standard deviation of the Gaussian over head direction (Equation 1). The solid curve represents the best-fitting gamma function with parameters indicated on the graph.

.06. Such statistical relations for HD cells have not been reported before. This type of data may prove valuable in later developments of physiological models concerning the ability of animals to navigate through the environment.

Descriptive statistics are given in Table 1, together with a comparable summary for the *SD* of the Gaussian fits.

The distribution of cells throughout the range of 360° was assessed by multidimensional scaling (MDS). This analysis (Baird & Noma, 1978) was conducted separately for the two classes of cells: ADN and PoS. A matrix was constructed of $N = 90$ (ADN) or 84 (PoS) cells by $M = 60$ bins (head directions). A similarity measure between cells was taken as the correlation between firing rates across all 60 bins, representing equally spaced directions from 0° to 360° . This correlation matrix was converted into distance measures by subtracting 1.0 from each value and then submitting the matrix to MDS. The resulting Euclidean, two-dimensional solutions are shown in Figure 4. The locations of the points are roughly circular, both for cells from the ADN (Figure 4A) and from the PoS (Figure 4B). This result indicates that HD cells are equally represented in all directions. The finding of uniform angular representation agrees with previous work based on Chi-square analyses; the present Euclidean solution graphically depicts the same outcome.

Applications of Information Theory

Recent studies have applied information theory to the firing rates of single cells that code for place or head direction in freely moving rats (Gothard et al., 1996; Jung & McNaughton, 1993; Jung, Wiener, & McNaughton, 1994; Markus et al., 1994; Poucet et al., 1994; Skaggs et al., 1993; Taube & Muller, 1998). The main goal of these approaches was either to quantify the link between stimulus features in the physical environment and the firing rates of single cells or to compare the amount of information transfer across cells in different parts of the hippocampal formation or HD cell network. All these ap-

proaches used the measure of bits/spike as the dependent variable, and, as explained below, we have serious questions about whether this measure is appropriate for indicating the behavioral ability of the animal to identify head directions.

We propose here two new models of information transmission and provide formulas to compute a measure in bits/heading of a cell's ability to selectively code environmental directions and a rate measure in bits/sec that permits a direct comparison with the physiological properties of these cells. The "stimulus" in this formulation is best treated as an abstract concept, because, as noted in the introduction, HD cells respond to visual and vestibular cues (e.g., they still fire selectively when the animal is in the dark). The first model applies to single cells and is referred to as the *cell response model*. The second model applies to ensembles of cells and is referred to as the *population response model*.

Our theoretical strategy builds on the standard Shannon and Weaver (1963) analysis as applied to psychophysical data by experimental psychologists (for review, see Baird, 1997; Baird & Noma, 1978) and to physiological data by neuroscientists (for review, see Georgopoulos, 1995; Theunissen & Miller, 1991; Tovée et al., 1993). The essential equation for calculating the average uncertainty (information) ($U(A)$) over a set of A discrete alternatives is

$$U(A) = -\sum_{i=1}^N p(a_i) \log_2 p(a_i), \quad (2)$$

where $p(a_i)$ is the probability of occurrence of alternative i . All the information measures used to derive models in this paper are variations on Equation 2 (for details, see Baird & Noma, 1978; Cover & Thomas, 1991).

Cell Response Model

The information transmitted by single cells is computed by the cell response model as

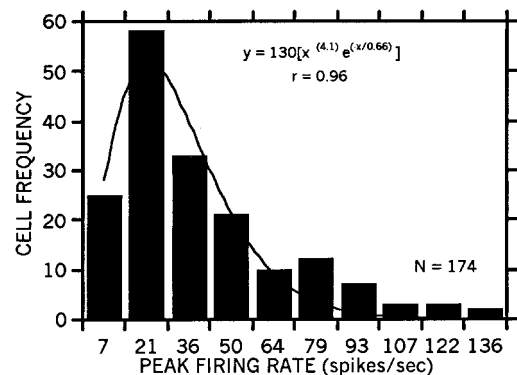


Figure 3. Histogram of cell frequency as a function of peak firing rate. The solid curve represents the best-fitting gamma function with the parameters indicated.

Table 1
Firing Rate and Head Direction Statistics

| | Peak | Mean* | SD** | SD*** |
|----------|-------|-------|------|-------|
| Minimum | 6.0 | 3.0 | 2.0 | 10.9 |
| Maximum | 143.0 | 47.0 | 47.0 | 51.5 |
| Mean | 39.5 | 12.2 | 12.2 | 24.0 |
| Median | 30.0 | 9.5 | 9.0 | 22.0 |
| Standard | 28.5 | 8.4 | 9.5 | 8.4 |
| Skewness | 1.3 | 1.6 | 1.4 | 1.1 |

Note—*Equivalent to the overall mean firing rate (direction independent). **Standard deviation of firing rate. ***Standard deviation of Gaussian (head direction in degrees).

$$I_t = \sum_i^N p(x_i) \sum_j^M p(y_j | x_i) \log_2 \frac{p(y_j | x_i)}{p(y_j)}. \quad (3)$$

The derivation of this equation is given in Appendix A (Equation A11 is the same as Equation 3), so here we will only summarize the main points. Define the environmental stimulus as a head direction (x_i) with a specific probability of occurrence. The average stimulus uncertainty over N headings is then given by Equation A1, which is obtained by substituting x for a in Equation 2. For purposes of the cell response model, we define the headings as equally likely; thus, Equation A1 reduces to $-\log_2(1/N)$. For our particular case there are 60 bins (head directions), yielding an average stimulus uncertainty of 5.9 bits.

Next, define an interval of time, Δt , during which a cell is in one of two response states: *spike* or *no spike*. A spike occurring anytime during Δt means the cell exhibits state *spike*. Each such time window can be treated as a single “trial” in the experiment. There are $T/\Delta t$ such intervals over some period T ($\Delta t \leq T$). We have seen cells that fire over 200 spikes/sec. An upper estimate of the maximum that might be achieved is 250 spikes/sec, a value also suggested by MacKay and McCulloch (1952) for sensory neurons in general. The physiological basis for this choice is as follows: The time course of a neural spike is between 1 and 2 msec, and the spike’s refractory period is similar in magnitude (Katz, 1966). Thus, the maximum number of spikes a neuron can deliver per second is between 250 and 500 spikes/sec, and therefore, $\Delta t = 2\text{--}4$ msec. We selected $\Delta t = 4$ msec (0.004 sec) because all our cells fire below 250 spikes/sec ($1/0.004$). Only a single spike can occur within the window Δt .

For each heading (x_i), we define the probability of a spike event (y_1) as $p(y_1 | x_i) = \lambda(\Delta t)$. For example, a cell with a firing rate of 50 spikes/sec has 50 of its Δt windows filled by a spike, and, hence, the probability of a spike event for this cell at this heading is $p(y_1 | x_i) = 50(0.004) = .20$. The probability of no spike for the same heading is then $p(y_2 | x_i) = 1 - p(y_1 | x_i) = .80$. The two joint probabilities, $p(x_i y_j)$, for a single heading (Equation A3) are $p(x_i) p(y_1 | x_i)$ and $p(x_i) p(y_2 | x_i)$. Substituting these values into Equation A3 for all possible headings yields the average joint uncertainty. The transmitted information is then given by Equation A7.

Appendix B presents the derivation of a single formula to compute information transmitted using the terms of the model. The final equation (Equation B6) gives the same answers as Equations A6, A7, and A8, but in a more compressed form, reproduced here as Equation 4.

$$I_t = \frac{1}{N} \sum_i^N \left[\lambda_i \Delta t \log_2 \frac{\lambda_i \Delta t}{\lambda \Delta t} + (1 - \lambda_i \Delta t) \log_2 \frac{(1 - \lambda_i \Delta t)}{(1 - \lambda \Delta t)} \right], \quad (4)$$

where N = number of discrete head directions (the number of bins in which head directions are sorted), Δt = time course of a single discrete event, λ_i = mean firing rate of the cell in the i th bin, and λ = overall firing rate of the cell for the entire recording session, expressed as

$$\lambda = \frac{1}{N} \sum_{i=1}^N \lambda_j.$$

In this model, an upper limit on information transmission is determined by the response uncertainty, which in our applications is always less than the stimulus (location or bin) uncertainty. Because a cell only has *on* and *off*

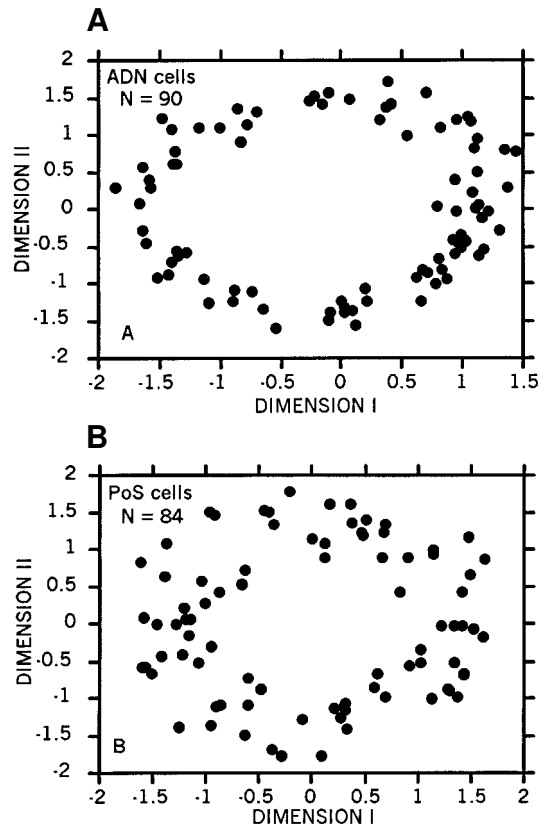


Figure 4. Two-dimensional depiction (based on multidimensional scaling) of the tuning of head-direction cells recorded in (A) the anterior dorsal thalamus (ADN), and in (B) the post-subiculum (PoS). The circular patterns suggest that cells from both locations are uniformly distributed from 0° to 360°.

states, the maximum average response uncertainty is 1 bit for a single time interval Δt . Therefore, according to the cell response model, the maximum information/heading that can be transmitted by a single cell is 1 bit. All empirical measures of information must be less than or equal to this value. The extent to which the empirical value approaches the theoretical maximum depends on the size of Δt and on the distribution of firing rates at each head direction.² On the other hand, a cell can transmit more than 1 bit over some longer time interval, and a rate measure per sec is obtained by multiplying the information/heading (I_i) by the number of time intervals ($T/\Delta t$).

Table 2 presents a statistical summary of the results. All information measures (left column) are well below the theoretical maximum of 1 bit; indeed, the values indicate that the ability of a single cell to absolutely discriminate among head directions is extremely limited. One of the reasons for these low values is that the maximum firing rate of 250 is considerably greater than the peak firing rates actually achieved by most of the cells. This means that there is an asymmetry between the states of *spike* and *no spike*, with the latter far outweighing the former. This asymmetry in turn leads to a low theoretical maximum information transmission for the cell. Given that animals are capable of making rather fine behavioral discriminations in head direction (on the order of 10°–20°; Olton, Collison, & Werz, 1977; Tolman, Ritchie, & Kalish, 1946), the low information content emphasizes the point that network properties of HD cells must ultimately play a central role in the discrimination of head direction. In this regard, however, it is important to realize that discrimination is not the same thing as absolute identification. We will return to this point at the end of the paper.

It is possible to compute a measure for the cell response model that takes into account the differences in peak firing rates among the cells. For example, one can compute the amount of information transmitted relative to the amount possible, given the constraints of the firing rate—a measure of cell efficiency. In our case,³ this amount is the information transmitted divided by the response uncertainty (multiplied by 100 to convert measures into percentages). Descriptive statistics for cell efficiency are presented in the second column of Table 2. The mean efficiency is 23% with an *SD* of 9%. Thus, although the information measures produced by the cell response model are small in absolute terms, the efficiency of the cells is about one-quarter of the maximum achievable.

Information Measures and Statistical Properties

In order to compare information measures with other physiological properties of neurons, we computed an information/sec value for each cell. This was obtained by multiplying the information/heading value by 250 (maximum rate). Figure 5 shows the relation between information transmitted in bits/sec and the mean firing rate in spikes/sec. The correlation is highly significant, and, since the mean firing rate is highly correlated with both peak

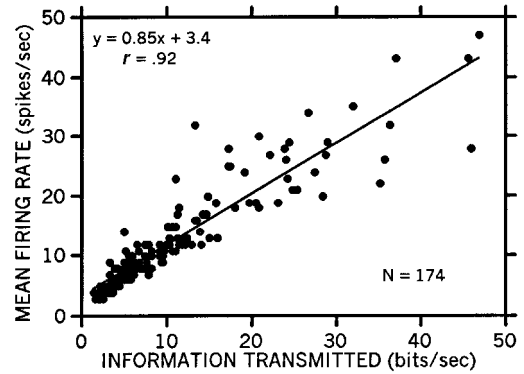


Figure 5. Mean firing rate (spikes/sec) as a function of information transmitted (bits/sec) according to the cell response model.

firing rate and standard deviation, the latter measures also are highly correlated with information transmitted.

Figure 6 shows a similar plot for the relation between information transmitted and the *SD* of the best-fitting Gaussian (Equation 1). The correlation is essentially zero ($r = .05$). Therefore, according to the cell response model, there is a strong link between information transmitted and a cell's peak and mean firing rates, but virtually no link between information transmitted and the width of a cell's tuning curve.

Previous models of the information transmitted by HD cells have expressed the relative sensitivity of the cells as the number of bits/spike (Gothard et al., 1996; Jung & McNaughton, 1993; Markus et al., 1994; Poucet et al., 1994; Skaggs et al., 1993; Taube, 1995b; Taube & Muller, 1998). This measure (bits/spike) is obtained by dividing the information measure (bits/sec) by the mean firing rate of the cell (spikes/sec). Assuming the relation between the information rate and the mean firing rate is of the linear form shown in Figure 5, the relationship between (bits/sec) = x , and (bits/spike) = y is given by Equation 5:

$$y = \frac{x}{\alpha x + \beta}, \quad (5)$$

where, for our particular example, $\alpha = 0.85$, and $\beta = 3.4$. This relationship is graphed in Figure 7. Converting from bits/sec to bits/spike greatly compresses the information content of cells with high firing rates (the function is approximately logarithmic). Since there is a linear relation between bits/sec and bits/HD, the same nonlinear function shown in Figure 7 applies as well when converting from bits/HD into bits/spike. It is important to realize, therefore, that the *theoretical* relative sensitivity of cells to head direction depends critically on the units used to express information transmission. In particular, the practice in the literature has been to interpret the information measure of bits/spike as though it reflected the rat's ability to identify or discriminate among head directions. This interpretation is, however, misleading, as emphasized by

Table 2
Information and Efficiency Statistics

| | Information* | Efficiency** |
|----------|--------------|--------------|
| Minimum | 0.005 | 4 |
| Maximum | 0.19 | 44 |
| Mean | 0.04 | 23.3 |
| Median | 0.03 | 22.6 |
| Standard | 0.04 | 9.4 |
| Skewness | 1.8 | 0.06 |

*Bits/heading (except for skewness). **Percentage of maximum achievable (except for skewness).

the nonlinear relationship in Figure 7. By converting the information values into bits/spike, one essentially is standardizing each cell by its firing rate. Hence, the observed relationships between bits/heading or bits/sec and the statistics of cell firing rate are not captured by this measure.

The broader implication of the cell response model is that any demonstration at the behavioral level of discrimination among head directions that exceeds the transmission properties of single HD cells must be attributed to other factors. One strong possibility is that network properties of cells acting together are responsible for discriminating finer differences in head directions and could well be implicated in the ability of an animal to identify a larger set of head directions. We now explore this possibility by introducing a population model in which each cell is treated as a single response option.

Population Response Model

Assume the animal’s head points equally often in all directions (x). Then, the average directional uncertainty [$U(X)$] is $\log_2(360) = 8.49$ bits (Equation A1). Consider each cell in the population as a response generator (y_j). Assume the spike produced by such a cell represents its observable output and that its firing rate (spikes/sec) is normally distributed over variations in head direction (Equation 1), with SD (σ) and mean (μ). For our hypothetical example, the peak firing rate occurs at the preferred head direction of a cell. The preferred directions among all the cells in the population vary from 0° to 360° in 1° steps (the analysis is unaffected by step size). Then, the probability of each response generator [$p(y_j)$ in Equation A2] is obtained by adding up all the spikes of a cell across all head directions and dividing by the total number of spikes for each cell is the same for each member of the population, the average response uncertainty (Equation A2) reduces to [$U(Y)$] = $\log_2 360 = 8.49$ bits—the same as the directional uncertainty. In addition, because the firing rate profile of each cell across head directions is an identical Gaussian,⁴ the average conditional uncertainty $U(X | Y)$ (Equation A4) is the same as one such profile. Given a single cell (j), the probability of head direction i is

$$p(x_i | y_j) = \frac{\lambda_{ij}}{\sum_i \lambda_{ij}}, \tag{6}$$

where λ_{ij} is the firing rate of cell j at head direction i . The denominator is a standardization factor to convert firing rate into a probability. Both the numerator and denominator are determined by the Gaussian (Equation 1), describing firing rate ($\lambda_i = f$) as a function of head direction (d_i). Hence,

$$\lambda_{ij} = Ke^{-(1/2)[(d_i - \mu)/\sigma]^2}, \tag{7}$$

and

$$\sum_{i=1}^N \lambda_{ij} = K \sum_{i=1}^N e^{-(1/2)[d_i - \mu/\sigma]^2}. \tag{8}$$

Substituting Equations 7 and 8 into Equation 6 and averaging over all head directions, yields

$$U(X|Y) = -\sum_i \frac{N}{i} \frac{e^{-(1/2)[d_i - \mu/\sigma]^2}}{\sum_i e^{-(1/2)[d_i - \mu/\sigma]^2}} \log_2 \frac{e^{-(1/2)[d_i - \mu/\sigma]^2}}{\sum_i e^{-(1/2)[d_i - \mu/\sigma]^2}}. \tag{9}$$

Information transmitted can be found by Equation A7 in the Appendix:

$$I_t = U(X) - U(X | Y).$$

Substituting for $U(X)$ and $U(X | Y)$ in Equation A7 yields the information transmitted by the population

$$I_t = \log_2(N) + \sum_i \frac{N}{i} \frac{e^{-(1/2)[d_i - \mu/\sigma]^2}}{\sum_i e^{-(1/2)[d_i - \mu/\sigma]^2}} \log_2 \frac{e^{-(1/2)[d_i - \mu/\sigma]^2}}{\sum_i e^{-(1/2)[d_i - \mu/\sigma]^2}}. \tag{10}$$

This model was evaluated as follows. The value of N was taken as 360, the mean (μ) was arbitrarily set at 180° , and three separate parameter values for the SD (σ) were used to determine three different measures of the infor-

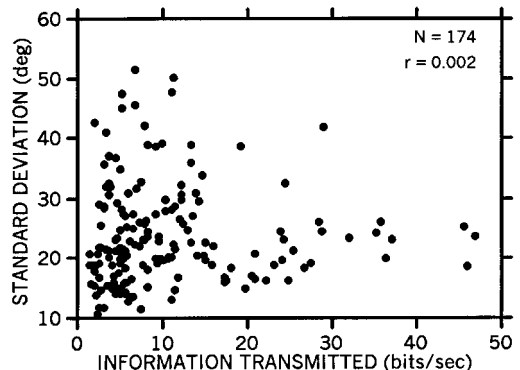


Figure 6. Standard deviation of Gaussian tuning curve (Equation 1) as a function of information transmitted according to the cell response model.

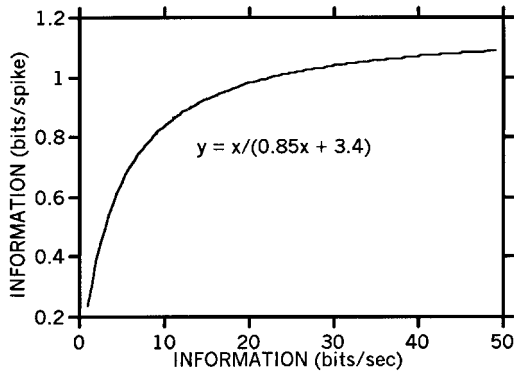


Figure 7. Theoretical relationship between information transmission of single cells expressed as bits/spike and bits/second.

mation transmitted. The value of d_i was varied over a range of plus and minus two SD s in integer steps (smaller step sizes do not influence the results). It was assumed in each case that all cells in the population had the same SD [either the minimum (10.9°), the median (22.0°), or the maximum (51.5°) obtained empirically (Table 1)]. This gives us a good indication of the range of values that could be expected in an actual population of cells with different widths of their tuning curves.

An evaluation of Equation 10 indicates that these three hypothetical populations of HD cells could transmit 3.2 (based on the minimum SD), 2.2 (median SD), and 0.98 (maximum SD) bits of information/heading. This result implies that the populations could distinguish absolutely (i.e., with 100% certainty) between 9, 5, and 2 different head directions, respectively. The largest value is reasonably close to the lower limit on information transmission of air current directional cells in the cricket, as reported by Theunissen and Miller (1991), though, of course, there is no good reason to expect the two species to yield the same values, since the two models are based on different assumptions.

DISCUSSION

No single theoretical perspective dictates the application of information theory to neural processing. Depending on the assumptions about the physiological meaning of the terms in the information-theoretic equations, different absolute and relative measures of information content are predicted for single cells, as well as for populations.

Bearing in mind that the maximum value for information content using the cell response model is 1 and that the measure is sensitive to changes in the assumed maximum firing rate, note how low the absolute numbers are with this model. The mean information transmitted is 0.04, with an SD of 0.04. Intuitively, this value appears low, especially for a signal that is so robust and contains a relatively high signal-to-noise ratio. In addition to stressing the importance of network properties, the small values of information content predicted by the cell response model also helps to explain why the nervous system needs

thousands of cells in order to perform many of its functions accurately.

The number of head directions that can be identified according to the population response model is approximately 5, based on the median value for SD of the Gaussian tuning curves. This amount is roughly one direction every 72° —a value somewhat less than the mean ($\sim 90^\circ$) reported for the directional firing range (Taube, 1995a, 1995b; Taube et al., 1990a).

One of the unresolved issues in regard to the statistical and informational properties of HD cells concerns inter-cell variability. It is not clear from a physiological standpoint why there is so much variability among cells in measures such as information content and peak firing rates, or why the distributions of these population statistics follow the gamma distribution. We are not speaking here of the *coarse coding* of stimulus features, but rather of the substantial differences among cells that ostensibly perform the same function. One possibility is that HD cells are involved in multiple tasks and that our methods reveal their role in only one of these tasks: the coding of head direction. They may exhibit more or less vigorous discharge rates to other stimulus conditions about which we are unaware at the present time, perhaps because our measurement techniques are not designed to record responses to such conditions. Although Dudchenko and Taube (1997) did not report any behavioral correlates, other than head direction, for HD cells in a spatial reference memory task, cell firing was modulated by angular and linear speed of head movements (Blair & Sharp, 1995; Taube, 1995a, 1995b; Taube & Muller, 1998; Taube et al., 1990a). Perhaps these variables are somehow implicated in the wide variations we see in mean firing rates.

We have shown that the information transmitted by individual cells is less than can be achieved by a population of cells whose members are differentially tuned to different head directions. Although this is not surprising, information theory alone does not reveal anything about the functional links among cells in a population. For example, our analyses are blind to the possibility that adjacent head directions may be coded by cells that mutually influence each other through synaptic connections. Any meaningful pattern among the firing rates within the cell population is therefore missed by this type of analysis. For example, it is possible to have a cell firing at a low rate in an apparently random fashion across head directions that nonetheless exhibits a high amount of information transmitted. Because of this limitation on information theory, it should only be applied when a cell's firing rate is coherently related to adjacent head directions, such as is the case when a Gaussian distribution describes the tuning curve over head direction.

The human channel capacity for a wide variety of sensory inputs generally falls between 2 and 3 bits/stimulus, indicating that a person is able to distinguish absolutely between four and eight alternatives that vary along a single physical dimension (e.g., light intensity, sound amplitude, and line length; Baird & Noma, 1978; Garner, 1962; G. A. Miller, 1956). Our results imply that

the rat's ability to distinguish among head directions on the basis of neural firing rates is well within this range, and, therefore, one is encouraged to model the neurophysiological basis of human channel capacity by emphasizing parameters associated with the activity of cell populations (Baird, 1997, Chap. 2).

It should be noted, however, that a measure of absolute identification is not the same as a measure of sensory discrimination. For example, a person is able to distinguish, on the basis of vision, millions of hues in a paired comparison discrimination task, whereas, the same individual is limited to about 8 to 12 hues when each stimulus must be identified by a unique color name (Eriksen & Hake, 1955). We are unaware of any behavioral or psychophysical studies (behavioral) with rats or humans in which absolute identification of head directions was assessed.

REFERENCES

- ARBIB, M. A., ÉRDI, P., & SZENTÁGOTHAJ, J. (1998). *Neural organization: Structure, function, and dynamics*. Cambridge, MA: MIT press.
- ATTNEAVE, F. (1959). *Applications of information theory to psychology*. New York: Holt, Rinehart & Winston.
- BAIRD, J. C. (1984). Information theory and information processing. *Information Processing & Management*, **20**, 373-381.
- BAIRD, J. C. (1997). *Sensation and judgment: Complementarity theory of psychophysics*. Mahwah, NJ: Erlbaum.
- BAIRD, J. C., & NOMA, E. (1978). *Fundamentals of scaling and psychophysics*. New York: Wiley.
- BARLOW, H. (1995). The neuron doctrine in perception. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 415-435). Cambridge, MA: MIT Press, Bradford Books.
- BLAIR, H. T., & SHARP, P. E. (1995). Anticipatory head direction signals in anterior thalamus: Evidence for a thalamocortical circuit that integrates angular head motion to compute head direction. *Journal of Neuroscience*, **15**, 6260-6270.
- CHEN, L. L., LIN, L. H., GREEN, E. J., BARNES, C. A., & McNAUGHTON, B. L. (1994). Head-direction cells in the rat posterior cortex. I. Anatomical distribution and behavioral modulation. *Experimental Brain Research*, **101**, 8-23.
- COVER, T. M., & THOMAS, J. A. (1991). *Elements of information theory*. New York: Wiley.
- DUDCHENKO, P. A., & TAUBE, J. S. (1997). Correlation between head direction cell activity and spatial behavior on a radial arm maze. *Behavioral Neuroscience*, **111**, 3-19.
- ERIKSEN, C. W., & HAKE, H. W. (1955). Multidimensional stimulus differences and accuracy of discrimination. *Journal of Experimental Psychology*, **50**, 153-160.
- FISHER, N. I. (1993). *Statistical analysis of circular data*. Cambridge: Cambridge University Press.
- GARNER, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.
- GEORGOPOULOS, A. P. (1995). Motor cortex and cognitive processing. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 507-517). Cambridge, MA: MIT Press, Bradford Books.
- GEORGOPOULOS, A. P., CAMINITI, R., KALASKA, J. F., & MASSEY, J. T. (1983). Spatial coding of movement: A hypothesis concerning the coding of movement direction by motor cortical populations. *Experimental Brain Research*, **7**, 327-336.
- GEORGOPOULOS, A. P., & MASSEY, J. T. (1988). Cognitive spatial motor processes: 2. Information transmitted by the direction of two-dimensional arm movements and by neuronal populations in primate motor cortex and area 5. *Experimental Brain Research*, **69**, 315-326.
- GERSHON, E. D., WIENER, M. C., LATHAM, P. E., & RICHMOND, B. J. (1998). Coding strategies in monkey V1 and inferior temporal cortices. *Journal of Neurophysiology*, **79**, 1135-1144.
- GOTHARD, K. M., SKAGGS, W. E., MOORE, K. M., & McNAUGHTON, B. L. (1996). Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task. *Journal of Neuroscience*, **16**, 823-835.
- JUNG, M. W., & McNAUGHTON, B. L. (1993). Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus*, **3**, 165-182.
- JUNG, M. W., WIENER, S. I., & McNAUGHTON, B. L. (1994). Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *Journal of Neuroscience*, **14**, 7347-7356.
- KATZ, B. (1966). *Nerve, muscle, and synapse*. New York: McGraw-Hill.
- MACKAY, D. M., & McCULLOCH, W. S. (1952). The limiting information capacity of a neuronal link. *Bulletin of Mathematical Biophysics*, **14**, 127-135.
- MARKUS, E. J., BARNES, C. A., McNAUGHTON, B. L., GLADDEN, V. L., & SKAGGS, W. E. (1994). Spatial information content and reliability of hippocampal CA1 neurons: Effects of visual input. *Hippocampus*, **4**, 410-421.
- MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on capacity for processing information. *Psychological Review*, **63**, 81-97.
- MILLER, J. P., JACOBS, G. A., & THEUNISSEN, F. E. (1991). Representation of sensory information in the cricket cercal sensory system. I. Response properties of the primary interneurons. *Journal of Neurophysiology*, **66**, 1680-1689.
- MIZUMORI, S. J. Y., & WILLIAMS, J. D. (1993). Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats. *Journal of Neuroscience*, **13**, 4015-4028.
- MULLER, R. U. (1996). A quarter of a century of place cells. *Neuron*, **17**, 813-822.
- NORWICH, K. H. (1993). *Information, sensation, and perception*. New York: Academic Press.
- O'KEEFE, J., & DOSTROVSKY, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, **34**, 171-175.
- O'KEEFE, J., & NADEL, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press, Clarendon Press.
- OLTON, D. S., COLLISON, C., & WERZ, M. A. (1977). Spatial memory and radial arm maze performance of rats. *Learning & Motivation*, **8**, 289-314.
- O'MARA, S. M. (1995). Spatially selective firing properties of hippocampal formation neurons in rodents and primates. *Progress in Neurobiology*, **45**, 253-274.
- OPTICAN, L. M., & RICHMOND, B. J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *Journal of Neurophysiology*, **57**, 162-178.
- POUCET, B., THINUS-BLANC, C., & MULLER, R. U. (1994). Place cells in the ventral hippocampus of rats. *Neuroreport*, **5**, 2045-2048.
- RICHMOND, B. J., & OPTICAN, L. M. (1990). Temporal encoding of two-dimensional patterns by single units in primate visual cortex: II. Information transmission. *Journal of Neurophysiology*, **64**, 370-380.
- RIEKE, F., WARLAND, D., DE RUYTER VAN STEVENINK, R., & BIALEK, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- SHANNON, C. E., WEAVER, W. (1963). *The mathematical theory of communication*. Urbana: University of Illinois Press.
- SKAGGS, W. E., McNAUGHTON, B. L., GOTHARD, K. M., & MARKUS, E. J. (1993). An information-theoretic approach to deciphering the hippocampal code. *Advances in Neural Information Processing Systems*, **5**, 1030-1037.
- STACKMAN, R. W., & TAUBE, J. S. (1997). Firing properties of head direction cells in rat anterior thalamic neurons: Dependence upon vestibular input. *Journal of Neuroscience*, **17**, 4349-4358.
- STACKMAN, R. W., & TAUBE, J. S. (1998). Firing properties of cells recorded in the lateral mammillary nuclei of rats: Head direction, angular velocity, and pitch cells. *Journal of Neuroscience*, **18**, 9020-9037.
- TAUBE, J. S. (1995a). Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *Journal of Neuroscience*, **15**, 70-86.
- TAUBE, J. S. (1995b). Place cells recorded in the parasubiculum of freely-moving rats. *Hippocampus*, **5**, 569-583.

- TAUBE, J. S. (1998). Head direction cells and the neurophysiological basis for a sense of direction. *Progress in Neurobiology*, **55**, 225-256.
- TAUBE, J. S., & BURTON, H. L. (1995). Head direction cell activity monitored in a novel environment and during a cue conflict situation. *Journal of Neurophysiology*, **74**, 1953-1971.
- TAUBE, J. S., & MULLER, R. U. (1998). Comparison of head direction cell activity in the postsubiculum and anterior thalamus of freely moving rats. *Hippocampus*, **8**, 87-108.
- TAUBE, J. S., MULLER, R. U., & RANCK, J. B., JR. (1990a). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *Journal of Neuroscience*, **10**, 420-435.
- TAUBE, J. S., MULLER, R. U., & RANCK, J. B., JR. (1990b). Head-direction cells recorded from the postsubiculum in freely moving rats: II. Effects of environmental manipulations. *Journal of Neuroscience*, **10**, 436-447.
- TAUBE, J. S., STACKMAN, R. W., & DUDCHENKO, P. A. (1996). Head-direction cell activity monitored following passive transport into a novel environment. *Society for Neuroscience Abstracts*, **22**, 1873.
- THEUNISSEN, F. E., & MILLER, J. P. (1991). Representation of sensory information in the cricket cercal sensory system: II. Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. *Journal of Neurophysiology*, **66**, 1690-1703.
- TOLMAN, E. C., RITCHIE, B. G., & KALISH, D. (1946). Studies in spatial learning: I. Orientation and the short-cut. *Journal of Experimental Psychology*, **36**, 13-24.
- TOVÉE, M. J., ROLLS, E. T., TREVES, A., & BELLIS, R. P. (1993). Information encoding and the responses of single neurons in the primate temporal visual cortex. *Journal of Neurophysiology*, **70**, 640-654.
- WIENER, S. I. (1993). Spatial and behavioral correlates of striatal neurons in rats performing a self-initiated navigation task. *Journal of Neuroscience*, **13**, 3802-3817.
- WILSON, M. A., & MCNAUGHTON, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, **261**, 1055-1058.
- ZHANG, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *Journal of Neuroscience*, **16**, 2112-2126.

NOTES

1. A circular normal function was also fit to the data. It yielded virtually identical results (see Fisher, 1993, and Zhang, 1996).
2. The maximum firing rate attained in our sample was 143 spikes/sec. Using the inverse of this rate as Δt does not alter any of the conclusions of this paper. Although the information transmission values are larger, they are still small relative to the theoretical maximum (mean = 0.08 bits/HD, $SD = 0.07$).
3. In the early literature on information theory, this was referred to as "relative entropy" (Attneave, 1959).
4. An alternative approach is to derive the information in a Gaussian distribution, discretize the result, and use this form to represent conditional uncertainty (see Baird, 1984).

APPENDIX A
Information Measures

The following equations are employed in the theoretical models discussed in the text. Details concerning the rationale for each equation are given in Baird and Noma (1978, Chap. 12).

Average stimulus uncertainty is

$$U(X) = -\sum_i^N p(x_i) \log_2 p(x_i), \quad (\text{A1})$$

where $p(x_i)$ is the probability of the i th environmental direction in a set of N . Average response uncertainty is

$$U(Y) = -\sum_j^M p(y_j) \log_2 p(y_j). \quad (\text{A2})$$

where $p(y_j)$ is the probability of the j th response (defined differently for each model). Average joint uncertainty is

$$U(X, Y) = -\sum_i^N \sum_j^M p(x_i, y_j) \log_2 p(x_i, y_j). \quad (\text{A3})$$

Average conditional uncertainty of $X|Y$ is

$$U(X|Y) = -\sum_j^M p(y_j) \sum_i^N p(x_i|y_j) \log_2 p(x_i|y_j). \quad (\text{A4})$$

Average conditional uncertainty of $Y|X$ is

$$U(Y|X) = \sum_i^N p(x_i) \sum_j^M p(y_j|x_i) \log_2 p(y_j|x_i). \quad (\text{A5})$$

Information transmitted is

$$I_t = U(X) + U(Y) - U(X, Y), \quad (\text{A6})$$

$$I_t = U(X) - U(X|Y), \text{ and} \quad (\text{A7})$$

$$I_t = U(Y) - U(Y|X), \quad (\text{A8})$$

It is also possible to derive an expanded version for information transmission that may prove useful for computational purposes. From Equations A2, A5, and A8 we get

$$I_t = -\sum_j^M p(y_j) \log_2 p(y_j) + \sum_i^N p(x_i) \sum_j^M p(y_j|x_i) \log_2 p(y_j|x_i), \quad (\text{A9})$$

but

$$\sum_j^M p(y_j) \log_2 p(y_j) = \sum_i^N p(x_i) \sum_j^M p(y_j|x_i) \log_2 p(y_j|x_i). \quad (\text{A10})$$

Thus,

$$I_t = -\sum_i^N p(x_i) \sum_j^M p(y_j|x_i) \log_2 p(y_j|x_i) + \sum_i^N p(x_i) \sum_j^M p(y_j|x_i) \log_2 p(y_j|x_i). \quad (\text{A11})$$

Equation A11 is given as Equation 3 in the text.

APPENDIX B
Derivation of Cell Response Model

Define x_i as head pointing in direction i during a time Δt long. Define y_1 as cell fires within the time Δt . Define y_2 as cell does not fire in time Δt . Substituting into Equation A11 yields

$$l_t = \sum_i^N p(x_i)p(y_1 | x_i) \log_2 \frac{p(y_1 | x_i)}{p(y_1)} + \sum_i^N p(x_i)p(y_2 | x_i) \log_2 \frac{p(y_2 | x_i)}{p(y_2)}, \quad (B1)$$

where $p(y_1)$ = (total spikes)/(number of time intervals of length Δt), and the total spikes is the sum of the firing rate λ_i times the amount of time (t_i) the head is pointing in direction i ; that is, if

$$\text{total spikes} = \sum_i^N \lambda_i t_i \quad (B2)$$

and

$$\text{number of time intervals} = \frac{T}{\Delta t}, \quad (B3)$$

then

$$p(y_1) = \sum_i^N \lambda_i t_i \frac{\Delta t}{T} = \lambda \Delta t, \quad (B4)$$

where the mean firing rate, λ , is

$$\lambda = \frac{1}{T} \sum_i^N \lambda_i t_i.$$

Then define

$$p(y_2) = 1 - \lambda \Delta t, \\ p(y_1 | x_i) = \lambda_i \Delta t, \\ p(y_2 | x_i) = 1 - \lambda_i \Delta t,$$

and

$$p(x_i) = \frac{t_i}{T},$$

Substituting into Equation B1 yields

$$l_t = \sum_i^N \frac{t_i}{T} \lambda_i \Delta t \log_2 \frac{\lambda_i \Delta t}{\lambda \Delta t} + \sum_i^N \frac{t_i}{T} (1 - \lambda_i \Delta t) \log_2 \frac{(1 - \lambda_i \Delta t)}{(1 - \lambda \Delta t)} \\ = \sum_i^N \frac{t_i}{T} \left[\lambda_i \Delta t \log_2 \frac{\lambda_i \Delta t}{\lambda \Delta t} + (1 - \lambda_i \Delta t) \log_2 \frac{(1 - \lambda_i \Delta t)}{(1 - \lambda \Delta t)} \right]. \quad (B5)$$

For the case where the times are the same for all head directions,

$$l_t = \frac{1}{N} \sum_i^N \left[\lambda_i \Delta t \log_2 \frac{\lambda_i \Delta t}{\lambda \Delta t} + (1 - \lambda_i \Delta t) \log_2 \frac{(1 - \lambda_i \Delta t)}{(1 - \lambda \Delta t)} \right], \quad (B6)$$

where

$$\lambda = \frac{1}{N} \sum_{i=1}^N \lambda_i.$$

Equation B6 is given as Equation 4 in the text.