Information Theory's Failure in Neuroscience

On the Limitations of Cybernetics

Lance Nizami
Independent Research Scholar
Palo Alto CA, USA
nizamii2@att.net

Abstract—In *Cybernetics* **(1961 Edition), Professor Norbert Wiener noted that "The role of information and the technique of measuring and transmitting information constitute a whole discipline for the engineer, for the neuroscientist, for the psychologist, and for the sociologist". Sociology aside, the neuroscientists and the psychologists inferred "information transmitted" using the discrete summations from Shannon Information Theory. The present author has since scrutinized the psychologists' approach in depth, and found it wrong. The neuroscientists' approach is highly related, but remains unexamined. Neuroscientists quantified "the ability of [physiological sensory] receptors (or other signal-processing elements) to transmit information about stimulus parameters". Such parameters could vary along a single continuum (e.g., intensity), or along multiple dimensions that altogether provide a Gestalt – such as a face. Here, unprecedented scrutiny is given to how 23 neuroscience papers computed "information transmitted" in terms of stimulus parameters and the evoked neuronal spikes. The computations relied upon Shannon's "confusion matrix", which quantifies the fidelity of a "general communication system". Shannon's matrix is square, with the same labels for columns and for rows. Nonetheless, neuroscientists labelled the columns by "stimulus category" and the rows by "spike-count category". The resulting "information transmitted" is spurious, unless the evoked spike-counts are worked backwards to infer the hypothetical evoking stimuli. The latter task is probabilistic and, regardless, requires that the confusion matrix be square. Was it? For these 23 significant papers, the answer is No.**

Keywords—**cybernetics, Wiener, information theory, neuroscience, spikes, confusion matrix**

I. INTRODUCTION

In Cybernetics [1, p. vii], Professor Norbert Wiener noted that "The role of information and the technique of measuring and transmitting information constitute a whole discipline for the engineer, for the neuroscientist, for the psychologist, and for the sociologist". Note well the mention of non-engineering disciplines. In particular, how psychologists have inferred transmitted information using the discrete summations from Shannon Information Theory [2] has been critically examined by the present author [3]-[6]. However, there has been no comparable scrutiny of the neuroscientists' use of Shannon's summations. The latter use greatly resembles what was done by the psychologists, making criticism not just timely but well-overdue. The task proves non-trivial, and it exposes misinterpretations that have persisted for decades. Further

problems exist, but are relegated to future papers, because of the depth required. Regardless, the present work challenges neuroscientists to re-evaluate what can be inferred from animal recordings, especially when taken as models of man.

Sensory neuroscientists have used Shannon Information Theory to quantify "the ability of [physiological sensory] receptors (or other signal-processing elements) to transmit information about stimulus parameters" [7, p. 82]. That usage is scrutinized here, starting with Werner and Mountcastle [8], who provided the interpretation of Shannon which became the blueprint for neuroscience. Table 1 describes Werner and Mountcastle and some of the significant papers that followed it, which cover a range of dates, species, sensory modalities, and types of stimuli. Some papers varied a single characteristic of their stimuli, one that fills a continuum, such as the depth of thrust of a blunt probe against the skin (experienced as differing pressure). As such, the neurons studied can be those closest to the outside world, the primary afferents. They are typically attached to, or themselves contain, sensory receptors. Other neurons that respond to sensory stimuli rely on input from the peripheral afferents, but are closer to the brain. The cerebral cortex itself can contain neurons that preferentially respond to a Gestalt formed from multiple stimulus features.

Neuroscientists' use of Information Theory has been extremely influential in behavioral science. Unfortunately, as will be shown, the information-theory analyses in the papers listed in Table 1 are meaningless. Attempts to correct the computation, by relating responses back to stimuli ("Decoding"; see below), are likewise futile.

Note well that the information-theory analyses discussed here employed the original discrete *summations* of Shannon [2]. Smooth *integral* forms exist, and some neuroscientists have used them. But those are special approaches (e.g., [30]) having special problems, deserving later critique. However, given the present revelations, neuroscientists' use of the integrals likely will also prove faulty.

II. THE "GENERAL COMMUNICATION SYSTEM", THE "CONFUSION MATRIX", AND THE INFORMATION TRANSMITTED

In Shannon's own words [2], his "general communication system" consists of (1) "An *information source* which produces a message or sequence of messages to be communicated to the receiving terminal", (2) "A *transmitter* which operates on the message in some way to produce a

signal suitable for transmission over the channel", (3) a *channel*, where "The *channel* is merely the medium used to transmit the signal from transmitter to receiver", (4) a *receiver*, where "The *receiver* ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal", and (5) a *destination*, where "The *destination* is the person (or thing) for whom the message is intended" (the italics within the quotes are the originals).

The sent messages are groups of symbols, such as letters arranged into words and then into sentences [2]. Symbols have probabilities from which the information transmitted in a message can be computed, as follows [2]. Consider n possible *events*; the one that occurs is the *outcome*, which is uncertain if n > 1. If each event has a known probability of occurrence, p_i, i = 1,..., n, where $\sum_{i=1}^{n} p_i = 1$, then the requisite [amount of] "uncertainty", "choice", or "information", called I_S, is

$$I_S = -K \sum_{i=1}^{n} p_i \log p_i, \ K > 0 \qquad (1)$$

Shannon set K =1. If events are symbols denoted "k", then signal uncertainty/information I_S is

$$I_S = -\sum_{k} p(k) \log p(k) \qquad (2)$$

In experiments, $p(k)$ is set by the experimenters.

A symbol received is from the set of symbols that can be sent, but not all symbols will be received as sent. That is, the system is "noisy". Information transmitted, denoted I_t, can be calculated knowing (a) what symbols "k" (events) were sent, (b) what symbols "j" (outcomes) were received, and (c) the number of times a symbol sent as "k" was received as "j", denoted N_{jk}. The latter form the "confusion matrix".

Fig. 1 shows the confusion matrix for a *total number N* symbols sent. Let $p(k) = N_{.k}/N$ = the probability that k was sent, and $p_j(k) = N_{jk}/N_{j.}$ = the probability that k was sent if j was received, from which

$$E_S = -\sum_{j}\sum_{k} p_j(k) \ \log p_j(k) \qquad (3)$$

is the signal equivocation/entropy, and

$$I_t = I_S - E_S = -\sum_{k} p(k) \log p(k) + \sum_{j}\sum_{k} p_j(k) \ \log p_j(k) \quad (4)$$

is the information transmitted. These are averages taken over the course of the message. The base of the logarithms is a positive integer; when set to 2, the information transmitted is in binary information units per symbol, or "bits/symbol". Note that $I_S \ge I_t \ge E_S \ge 0$; $I_t = I_S$ represents perfect transmission. In the coming details, "signal" will be replaced by "stimulus". Note also that if the n symbols or stimuli are equiprobable (that is, $p_i = 1/n$ in Eq. (1), $p(k) = 1/n$ in Eq. (2)), then for

base b we have $n = b^{I_S}$, and the number of symbols or stimuli that the system can potentially correctly indicate is b^{I_t}.

	Symbol sent (event)						Row totals
	1	2	-	k	-	n	
Symbol received (outcome)							
1	N_{11}	N_{12}	-	N_{1k}	-	N_{1n}	$N_{1.}$
2	N_{21}	N_{22}	-	N_{2k}	-	N_{2n}	$N_{2.}$
-	-	-	-	-	-	-	-
j	N_{j1}	N_{j2}	-	N_{jk}	-	N_{jn}	$N_{j.}$
-	-	-	-	-	-	-	-
n	N_{n1}	N_{n2}	-	N_{nk}	-	N_{nn}	$N_{n.}$
Column totals	$N_{.1}$	$N_{.2}$	-	$N_{.k}$	-	$N_{.n}$	**Sum =N**

Fig. 1. Shannon's Information Theory "confusion matrix".

Henceforth, the information transmitted computed according to Shannon [2] is denoted $I_{t,true}$, to distinguish it from the information transmitted in neuroscientists' Werner and Mountcastle [8] interpretation, here denoted $I_{t,WM}$.

III. SHANNON INFORMATION THEORY IN NEUROSCIENCE

A. Why use Information Theory in Neuroscience? The Variability of Spike Counts in Sensory Neurons

Applying a sensory stimulus to an organism that has neurons, such as a man or a cat, may evoke firing of voltage spikes in some of those neurons. Such neurons may respond predictably, within some limit of variability, to changes in some characterizing feature(s) of the stimulus. In particular, the number of neuronal voltage spikes evoked during some sub-interval of the interval in which the stimulus is applied – the sub-interval called the *spike-counting time* – will change systematically as the stimulus characteristic is changed. This spike count is the most commonly used of numerous quantities obtainable from neuronal spike trains. For such measures, the mean stimulus-evoked spike count tends to smoothly, systematically change in response to smooth, systematic change(s) in the sensory stimulus. Those measures include the latency of the first stimulus-evoked spike (see for example [31]; [25], for mean response time; [27]) or interspike intervals (e.g., [23]) or "the principal components of the temporal waveform of the response" ([12, p. 168]; also [14]-[16], [18]). Conclusions reached presently about the viability of information theory in neuroscience will apply equally well regardless of the particular neuronal response that has been used in computing $I_{t,WM}$, because those conclusions do not intrinsically depend upon the particulars of the neuronal response measure, but rather on how it is employed in

computing $I_{t,WM}$. Likewise, the present conclusions also apply when information-transmitted is computed from more than one simultaneous measurable stimulus-evoked neuronal response, i.e., a vector of responses, representing an ensemble of (usually mutually neighboring) neurons (e.g., [32]).

Whether the stimulus is varied along a single continuum or along many simultaneously (in the case of a Gestalt), each application of the *same* stimulus could, empirically, produce a *different* count of the voltage spikes evoked within the counting time. The likelihood that a given count lies within any two limits, over repeated applications of the stimulus, can be approximated for infinitely close limits by a probability density function having a mean and a variance. That is, spike count is *stochastic*, so that the response to the stimulus appears to be inherently "confused". The degree of that confusion is what neuroscientists *believe* that they quantify through their use of Information Theory. (The reason for the qualifying italics will hopefully soon become clear.) Experiments are typically done under conditions such that each spike-count distribution is believed to be unchanging (i.e., *stationary*). To sensory neuroscientists, all of this makes a particular stimulus a potential Shannon "event", because it evokes in the neuron, *on average*, a predictable "outcome". The measured spike-train feature was taken as Shannon's "symbol received" and the stimulus characteristic being varied (e.g., intensity) was taken as Shannon's "symbol sent". Of course, if the latter bore no relation at all to the former, then the transmitted information was zero; $I_S = E_S$ in Eq. (4).

The difference between any "event" and its evoked "outcome" could be attributed to "noise", which neuroscientists interpreted as the probabilistic nature of the evoked spike count. As Smith et al. [7, p. 94] declared concerning spike counts taken from primary afferent neurons, "The ability of an organism to make decisions about stimulus events is limited by the resolution of its peripheral sensory receptors … The concepts and methods of information theory [omitted citation] have provided a means for addressing the resolving power of these receptors". Other workers replaced "peripheral sensory receptors" with "sensory neurons", with the understanding that peripheral sensory receptors, regardless of further processing of their output at "higher" brain-ward neurons, did indeed provide the ultimate limit. That is, $I_{t,WM}$ was computed from neuronal responses (whether one neuron or several, and regardless of their bodily location) on the tacit presumption that $I_{t,WM}$ had a putative maximum $I_{t,WM\,\text{max}}$ that identified the neuron's ability to encode the stimulus characteristic(s) altered by the experimenter.

B. The Sensory Neuroscientists' Confusion Matrix

1) General description

In the neuroscientists' confusion matrix, in contrast to Shannon's, the "symbols sent" ("events") are stimuli, divided into ranges called *stimulus categories*. The "symbols received" ("outcomes") are numbers of voltage spikes evoked in the neuron over some counting time. As noted above, repetitions of a given stimulus can produce different numbers of spikes within some agreed-upon counting time; there is presently no means of determining the "correct" counting time, *should one exist*. Perhaps for this reason, neuroscientists treated *ranges* of spike counts as the labels for the rows of the confusion matrix, *spike-count categories*. Fig. 2 shows such a confusion matrix, but in which (unlike the literature) the number of stimulus categories equals the number of spike-count categories, i.e., a *square* matrix as in Fig. 1. Experiment-wise, each stimulus category was typically represented by a single stimulus, which differed from those in other categories (for a rare exception see [17], in which 37 stimuli were sometimes grouped into 3 categories). Thus, each matrix entry in Fig. 2 is typically the number of times that a particular stimulus evoked a spike count that fell within the particular spike-count category, i.e. $p(k)$ is "the probability of giving stimulus k" whereas $p_j(k)$

is "the probability that stimulus k was given, given a spike count in spike-count category j".

As noted, square matrices are *not* typical of the literature. Fig. 2 represents what a confusion matrix might look like if firing characteristics were to be used to "work backwards" to infer what the evoking stimuli had been. That exercise has sometimes been attempted (see below). The actual construction of the majority of neuroscientists' confusion matrices will now be described.

2) The neuroscientists' confusion matrix for stimuli characterized by values along a smooth continuum

Werner and Mountcastle [8] provided the archtype sensory neuroscience confusion matrix. All subsequent papers, whether using sums or integrals to compute information transmitted, follow their model. Werner and Mountcastle recorded spikes evoked in primary afferent neurons by a narrow probe quickly applied and held for 1 second or less against the center of touch-sensitive Iggo corpuscles in the shaven hairy skin of the hind limb of cats and monkeys (the specific monkey species was not mentioned, which became a common error). The stimulus intensities were the depths-of-depression of the probe, each possible value equidistant in micrometers (microns) from its nearest neighbors, each its own stimulus-intensity category. Probe-evoked spikes were counted for "a time which begins within 100-150 msec. [milliseconds] of stimulus onset and lasts for about 4-500 msec." [8, p. 379]. Spike-count categories were made by dividing the stimulus-evoked spike counts into groups differing by 2 spikes.

Table 1 lists the widths (W, in numbers of spikes) of the spike-count categories, as well as C, the numbers of columns, and R, the numbers of rows, in the confusion matrices, for Werner and Mountcastle [8] and subsequent others. If C=R, a matrix is square. The duration of application of the stimulus could be varied, as could the spike-counting duration. The spike-count categories would inevitably not all have equal widths; empirically, mean spike-firing rates are nonlinear functions of relevant stimulus characteristics. Hence, experimenters usually chose *stimuli* at equal intervals, along a range of the relevant stimulus characteristic within which the evoked firing rate is a straight line when plotted versus either linear or logarithmic scales of said characteristic. The experimenters likewise chose equal widths for their spike-count categories, except perhaps the uppermost and lowermost

categories, which might be extended to their respective infinities.

3) The neuroscientists' confusion matrix for stimuli characterized by their Gestalt

Here, as in the examples above which varied the stimulus intensity, each stimulus is its own stimulus category.

Tovee et al. [15] provide an example of brain recordings in response to Gestalt stimuli. The stimuli were "4 face stimuli" [15, p. 642] which "were monkey or human, usually in frontal view, but in some cases in profile view" [15, p. 642]. The faces were presented to attentive macaque monkeys who were trained to fixate at 5 different points of any image, actual visible *points* that were presented either at the center of the image, or at the edge of the image and towards any of the four corners of the video monitor used for viewing. The subjects thus maintained a constant behavior, giving altogether 4x5=20 "stimulus categories". Fixation points were "blinked off 200 ms [milliseconds] before the test stimulus appeared" [15, p. 641]. The latter lasted 500 ms.

IV. PROBLEMS WITH THE USE OF SHANNON INFORMATION THEORY IN SENSORY NEUROSCIENCE

Neuroscientists' use of Shannon Information Theory has a fatal problem concerning the interpretation of the Shannon concepts of "event" and "outcome", as follows.

A. Shannon's Requirements (i.e., The Theory)

1) Each "event" is from a limited set of distinct things, and evokes an "outcome", one of the possible "events"

For "events" (symbols transmitted) and "outcomes" (symbols received), Shannon [2] used, as examples, English letters. They form a limited set whose members are mutually distinguishable, insofar as a sent "A" received as an "A" will not normally be confused with a "B". To Shannon, each "event" had just one "outcome", that is one of the "events".

2) The Shannon confusion matrix is square

The symbol received need not be the one transmitted. Regardless, Shannon's confusion matrix has rows labelled by the same symbols as its columns; the matrix is square (Fig. 1).

B. Sensory Neuroscience (i.e., The Practice)

1) Each "event" is from a limited set of distinct things, and evokes an "outcome" – not one of the possible "events", but altogether different in character

a) Each "event" is from a limited set of distinct things

The sensory neuroscientists' "events" are stimuli – or perhaps stimulus categories. This uncertainty arises because a stimulus characterization (such as intensity) forms what, to our measuring instruments, seems to be a continuum (i.e., macroscopically smooth, even if bumpy at the quantum scale). And even when the stimulus is a Gestalt such as a face, small changes in the face (along one or more dimensions) might be indiscriminable to the judge (the human eye, and the mind connected to it) – such that faces represent a multi-dimensional continuum. In fact, in all of the 23 papers scrutinized here, each stimulus category k contained a stimulus

of just one characterization from the relevant set of characterizations (face, shape, intensity, or whatever).

Compelling reasons for that, which are not mentioned in the literature, are as follows. Consider that if different stimuli within a given stimulus category differ sufficiently that the resulting evoked spikes *consistently* fall into different spike-count categories, then each of the differing stimuli is acting as if it could be placed within a separate stimulus category. Conversely, if m different stimuli within the given stimulus category are sufficiently close in character that the resulting spike counts fall within just one spike-count category, then those m different stimuli are emulating a single stimulus repeated m times. These problems confuse the interpretation of $I_{t,WM}$. The confusion is avoided only when each stimulus category is restricted to a single stimulus which differs from all the others employed, such that the difference between "stimulus" and "stimulus category" disappears.

	Stimulus category					
	1	**2**	**-**	**k**	**-**	**n**	**Row totals**
Spike-count category						
1	N_{11}	N_{12}	-	N_{1k}	-	N_{1n}	$N_{1.}$
2	N_{21}	N_{22}	-	N_{2k}	-	N_{2n}	$N_{2.}$
-	-	-	-	-	-	-	-
j	N_{j1}	N_{j2}	-	N_{jk}	-	N_{jn}	$N_{j.}$
-	-	-	-	-	-	-	-
n	N_{n1}	N_{n2}	-	N_{nk}	-	N_{nn}	$N_{n.}$
Column totals	$N_{.1}$	$N_{.2}$	-	$N_{.k}$	-	$N_{.n}$	**Sum = N**

Fig. 2. The "confusion matrix" of sensory neuroscience.

b) Each "event" evokes an "outcome" – not one of the possible "events", but altogether different in character

The neuroscientists' "outcome" is a spike count. Or is it a spike-count *category*? Neuroscientists don't clarify this. Regardless, neither a spike-count category nor a spike count is a stimulus. Hence, regardless of whether the "outcomes" are taken as spike counts or as spike-count categories, "outcomes" do not have the same character as "events".

But can "outcomes" be made to have the same character as "events"? To compute a meaningful $I_{t,true}$ would require working backwards from each stimulus-evoked spike count to the stimulus that the spike-count *implied*, which would not necessarily be the stimulus actually given. To attempt working backwards, how would the number of stimulus categories have to compare to the number of spike-count categories? The answer is that each spike-count category j (representing a row in the sensory neuroscientists' confusion matrix) must have a single corresponding, "correct" stimulus category k (representing a column of the matrix). And even so, specification of the *implied* stimulus category would be

uncertain, because (1) the relation between stimulus and *mean* spike count, is usually nonlinear (there are too many supporting references for this) and (2) the spike count evoked by a given stimulus is stochastic, as noted above. This issue is pursued further in Section V, below.

2) The neuroscience confusion matrix is not square

Did the papers discussed above meet the requirement of equal numbers of spike-count categories and stimulus categories, that is, a square confusion matrix? Table 1 lists the numbers of rows and columns in the respective published matrices, and shows that none were square. Indeed, their creators revealed no inkling that squareness was required.

V. APPARENT EXCEPTIONS: "DECODING"

Shannon's system [2] contains an encoder (the transmitter) and a decoder (the receiver). The latter was assumed to convert the "noisy signal" back into a message containing a subset of the symbols that are available to form the message that was originally given to the transmitter. Thus the engineer's confusion matrix has rows and columns having the same labels. Did sensory neuroscientists ever attempt to "decode" the stimulus-evoked neuronal firing, in the sense that a decoder "receives a spike train as input and guesses which stimulus evoked the spike train" [25, p. 200]? Yes – for example, Georgopoulos and Massey [33] and Furukawa and Middlebrooks [31] had rows labelled the same as columns, that is, by the stimulus characteristic employed. They accomplished the seemingly impossible task that is sometimes called "reconstruction", "simulation", or "decoding" (e.g., [25], [35]) by *inferring* the entries in their square confusion matrices, based upon their recorded responses of neurons, and using statistical models based upon varieties of assumptions. The methods are too elaborate to be described here. Importantly, neither paper ([31], [33]) showed any explicit understanding of *why* their matrix was square. Georgopoulos and Massey [33], for example, merely cited Sakitt [34], who had used square matrices in psychophysics.

Let us denote the *inferred* values of $I_{t,true}$ as $I_{t,true}$(inferred) . Consider what would happen if circumstances existed which favored back-calculation. For example, the neurons examined in [8]-[11] showed remarkably reproducible spike firing, in that the variance of the firing rate to a given stimulus was remarkably low. And there was evidently a method to lower the variance-to-mean ratio in the counted spikes, namely, to employ a longer recording epoch (for explicit evidence see, for example, Rogers et al. [21, p. 457]), as $I_{t,WM}$ tended to rise as recording epoch was lengthened ([8, Fig. 22]; [10, Fig. 6]; [7, Fig. 8]; [15, Fig. 13]; [20, Fig. 9]; [24, Fig. 3]; [27, Fig. 3]; but for an exceptional momentary dip see [21, Fig. 6, top panel, several traces] and [22, Fig. 3]). As such, $I_{t,WM}$ could be made as arbitrarily close to I_S as neuronal biochemistry allowed, or as close to 2.81 as desired (such that the number of discernable categories was $2^{2.81} \approx 7$, the result from psychophysical "absolute identification" experiments; see [3]), by simply (1) altering the number of different stimuli (i.e., the size of the stimulus set)

and their spacing along the stimulus characteristic that was manipulated, while simultaneously (2) gerrymandering the number and widths of the spike-count categories. And close reading of [7] and [8]-[11] suggests that this is, in fact, what had been attempted. Fig. 3 illustrates such a case, given simplifying assumptions (for illustration's sake) of equal numbers of spike-count and stimulus categories, and distributions of equal variance. Here, the distribution of spike-counts evoked by each different stimulus lies effectively within the corresponding spike-count category. In such cases, there would be little purpose to computing $I_{t,true}$(inferred), as it, like $I_{t,WM}$, would closely approach I_S .
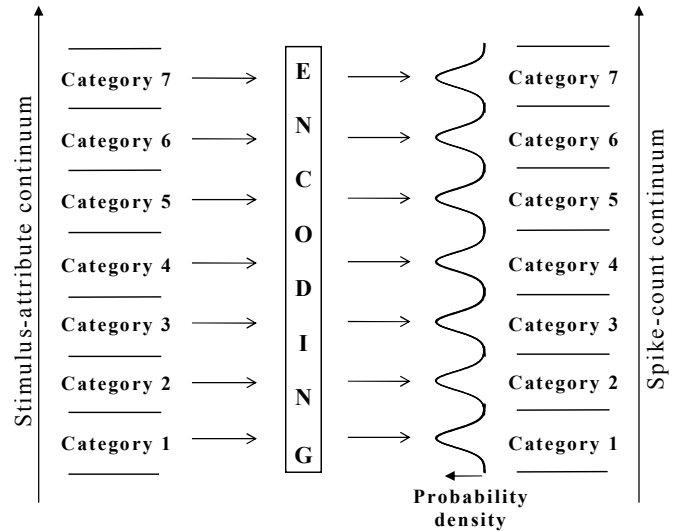

Fig. 3. Relation of stimulus categories to distributions of spike counts when information transmitted, I_t, approximates the stimulus information, I_S. A stimulus representing each stimulus category evokes neuronal voltage spikes which are counted over some counting time. Evoked spike count, over repeated presentations of the stimulus, is imagined as being distributed according to a Gaussian probability density function, assumed identical for the different given stimulus intensities. Tracing backward from spike-count category to stimulus category leaves little ambiguity about which recorded spike-count corresponded to which stimulus category.

Now consider the opposite assumption about the distribution of spike-counts for any spike-count category, namely, that it is very broad – so broad, in fact, that $I_{t,WM} << I_S$. Fig. 4 illustrates broad distributions. This is the sort of situation found, for example, in Tolhurst [13], in which spike-count "variance is usually 2-5 times the average response" [13, p. 410], resulting in a mean $I_{t,WM}$ of 0.71 bits (i.e., less than 2 correctly identified choices) across 23 neurons, given 14 different stimuli ($I_S = 3.81$ bits). Then, backwards-calculation of stimulus category from spike count would be a matter of probability (as used in Signal Detection Theory, for example; see [36], [37]) rather than certainty, so that $I_{t,true}$(inferred) $<< I_S$; indeed, $I_{t,true}$ (inferred) could well approach zero. $I_{t,WM}$ was indeed, in many of the papers listed in Table 1, remarkably small (data omitted).

There is great irony here, in that the very reason that Shannon's Information Theory was applied to neurons in the

first place – their stochastic firing – may be the reason why such a computation cannot work. I_t could be yet closer to zero if the numbers of spike-count categories and stimulus categories were unequal; in such a case, the back-assignment of spike-count to the hypothetical evoking stimulus would involve even more miss-assignment, which, on average, would reduce (rather than inflate) $I_{t,true}$ (inferred).
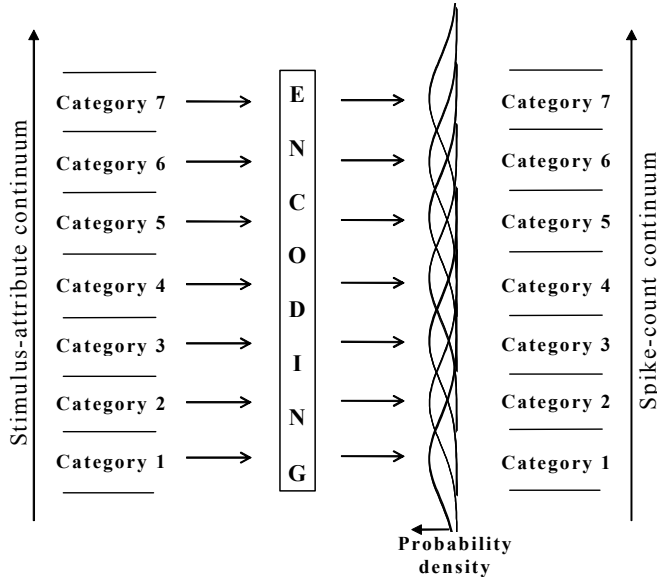

Fig. 4. Relation of stimulus categories to distributions of spike counts when information transmitted, I_t, is likely to be substantially less than the stimulus information, I_S. After Fig. 3, each spike-count distribution is Gaussian, and is identical for the different given stimulus intensities, but the variances are substantially broader than in Fig. 3 – so broad that tracing backward from a given spike-count to the stimulus that evoked it is effectively impossible. The stimulus-attribute and spike-count continua are pictured as being open at their top and bottom; otherwise, the spike-count distributions would become increasingly skewed as the end spike-count categories are approached (as seen in psychology "absolute identifications" [4]).

For there to be any possibility of inferring $I_{t,true}$ from neuronal responses, the neuronal code must be understood for the stimulus property that the experimenter is manipulating, in order that neuronal firing can be appropriately apportioned into response categories that are appropriate for the given stimulus categories. And yet, without such knowledge, $I_{t,WM}$ is currently used to choose between different possible neuronal codes (e.g., spike count versus first-spike latency versus principal components; see for example [12], [14]-[16], [18], [25], [27], [28], [31]).

VI. SUMMARY AND CONCLUSIONS

Werner and Mountcastle [8] first described an *interpretation* of Shannon's Information Theory [2], one later adopted by many neuroscientists, to estimate I_t, the information transmitted by sensory neurons. The computation depends upon the numerical entries in Shannon's "confusion matrix", which quantifies the fidelity of Shannon's "general communication system". The Shannon matrix has columns

labelled by symbol sent ("event") and rows labelled by symbol received ("outcome"). Shannon assumed that for each "event" there is only one "outcome", one of the possible "events". Therefore Shannon's matrix is necessarily square. But in the sensory neuroscientists' matrix, columns were labelled by "stimulus category" and rows were labelled by "spike-count category", neuronal voltage spikes being the response to stimuli. This does not yield a true I_t, because spike-generation is stochastic, with a mean stimulus-dependence that is nonlinear. $I_{t,true}$ could hypothetically still be computed if spike counts could somehow be related backwards to their implied stimulus, *given a square confusion matrix*. Squareness was therefore checked for, in 23 significant sensory neuroscience studies. Squareness was not found, and indeed, neither Werner and Mountcastle [8] nor any of their successors explained why they had ignored Shannon's requirement of identical labelling for rows and columns. The I_t's computed in those studies over five decades, here denoted $I_{t,WM}$, are gratuitous, yet they have been given great cachet and continue to be computed.

REFERENCES

[1] N. Wiener, Cybernetics: or Control and Communication in the Animal and the Machine. Cambridge, MA: MIT Press, 1961.

[2] C.E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623-656, 1948.

[3] L. Nizami, "Interpretation of absolute judgments using information theory: 'channel capacity' or memory capacity?," *Cyber. Hum. Know.*, vol. 17, pp. 111-155, 2010.

[4] L. Nizami, "Memory model of information transmitted in absolute judgment," *Kybernetes*, vol. 40, pp. 80-109, 2011.

[5] L. Nizami, "Norwich's Entropy Theory: how not to go from abstract to actual," *Kybernetes*, vol. 40, pp. 1102-1118, 2011.

[6] L. Nizami, "Paradigm versus praxis: why psychology 'absolute identification' experiments do not reveal sensory processes," *Kybernetes*, vol. 42, pp. 1447-1456, 2013.

[7] D.V. Smith, E. Bowdan, and V.G. Dethier, "Information transmission in tarsal sugar receptors of the blowfly," *Chem. Senses*, vol. 8, pp. 81-101, 1983.

[8] G. Werner and V.B. Mountcastle, "Neural activity in mechanoreceptive cutaneous afferents: stimulus-response relations, Weber functions, and information transmission," *J. Neurosci.*, vol. 28, pp. 359-397, 1965.

[9] I. Darian-Smith, M.J. Rowe, and B.J. Sessle, " 'Tactile' stimulus intensity: information transmission by relay neurons in different trigeminal nuclei," *Science*, vol. 160, pp. 791-794, 1968.

[10] B. Kenton and L. Kruger, "Information transmission in slowly adapting mechanoreceptor fibers," *Exp. Neurol.*, vol. 31, pp. 114-139, 1971.

[11] A.G. Hannam and T.J. Farnsworth, "Information transmission in trigeminal mechanosensitive afferents from teeth in the cat," *Arch. Oral Bio.*, vol. 22, pp. 181-186, 1977.

[12] L.M. Optican and B.J. Richmond, "Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis," *J. Neurophysiol.*, vol. 57, pp. 162-178, 1987.

[13] D.J. Tolhurst, "The amount of information transmitted about contrast by neurones in the cat's visual cortex," *Vis. Neurosci.*, vol. 2, pp. 409-413, 1989.

[14] J.W. McClurkin, L.M. Optican, B.J. Richmond, and T.J. Gawne, "Concurrent processing and complexity of temporally encoded neuronal messages in visual perception," *Science*, vol. 253, pp. 675-677, 1991.

[15] M.J. Tovee, E.T. Rolls, A. Treves, and R.P. Bellis, "Information encoding and the responses of single neurons in the primate temporal visual cortex," *J. Neurosci.*, vol. 70, pp. 640-654, 1993.

[16] M.J. Tovee and E.T. Rolls, "Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex," *Vis. Cognit.*, vol. 2, pp. 35-58, 1995.

[17] J.W. Gnadt and B. Breznen, "Statistical analysis of the information content in the activity of cortical neurons," *Vis. Res.*, vol. 36, pp. 3525-3537, 1996.

[18] E.T. Rolls, H.D. Critchley, and A. Treves, "Representation of olfactory information in the primate orbitofrontal cortex," *J. Neurophysiol.*, vol. 75, pp. 1982-1996, 1996.

[19] Y. Sugase, S. Yamane, S. Ueno, and K. Kawano, "Global and fine information coded by single neurons in the temporal visual cortex," *Nature*, vol. 400, pp. 869-873, 1999.

[20] D.D. Gehr, H. Komiya, and J.J. Eggermont, "Neuronal responses in cat primary auditory cortex to natural and altered species-specific calls," *Hear. Res.*, vol. 150, pp. 27-42, 2000.

[21] R.F. Rogers, J.D. Runyan, A.G. Vaidyanathan, and J.S. Schwaber, "Information theoretic analysis of pulmonary stretch receptor spike trains," *J. Neurosci.*, vol. 85, pp. 448-461, 2001.

[22] E. Arabzadeh, S. Panzeri, and M.E. Diamond, "Whisker vibration information carried by rat barrel cortex neurons," *J. Neurosci.*, vol. 24, pp. 6011-6020, 2004.

[23] T. Lu and X. Wang, "Information content of auditory cortical responses to time-varying acoustic stimuli," *J. Neurophysiol.*, vol. 91, pp. 301-313, 2004.

[24] L.C. Osborne, W. Bialek, and S.G. Lisberger, "Time course of information about motion direction in visual area MT of macaque monkeys," *J. Neurosci.*, vol. 24, pp. 3210-3222, 2004.

[25] I. Nelken, G. Chechik, T.D. Mrsic-Flogel, A.J. King, and J.W.H. Schnupp, "Encoding stimulus information by spike numbers and mean response time in primary auditory cortex," *J. Comput. Neurosci.*, vol. 19, pp. 199-221, 2005.

[26] G. Chechik, M.J. Anderson, O. Bar-Yosef, E.D. Young, N. Tishby, and I. Nelken, "Reduction of information redundancy in the ascending auditory pathway," *Neuron*, vol. 51, pp. 359-368, 2006.

[27] H.P. Saal, S. Vijayakumar, and R.S. Johansson, "Information about complex fingertip parameters in individual human tactile afferent neurons," *J. Neurosci.*, vol. 29, pp. 8022-8031, 2009.

[28] E.T. Rolls, H.D. Critchley, J.V. Verhagen, and M. Kadohisa, "The representation of information about taste and odor in the orbitofrontal cortex," *Chemosens. Percept.*, vol. 3, pp. 16-33, 2010.

[29] F.D. Farfan, A.L. Albarracin, and C.J. Felice, "Electrophysiological characterization of texture information slip-resistance dependent in the rat vibrissal nerve," *BMC Neurosci.*, vol. 12:32, pp. 1-11, 2011.

[30] A. Borst and F.E. Theunissen, "Information theory and neural coding," *Nature Neurosci.*, vol. 2, pp. 947-957, 1999.

[31] S. Furukawa and J.C. Middlebrooks, "Cortical representation of auditory space: information-bearing features of spike patterns," *J. Neurosci.*, vol. 87, pp. 1749-1762, 2002.

[32] B.B. Averbeck, D.A. Crowe, M.V. Chafee, and A.P. Georgopoulos, "Neural activity in prefrontal cortex during copying geometrical shapes. II. Decoding shape segments from neural ensembles," *Exp. Brain Res.*, vol. 150, pp. 142-153, 2003.

[33] A.P. Georgopoulos and J.T. Massey, "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," *Exp. Brain Res.*, vol. 69, pp. 315-326, 1988.

[34] B. Sakitt, "Visual-motor efficiency (VME) and the information transmitted in visual-motor tasks," *Bull. Psychonom. Soc.*, vol. 16, pp. 329-332, 1980.

[35] J. Heller, J.A. Hertz, T.W. Kjaer, and B.J. Richmond, "Information flow and temporal coding in primate pattern vision," *J. Comput. Neurosci.*, vol. 2, pp. 175-193, 1995.

[36] L. Nizami, "Dynamic range relations for auditory primary afferents," *Hear. Res.*, vol. 208, pp. 26-46, 2005.

[37] L. Nizami, "Intensity-difference limens predicted from the click-evoked peripheral N_1: the mid-level hump and its implications for intensity encoding," *Math. Biosci.*, vol. 197, pp. 15-34, 2005.

Table 1. Details of papers that express information transmitted based on stimulus-evoked spike counts, and details of the respective confusion matrices. The papers need not represent the relative contributions, in quantity or quality, of the respective authors or laboratories. Some details were unclear, but efforts to contact authors of papers that were more than a few years old sometimes proved futile. "W" is width of spike-count category in units of "spikes". "C" is number of confusion-matrix columns (stimulus categories). "R" is number of confusion-matrix rows (spike-count categories). V means "Varied". That is, when the number of spike-count categories is fixed, the number of spikes differentiating each spike-count category from its neighbors is determined by the range of firing rates of the neuron; likewise, the latter determines the number of spike-count categories when the number of spikes differentiating each spike-count category is fixed. An extreme case was Gnadt and Breznen [17, p. 3528], who established a different number of response categories for each neuron, by dividing the range of firing rate of that neuron by the "number of distinguishable levels" for that neuron (their Table 2), established from modelling of its firing-rate behaviour.

Ref.	Stimuli	Species	Type and location of recording	W	C	R
8	Square-wave thrusts of blunt probe, of amplitudes differing by equal increments, to pressure receptors in skin	Cat, Monkey	Single-neuron recordings from slowly-adapting mechanoreceptors on hind limb	2	4-30	23
9	Square-wave thrusts of blunt probe, of amplitudes differing by equal increments, to pressure receptors in skin	Cat	Single-neuron recordings from slowly-adapting mechanoreceptors on the face, and from the slowly-adapting "relay neurons" to which they project, in parallel, in nucleus oralis and in nucleus caudalis	V	2, 4, 5, 8, 16	22
10	Square-wave thrusts of blunt probe, of amplitudes differing by equal increments, to pressure receptors in skin	Cat, Caiman	Single-neuron recordings from slowly-adapting mechanoreceptors on hind limb	1, 2	4-149	21
11	Square-wave thrusts, of amplitudes differing by equal increments, to the tooth	Cat	Single neurons to the canine tooth	1	32, 128	≤ 100 (mean=60)
7	Sucrose-in-NaCl solutions of different sucrose concentrations, flowing through the recording electrode	Blowfly	Single sugar receptors of the tarsal hairs of prothoracic legs	5	11	18-19
12	Visual fixation of 2-dimensional black-&-white checkerboard Walsh patterns	Macaque	Single neurons in inferior temporal cortex	V	128	12

13	Varied contrasts, of one cycle of a high-contrast sinusoidal grating of optimal orientation and spatial frequency, moving past an arbitrary point in the neuron's receptive field	Cat	Single-neuron recordings from 10 simple cells, 14 complex cells, all in area 17	2	14	23
14	Visual fixation of 2-dimensional black-&-white checkerboard Walsh patterns	Macaque	Single retinal ganglion cell fibers, single neurons in lateral geniculate nucleus, single complex cells in layers 2 and 3 of primary visual cortex, and single neurons in inferior temporal cortex	V	128	12
15	Visual fixation on the center or "corners" of each of 4 faces in total (monkey or human), usually in frontal view, but in some cases in profile view	Macaque	Single neurons responding selectively to faces, in the superior temporal sulcus	V	20	15
16	Visual fixation on the center of each of 20 faces in total (monkey or human), usually in frontal view, but in two cases in profile view	Macaque	Single neurons responding selectively to faces, in the superior temporal sulcus	V	20	15
17	Visual fixation of small spots, crosses, or X's that moved momentarily from one point to another	Macaque	Single neurons in lateral bank of the intraparietal sulcus (area LIP)	V	3, 37	V
18	Up to 9 different odorants, pure chemicals representing different odor classes, sniffed in an odor-discrimination task	Macaque	Single neurons in orbitofrontal cortex	V	8-9	15
19	Visual fixation on the center of each of 4 monkey faces and 3 human faces, each having one of 4 different facial expressions; and 2 shapes, each in 5 different colors	Macaque	Single neurons responding selectively to faces, in the inferior temporal cortex, including both banks of the superior temporal sulcus	V	3, 4	15
20	Kitten's meow, natural and laboratory-altered in its temporal and spectral properties	Cat	Single-unit spike trains, sorted from multi-unit recordings, in primary auditory cortex	2	9	V
21	Manipulation of rate and volume of mechanical inflation of the lungs with a mechanical ventilator	N. Z. white rabbit	Single neurons of the cervical vagus nerve	1	2 to ≥20	2 to ≥22
22	Forced sinusoidal whisking of whisker according to different vibration frequencies and amplitudes	Wistar rat	Neuronal clusters at barrel cortex	1	49	V
23	Free-field click trains having regular or irregular (Poisson-distributed) inter-click intervals	Marmoset	Single neurons in primary auditory cortex	1	19	100
24	Random dot textures moving through a square aperture in a field of stationary random dots	Macaque	Single neurons in extrastriate visual area MT	V	13, 17	V
25	Ferret: Broadband noise bursts from different virtual source directions, presented over headphones Cat: Bird songs, natural and laboratory-altered in their temporal and spectral properties, delivered through a sealed acoustic system to the tympanic membrane	Ferret Cat	Single neurons in primary auditory cortex	1	24 15	V (3-9 at least; their Fig. 2)
26	Bird songs, natural and laboratory-altered in their temporal and spectral properties, delivered through a sealed acoustic system to the tympanic membrane	Cat	Single neurons in inferior colliculus or in medial geniculate body or in primary auditory cortex	1	15	V (10 in their Fig 3)
27	Half-sinusoidal force change applied to fingertip, at 5 different possible force directions, with 3 possible surface curvatures, of the applied surface	Man	Single tactile afferents that terminated in the distal segment of the index, long, or ring finger	1	3-5	13-14
28	Up to 6 different tastants, pure chemicals representing different taste classes, delivered as a fixed volume of liquid to the monkey's mouth via syringe	Macaque	Single neurons in orbitofrontal cortex	V	6	15
29	Square-wave forced whisking of whisker over sandpapers at 3 distances (3 "skip-resistance levels")	Wistar rat	Compound action potential of transected deep vibrissal nerve innervating a vibrissal follicle (DELTA vibrissal nerve)	See note*	5	35*

* Response categories were 1 millivolt wide (see Farfan et al., Fig. 3). They computed $I_{t,WM}$ for only 3 of the 5 stimuli (see their Figs. 6 and 7).