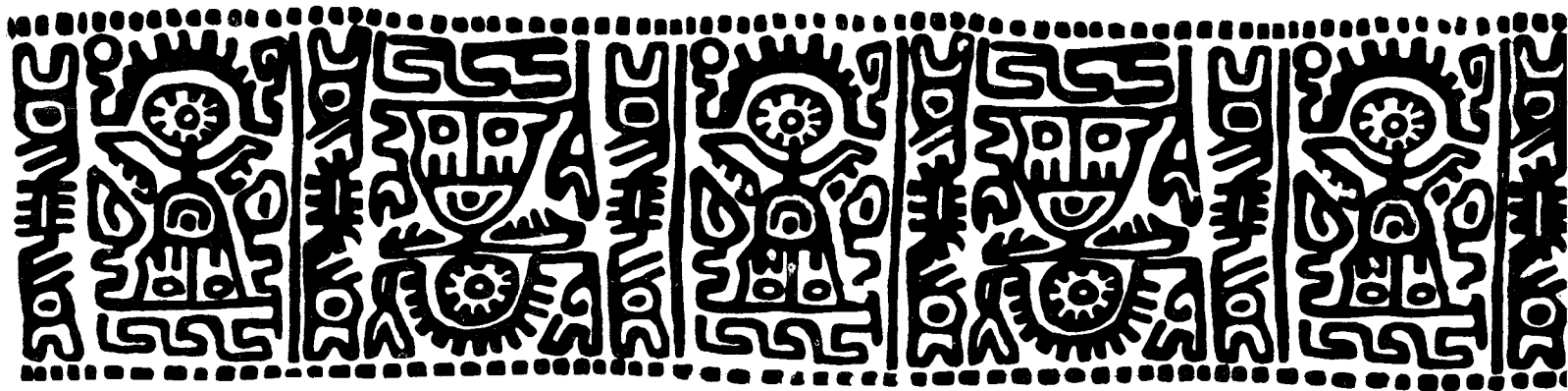


SIMPLE MODELS OF SENSORY TRANSMISSION

William J. McGill and Malvin C. Teich



CENTER FOR HUMAN INFORMATION PROCESSING

**UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CALIFORNIA 92093-0109**

**CHIP 132
JANUARY 1991**

Simple Models of Sensory Transmission

William J. McGill and Malvin C. Teich

Copyright © 1991 William J. McGill and Malvin C. Teich
Center for Human Information Processing
University of California, San Diego
La Jolla, CA 92093-0109

Simple Models of Sensory Transmission

William J. McGill and Malvin C. Teich

Puzzles abound in modern research on sensory thresholds.

The classical treatment of absolute visual detection devised by Hecht, Shlaer, and Pirenne (1942) still enjoys wide acceptance nearly 50 years after its introduction. It depicts the visual system as essentially noiseless, arguing that threshold phenomena are created by physical fluctuations in the energy content of the stimulus. The idea is attractively simple but at odds with much of what we know or suspect about the complexities of sensory processing.

Counting Every Quantum

A huge network exists, certainly more than 2 million neurons (Brown, 1965), connecting our eyes to the brain. If high level decision centers can access every photon absorbed in each retina, this transmission system must somehow be set up to count with perfect accuracy, permitting no variations of its own to intrude on threshold data.

Barlow's subsequent (1956) amendment of Hecht et al. did allow for an internal noise he labelled "dark light", but the interaction is seen as purely additive. Small detection errors remain small. They do not propagate into larger errors sowing greater confusion as they pass on up through the system.

Ideal Detection

Or consider the early Tanner and Swets (1954) treatment of ideal auditory detection. Why should this unique marriage of signal analysis and decision theory have retained its popularity in the front rank of auditory research for so many years? We knew from the beginning that ideal detection failed to account for either the shape or location of most auditory psychometric functions. It cannot, among other problems, explain the poor performance of the ear attempting to detect small intensity differences between pairs of phase-locked pure tones, or bursts of "frozen" noise, or indeed most other pairings of non-stochastic stimuli. Yet these inadequacies have not diminished the appeal of ideal detection in the slightest. Here again, analysis is limited entirely to the stimulus domain, bypassing any losses or transformations in the auditory system. The scheme finesses sensory mechanisms.

In Search of Simplicity

Evidently models that analyze thresholds by eliminating physiological parts of the process are enduringly popular in sensory psychology. This ought to be a mind-blowing contradiction, but in fact it is a fairly routine occurrence arising, we believe, from the operation of two unremarkable psychological principles.

First, a theory seldom falls out of favor just because it seems implausible. If it offers a simple, convenient calculus in which familiar findings can be interpreted with relative ease, it will continue in vogue even though everyone is convinced there is something wrong with it. This argument was originally promoted in Kuhn's (1957) landmark study of the history of astronomy. It was subsequently cast into good psychological form by Margolis (1987) writing on the role of context in scientific pattern recognition.

A second corollary principle is that people who preach a gospel of complexity generally end up talking to themselves, even when nearly everyone concedes they may well be basically correct.

Whenever a theory adopts a simplistic configuration in the face of complexity, it suggests not so much ignorance or obstinacy as a determined effort to outmaneuver difficulties by putting an ingenious twist on the facts. The very enormity of the reception

and transmission systems governing the major senses suggests they might provide a computer-like reconstruction of the external environment. The idea is at least as old as Aristotle. Our goal would then be to uncover stimulus dimensions decisive in human perception. An excellent example of the process is the concept of luminance in visual measurement, where light energy is scaled or weighted according to its effectiveness in stimulating the human eye (see Graham, 1965, chapter 13).

Classical Sensory Analysis

A bit less radical than disavowing the nervous system is the classical position of sensory psychophysics. Carefully controlled stimuli are used to explore the properties of receptor mechanisms. Transmission phenomena, if they exist, are ignored. Decision centers of the brain are thought capable of looking back through the long neural chain to the periphery so as to isolate events occurring in the sense organ. Hence, when we analyze a critical band, i.e. the masking region associated with a pure tone, it is almost a given that we are studying an activation pattern on the basilar membrane. In fact the electrophysiological evidence appears to be contrary. The critical band seems to be formed en route up the eighth nerve (see Ehret and Merzenich, 1985)

A highly instructive example of such problems can be found in an early attempt by Davis (see Stevens and Davis, 1938) to deduce

activation patterns in the inner ear by measuring psychophysical masking associated with pure tones. Later Békésy succeeded in making direct observations (Békésy and Rosenblith, 1951), successfully visualizing the vibratory motions of the basilar membrane. Receptor patterns were found to be much flatter, far more diffuse, than comparable psychophysical data. In the latter we appear to be examining sharply tuned responses of the entire system rather than receptor processes in isolation. It is possible, of course, that a transformation might be found taking Békésy's patterns into psychophysical data. To our knowledge no one has done it. Margolis (1987) explains why it is so difficult to give up on formulations yielding little progress. All of us hope devoutly for a psychophysical key that will unlock an accurate account of receptor processes. We keep searching for the key, and always it lies just beyond our grasp.

New Paradigms Needed

Time has run out on many of these reductionist ideas. The major senses are complex systems whose analysis begins in the realm of stimulus energy and ends in decision centers of the brain. The zones between reception and decision are mostly an uncharted communication region where little is known of what really happens. Sometimes messages are unaccountably lost. Some are contaminated with noise; effects that seem to multiply as events pass up through the system. Inevitably we must come to terms with all this

ambiguity. Ignoring critical elements of sensory mechanisms has led to half-explanations, analytical devices that do not quite work, and not a little frustration. It is time, we think, to move to a new level of discourse.

A serious attempt at coming to terms with ambiguity began with the emergence of modern versions of signal detection theory in both vision and audition. (see Geisler, 1989; Green and Swets, 1988 /1966). The development of a new and iconoclastic sensory paradigm took a great leap forward when Laming's book, Sensory Analysis (1986) appeared. Laming jettisoned much of the traditional apparatus of stimulus and receptor analysis in favor of a few highly sophisticated "black boxes", said to exist somewhere between receptor and brain. It is not our intention to review or restate Laming's work here. We have attempted that elsewhere (McGill and Teich, 1989); but some discussion of Laming is required as a preliminary to understanding where contemporary study of sensory processes seems to be heading.

The centerpiece of Laming's paradigm is a mechanism he calls a differential coupler. In effect, Laming argued, sensory systems somehow manage to neutralize incoming stimulation if a continuous background remains at constant intensity. Increments or decrements are then detected as transients away from this neutral condition. The important point is that the system behaves as though it were a null detector. Laming (1986) offered a battery of arguments in

support of this view.

Differential coupling is equivalent to the creation of an adaptation level (Helson, 1947). In fact, we prefer to speak of the phenomenon in this more traditional psychological way, but minor nuances of language should not divert us from the importance of Laming's ideas. He has managed to put an entirely new face on the data of masking, increment detection, and intensity discrimination, uncovering regularities no one else had noticed.

We begin our study of stochastic sensory networks with Laming's differential coupler because, although our own work started from entirely different premises, and although we remain devoted to finding a precisely defined stochastic process governing sensory information transmission, Laming's footprints are all over the network mechanism we have developed. One way to describe what we have done is to say that it appears to be an elementary realization of Laming's differential coupler.

Stochastic Networks

Visual and auditory psychophysics deal principally with stimuli and receptor mechanisms; visual and auditory detection theory, with stimuli and decision rules. We propose to add, (perhaps "insert" would be a more accurate characterization), a third level of analysis concentrating on information transmission, as sketched in

Fig. 1.

FIG #1

Little doubt exists as to the necessity of a step like this given the current state of sensory analysis, but we are also acutely conscious of the dangers of preaching complexity. How can such mechanisms be kept simple or even manageable? Everyone agrees that transmission phenomena are probably important, but no one other than Laming (1986) has dared offer a comprehensive view of how they operate. Even Laming's differential coupler is a black box with elaborate functional properties, whereas we are searching for something far more primitive.

Signal detection encountered initial consumer resistance because auditory researchers did not at first see the need for a second level of analysis distinct from the study of receptor mechanisms. What chance exists for a third tier of analysis, especially one proposing to take us into the depths of an uncharted sensory communication system?

Conflicts of this sort between the demands of realism and the requirements of clarity are typically resolved by turning to a stochastic process. A network, even one seen as too complex to permit tracing individual events, is nonetheless decipherable. We need not know exactly how the network is organized or exactly how events move from stage to stage. The stochastic process operates

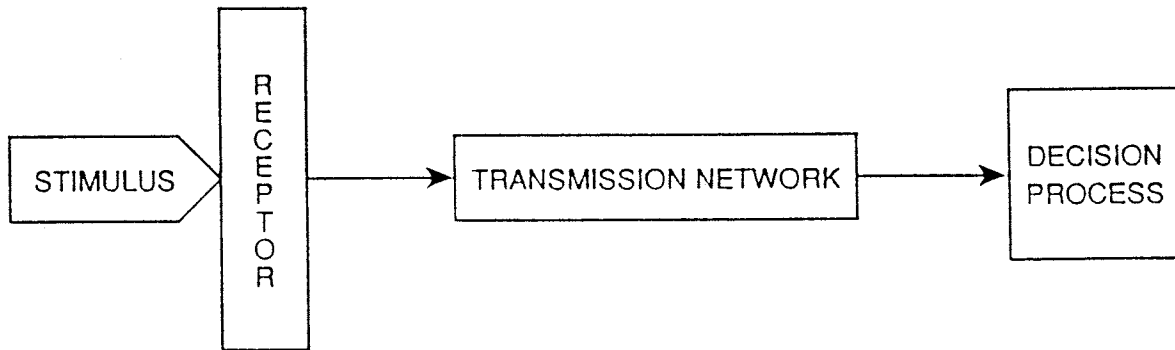


Figure 1: Flow diagram of a sensory process. Stimulus and receptor mechanisms constitute traditional realm of psychophysics. Stimulus and decision processes form domain of signal detection theory. Ordinarily transmission network linking receptor to decision is ignored. This network is focus of present paper.

at a different level of discourse in which transmission is represented as repetition of a probability operator. If we ever hit upon the right form for the operator, we may be able to construct the network functionally without exact knowledge of its interconnections. The resulting description is couched in probabilities of outcomes rather than precise tracings of events from input to output. The important point is that a simple transmission operator, applied repeatedly, generates a probability space requiring a very complicated network for its realization.

Simple Operators - Complex Networks

A stimulus activates a sense organ. The energy content of the stimulus is transferred to receptor cells either through chemical absorption or some form of physical resonance. However it happens, and whatever its efficiency, the reception process creates a charged condition generating a discharge of primary neurons just behind the receptor. The critical information marking the occurrence of a stimulus is contained in this discharge, but ordinarily it occurs against a backdrop of unanalyzed, hence "spontaneous", noise events. These are incorporated into the discharge leaving no way to distinguish signal from noise at the rear end of the system. When a decision is required on whether or not a stimulus actually occurred, it is made by counting events at the output. A small count would indicate that nothing new happened whereas a large one suggests the contrary. The formulation is the

traditional signal-noise decision option of detection theory most clearly exemplified in Barlow's (1956) refinement of the Hecht, Shlaer, and Pirenne (1942) model of visual energy detection.

Our first attempt (McGill, 1967; Teich and McGill, 1976) to introduce a transmission link into this picture required one additional step. The connecting link was said to be a Poisson transmission line obtained by smearing together all the information-bearing pathways leading back from a receptor. Momentary stimulus intensity was viewed as driving this Poisson rate. Counting distributions were worked out (McGill, 1967) and converted to detection laws for the transmission network driven by various well-defined inputs. These expectations were then compared with data.

Counting distributions for sinewave signals in wideband noise proved interesting because they produced results identical to those obtained in studies of laser energy output (Peřina, 1967; Teich and McGill, 1976). Their value as psychophysical tools was less apparent. Poisson transmission did lead to discovery of a near-miss to Weber's law in pure tone intensity discrimination (McGill and Goldberg, 1968a&b) as well as to the introduction of Neyman's Type A distribution in visual detection (see Teich, Prucnal, Vannucci, Breton and McGill, 1982), but the aggregate of our first attempt to formulate a simple transmission mechanism was not much better than what it sought to replace (McGill and Teich, 1990).

The main problem was that Poisson transmission creates a remarkably passive link. It fails to impose its own stamp on information passing through it or to alter that information significantly. Hence while the link behaved well where auditory detection theory had succeeded previously, Poisson transmission tended to fail where detection theory had previously failed. Something more sophisticated was needed, a more complex operator interacting with sensory messages as they pass to higher centers.

To achieve such realism we should first acknowledge that the receptor discharge and its noise contaminants are probably not directly available to decision centers. These initial events are the first stages of an intricately constructed chain. The latter does not passively transmit. It operates on its input information. A stimulus passes into this network. If all goes well, progeny emerge at the terminus ready for counting. We imagine the network to consist of stages arrayed in sequence, each one more or less like the others. Such an arrangement could easily amplify incoming signals. The network would be adequately represented by a model of a single stage together with an agreement that the number of stages is sizable (although not necessarily infinite). This takes us into near asymptotic effects produced by replicating an operator, for example the one depicted in Fig.2.

FIG.2

A transmission network is said to consist of a sequence of r such stages where, as we have said, r is thought to be large. At

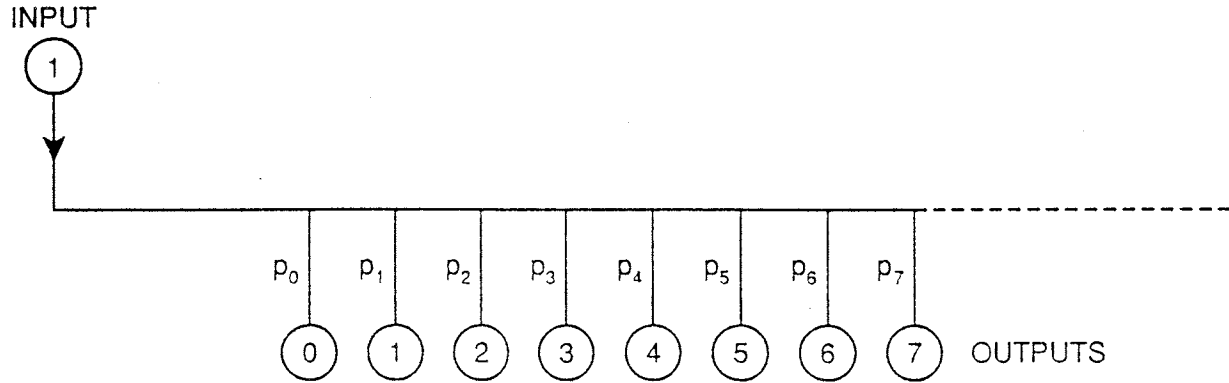


Figure 2: Markov operator diagram of stochastic transmission. Each input event is converted to one of several outputs with probabilities as indicated. A possible outcome is that an input may completely disappear (probability p_0). Multiplication occurs whenever output number equals or exceeds 2. A single input event starts process. Network then generates probability distribution of progeny at output of final stage.

each stage an input is transmitted only if it produces at least one output. To move a single message event through the network a run of successful transitions of the r stages is required. The network is deemed to amplify if progeny tend to increase as they pass from stage to stage, but passage is uncertain because there is also a probability that a given message may die out at any point. We characterize such a network as stochastic.

An important baseline condition is a sequence of stages that repeatedly generates a single output event (on the average) for each corresponding input. This is done by counterbalancing expected growth against expected loss at each stage.

It would seem that messages should pass through such a network more or less intact. But when growth and decay are perfectly balanced in a network that is sufficiently long, virtually all messages disappear in transit. This is a familiar result easily proved. It furnishes the key to a stochastic version of adaptation level or differential coupling. Long chains lose virtually all messages unless some amplification, however small, occurs in transmission. This implies further that a single input event subjected to stage-by-stage multiplication will eventually give birth to an unbounded number of progeny at the end of a long chain of multiplications.

At first thought no room at all seems visible between these

polarities: complete blockage or uncontrolled growth as the average birth rate passes from just under unity to a number just slightly in excess of unity. There is an easy way out of the dilemma but the very fact of its existence offers a sober warning about the complexities of branching networks. Moreover, any uncertainty encountered as information passes through a given stage introduces noise into the message.

A sequence of noisy stages multiplies uncertainty with each added stage. We cannot speak of such transition-noise as unrelated or additive. It is a built-in property of stochastic networks, varying in level with the length of the network. Hence, we expect to be contending with at least two different kinds of noise: 1) additive noise generated by unrelated events outside the transmission system that become incorporated into its messages; and 2) multiplicative noise created by the hazards of message transmission. Barlow's "dark light" (1956) is an excellent example of additive noise. Weber's law is typical of effects expected from multiplicative noise, as Laming has pointed out (1986 p. 71-75).

Simplest Network

The transform depicted in Fig. 2 is called a Markov operator. It is applied indiscriminately to every message event appearing at the input of each successive stage. A long chain of such stages brings on asymptotic properties of the Markov process. To

characterize the entire chain we need to work out these asymptotic properties. They are not limited to passive transmission. For example, the chain will amplify messages if the output of each stage produces an average of slightly more than one event for each input.

Even this representation is much too simple-minded. A realistic transmission mechanism would probably involve series-parallel chains in which different regions featured different operations. We intend to stay with the lowest level of function. Our goal here is to work out the properties of a simple neurological amplifier.

To put flesh on these arguments we first consider an extremely primitive network. It is not one we would ever use, but it is worth study because it was employed by Hecht et al. (1942) to explain the loss of photons moving from a measurement locus at the surface of the cornea to eventual absorption in retinal rods. The process is completely passive, producing only random deletion of message elements passing through it. This means that all transition probabilities other than p_0 or p_1 in Fig.2 are set equal to zero. In particular, multiplication is excluded. An input event can either reproduce itself or disappear as it moves through each stage. The Markov operator is diagramed in Fig.3.

Fig. 3

Transition probability from state 1 at any given stage to

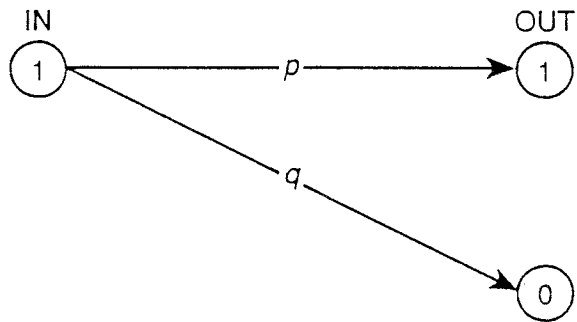


Figure 3: Markov diagram of a single stage in transmission network in which outputs are either transmitted or disappear at random. Transition from state 1 to state 1 implies that an input event is transmitted to next stage. Transition from state 1 to state 0 means that input is lost. No transmissions occur out of zero state. Any event entering this state is lost, terminating passage through network.

state 1 at the next stage is fixed at p . The corresponding dropout probability (transition from state 1 to state zero) is then also fixed and labeled q , where:

$$p + q = 1.$$

Evidently these are conditional probabilities characterizing the statistics of passage through each stage. Successive stages are independent and the process is multiplicative; (each step triggers the next step.) Accordingly, the probability generating function at any stage (see the Appendix at the end of the paper) is obtained by applying the operator to the generating function of the prior stage. (See Feller (1957), or Bharucha-Reid (1960) pp. 19-27 for a discussion of generating functions and their use in analyzing branching processes). Stage-to-stage linkages constrain the generating functions. Such constraints enable us to unravel the output statistics of the final stage, and thus to analyze the behavior of the entire network.

Start with a single input event. At stage 1 the transition operator generates:

$$G(s) = q + ps, \tag{1}$$

so that at stage 2 a fraction q of the input messages is expected to drop out. Repeated applications of the operator produce at the r th stage:

$$\begin{aligned}
 G_r(s) &= q + p G_{r-1}(s), \\
 &= 1 + p(G_{r-1}(s) - 1), \\
 \frac{G_r(s) - 1}{G_{r-1}(s) - 1} &= p.
 \end{aligned}$$

(2)

Multiplying this out over all r stages, and remembering that the initial condition is a single input (ie $G_0(s) = s$), we see that:

$$\frac{G_r(s) - 1}{G_{r-1}(s) - 1} \cdot \frac{G_{r-1}(s) - 1}{G_{r-2}(s) - 1} \cdots \frac{G_2(s) - 1}{G(s) - 1} \cdot \frac{G(s) - 1}{s - 1} = p^r.$$

(3)

Accordingly, the generating function for the output of the entire network must be:

$$G_r(s) = (1 - p^r) + p^r s,$$

(4)

where now the iterative expression has been replaced by unconditional probabilities at the final stage. Output proves to have only two states: zero and unity. Probability of successful passage (state 1 output of the final stage) is given by the transit probability at any given stage raised to a power set by the length of the network. As the latter lengthens, this state 1 output probability moves inexorably toward zero. Nearly all messages end up in state zero--lost along the way. This is, of course, not at

all surprising, but iterative restrictions on the probability generating function can be stated for far more complicated branching networks in which events multiply and outcomes are not nearly so obvious. These more challenging cases underscore the usefulness of generating functions for attacking this kind of problem.

Finally, a very easy extrapolation exists to a related process in continuous time. Suppose the transition at any stage requires a brief interval of time Δt . If the latter is very small, there is little likelihood of a change of state during Δt . (The only change possible would be a sudden transition to state zero.) Hence as Δt decreases, q must also decrease proportionally. This stable ratio is ordinarily expressed as a fixed loss parameter μ , where:

$$\lim_{\Delta t \rightarrow 0} \frac{q}{\Delta t} = \mu. \quad (5)$$

The network probability becomes:

$$p^r = \left(1 - \frac{\mu r \Delta t}{r}\right)^r, \quad (6)$$

and $r \Delta t$ is the transit time for the full network:

Accordingly:

$$\lim_{\substack{r \rightarrow \infty \\ \Delta t \rightarrow 0}} (p^r) = e^{-\mu t} \quad (7)$$

Length of the network is measured by its transit time t , regarded here as fixed, while loss or dropout probability at any given stage is measured by μ . Evidently, μ is not a probability since it may take on any positive value, whereas the limit of the state 1 output established in Eq. (7) is a probability. As in discrete cases, this probability moves toward zero when the network lengthens (t increases), or when the dropout tendency grows large (μ increases).

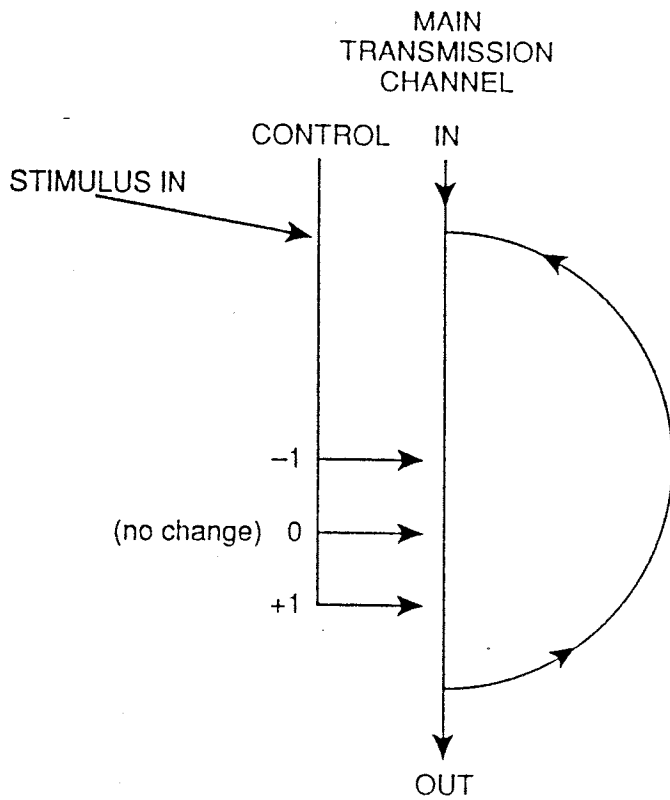
Amplifier Networks

Far more sophisticated is a closely related operator illustrated in Fig.4.

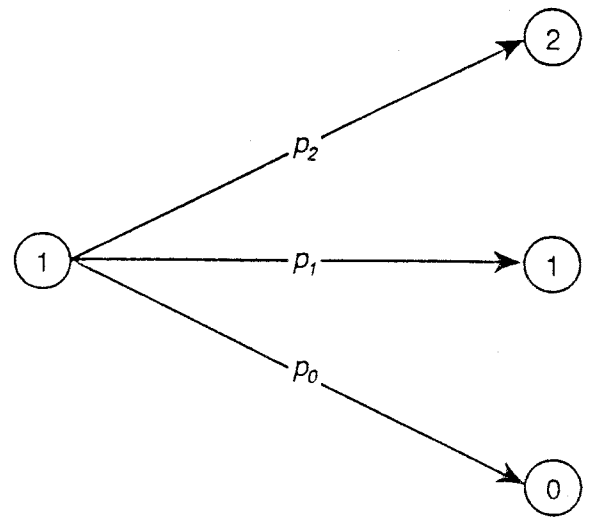
FIG 4

A network constructed from such stages will tend to amplify its input. Single events entering a given stage may be converted into pairs of events as they feed into the next stage. We say "may" because a possibility exists that message events may be wiped out in transit. There is also no requirement to stay with pairwise multiplication. Many other possibilities exist. Pairwise multiplication leads to a classical result in branching processes, so it is a good place to start.

How the various possibilities play themselves out will be determined by values assigned to the transition probabilities in Fig. 4. The terms p_0 , p_1 , and p_2 are conditional probabilities



AMPLIFIER STAGE



OPERATOR DIAGRAM

Figure 4: Simple amplifier network. At left is functional diagram of typical stage. Passage of message events moving down main transmission channel is modified by control mechanism on far left. Stimulus intensity sets configuration of control. Feedback may occur within a single neurological site or these little amplifiers may be arrayed in sequence. Diagram on right displays Markov operator characterizing this mechanism. Operator shows allowable transitions for events passing along main channel and probabilities associated with each possible transition.

governing the possible transitions for each message event moving through each stage. Hence:

$$p_0 + p_1 + p_2 = 1.$$

Whenever $p_2 > p_0$ the network will amplify. This conclusion follows because stage-to-stage birthrates are all greater than unity. The generating function of the operator in Fig. 4 is:

$$G(s) = p_0 + p_1 s + p_2 s^2.$$

This is the stochastic description of a single stage characterized by pairwise multiplication (see Appendix). We wish to find the output counting distribution for the entire network (i.e. the output of stage r). Since sequential networks are multiplicative, the best way to proceed is to use the multiplicative property to put an iterative restriction on the generating function for stage r , just as we did with random deletion:

$$G_r(s) = p_0 + p_1 G_{r-1}(s) + p_2 G_{r-1}^2(s).$$

(9)

Now, however, no easy way can be found to multiply the process out, isolating a single probability. But there is an easy way to construct a time derivative. First, subtract $G_{r-1}(s)$ from both sides:

$$G_r(s) - G_{r-1}(s) = p_0 - (p_0 + p_2) G_{r-1}(s) + p_2 G_{r-1}^2(s).$$

(10)

Next set $p_0 = \mu \Delta t$, $p_2 = \lambda \Delta t$, and let stages r and $r-1$ differ by Δt . (Notice that the coefficients of the polynomial on the right hand side sum to zero). Now convert the time difference between successive stages into a derivative:

$$\lim_{\substack{r \rightarrow \infty \\ \Delta t \rightarrow 0}} \left(\frac{G_r(s) - G_{r-1}(s)}{\Delta t} \right) = \lambda (\mu/\lambda - G_{r-1}(s)) (1 - G_{r-1}(s)),$$

(11)

$$G'_t(s) = \lambda (\mu/\lambda - G_t(s)) (1 - G_t(s)).$$

(12)

The derivative is taken with respect to time. This generating function depicts a process requiring time t to transit all stages. Our form for Eq (12) shows the polynomial factored so as to display its two roots. If we integrate Eq. (12), we then have a generating function at the output of stage r free of iterative restrictions. Moreover, with that problem solved it becomes a simple matter to work out the counting distribution from which the generating function was constructed. The solution of the differential equation is tabled in most standard tables (for example, the Chemical Rubber Handbook (1954), #47, p.250). Before addressing it, we first review Eq. (12) to assess its full meaning.

The derivative is a quadratic with two distinct roots. Since

its coefficients sum to zero, a root will be found where $G_t(s)=1$. This implies that as Δt decreases, each input making its transit of a given stage tends to resist change (in other words, to remain in state 1). The probability of a change up or down becomes vanishingly small in the limiting process that converts discrete stages to continuous time. When the coefficients of Eq (10) sum to zero, a root appears at $G_t(s) = 1$ in Eq. (12) reflecting such resistance.

A second root emerges when $G_t(s)$ reaches an asymptote as t increases without limit. In that case $G_t'(s) \rightarrow 0$. This second root proves to be the value of $G_t(s)$ at the asymptote, namely:

$$G_\infty(s) = \mu/\lambda$$

The limiting generating function depicts the outcome of a transmission process operating over infinite time. There is a spike of probability in state zero but nothing corresponding to any other finite state (since only the coefficient of s^0 survives at the limit). Hence, there exists a significant probability of an infinite count, a significant probability of a zero count, and essentially nothing in between.

From the generating function at infinite time we have:

$$P_\infty(0) = \mu/\lambda$$

for the limiting probability of a state zero output (message lost). In branching processes this probability is known as an *extinction probability*. Evidently when the loss parameter μ is larger than the gain parameter λ , the extinction probability must remain fixed at unity. Nearly all interesting cases will then arise when $\mu \leq \lambda$.

We see that the operator in Fig. 4 has two different asymptotic conditions. To establish the derivative in Eq. (12) r (the number of stages) goes to infinity while t (the network transit time) remains finite. A further limit for the generating function is obtained by letting t go to infinity. At this second level, the process becomes unmanageable. Every non-zero output is found to be infinite, whereas at the intermediate limit with t finite, conventional counting distributions are encountered over the full range of positive integers. Existence of two different limits explains many seeming contradictions that appear as we analyze network behavior. For example, when $\lambda = \mu$ in a perfectly balanced network, the extinction probability is found to be unity. This suggests that no messages can pass through the entire chain. Yet when Eq (12) is integrated in these same circumstances, a conventional probability distribution over nonzero counts is found. The latter case reflects an intermediate limit (r infinite; finite t) whereas the extinction probability is an asymptotic limit as t itself becomes infinite. It should then be the case that probability of a zero count for finite t rises systematically up to

the limiting extinction probability as t increases. This is exactly what happens.

Adaptation Level

Output distributions for the full network are obtained by integrating Eq. (12) and then extracting the counting distribution from the generating function. The original input is a single message event at the first stage, The network counting distribution spreads out over many values because multiplicative noise, generated in transmission, introduces variance with each added stage of processing. This noise variance becomes a signature, stamping the network's configuration on all information passing through it.

First consider the solution of the differential equation, Eq. (12), when expected gains just balance losses, i.e. when $\mu = \lambda$. We have:

$$\frac{G'_t(s)}{(1-G_t(s))^2} = \lambda \quad (12a)$$

Integrating both sides with respect to time produces:

$$\frac{1}{1-G_t(s)} \Bigg|_s^{G_t(s)} = \lambda t.$$

The limits on the left hand side establish that the generating function, $G_z(s)$, becomes s as $\tau \Rightarrow 0$, and $G_t(s)$ as $\tau = t$. We have:

$$\begin{aligned} \frac{1}{1-G_t(s)} &= \lambda t + \frac{1}{1-s}, \\ G_t(s) &= 1 + \frac{s-1}{1-\lambda t(s-1)}. \end{aligned} \quad (13)$$

This is the generating function of a Markov process with gains and losses precisely balanced, and operating in continuous time. To extract the counting distribution we need a way to expand the right hand side of Eq. (13) in powers of s . The coefficient of s^k will then be the probability, $p_t(k)$, that network output consists of exactly k message events. The expansion is easy. We show one way to do it in the Appendix. The output counting distribution at adaptation level is found to be:

$$\begin{aligned} p_t(0) &= \left(\frac{\lambda t}{1+\lambda t} \right), \\ p_t(k) &= \frac{(\lambda t)^{k-1}}{(1+\lambda t)^{k+1}}; \quad k=1,2,3,\dots,\infty. \end{aligned} \quad (14)$$

This distribution is actually much simpler than it looks. There is a large "spike" of probability in state zero as we have come to expect from transmission mechanisms with a propensity for losing information along the way. The distribution over non-zero states is not very significant but nonetheless there. Its form is geometric, reflecting difficulty in achieving long runs of multiplications as progeny pass from stage to stage in Fig. 4. The

distribution of surviving messages has a characteristic geometric or exponential form in such circumstances.

If the gain/loss parameter λ can be related somehow to the intensity of a continuous background, network output will obey Weber's law whenever background intensity remains constant through repeated stimulus trials. Multiplicative internal noise created as a steady background is neutralized, also regulates incremental detectability.

We are not invoking a scheme, à la Hecht et al. (1942), in which transmission parameters are fixed while stimulus intensity sets the number of triggering events appearing at the first stage. Instead, we conjecture a network in which stimulus properties determine flexible parameters controlling passage of information at every stage. The action resembles a grid in an electronic amplifier. Such effects are probably common in the nervous system, but it is important to pause over this one, and to underline it. The change in perspective reflects a genuinely new outlook on transmission of information from sense organs to higher brain centers. It was this change in perspective that led Laming (1986) to his concept of differential coupling.

Negligible transmission occurs at adaptation level with λ (gain) and μ (loss) precisely equal at every processing stage. The mean output is unity, while the variance is $2\lambda t$. These points

are easily verified by computing the mean and variance of the counting distribution in Eq. (14) (see Appendix).

Mean output at adaptation level does not vary with background intensity, but variance (noise level) does. A steady background will set up two opposed processes, one attempting to amplify the input, the other trying to knock it out. Multiplicative noise generated by such a Markov operator is not quite zero on the average nor is it gaussian as Laming (1986) prescribed, but its effects on detectability are similar to those suggested by Laming for the differential coupler. We would not be amiss in arguing that a differential coupler is probably a balanced amplifier network similar to Fig 4. In stochastic processes it would be called a "birth-death" process (see Bharucha-Reid, 1960, pp.88-91; Feller, 1957, pp. 407-411).

Adaptation level is established at a background intensity determined by the size of the equalized parameters. The network rests in precise balance at this adaptation level. Very little will then get through. Counting variance at network output will increase with background intensity, making transients of fixed size harder to detect as background level rises. Moreover, we now have a solid analytical method leading directly to the counting distribution when increments are added to the background. We need only establish a relation linking λ and μ , the gain and loss parameters, to stimulus intensity.

Output Counting Distribution

What is the output counting distribution of an amplifier network whose discrete form is as diagramed in Fig. 4? We no longer restrict consideration to adaptation level. The parameters have arbitrary values. In this circumstance the differential equation of the transmission process was shown earlier to be:

$$\frac{G'_\tau(s)}{(\mu/\lambda - G_\tau(s))(1 - G_\tau(s))} = \lambda \quad (12b)$$

where τ is the (time) variable of integration. Integration is carried out between limits $\tau = 0$ and $\tau = t$. When $\tau = 0$, $G_0(s) = s$, ie the process begins with a single input.

The solution can be found in most tables of integrals. Our notation produces:

$$G_t(s) = \frac{\mu/\lambda - \left(\frac{\mu/\lambda - s}{1-s}\right)e^{-(\lambda-\mu)t}}{1 - \left(\frac{\mu/\lambda - s}{1-s}\right)e^{-(\lambda-\mu)t}} \quad (15)$$

As $\lambda \rightarrow \mu$, this generating function is identical to that for a balanced network in Eq. (13). Hence, whatever the values of λ and μ , the generating function in Eq. (15) describes a continuous version of the process diagramed in Fig. 4.

The counting distribution is not quite so easy to extract as its counterpart in a precisely balanced network. Nonetheless, the distribution is quite well-known in stochastic processes. It is given in many sources (see for example Bharucha-Reid (1960) p.88). We do not present it here because we now propose to introduce an additional simplification.

A big problem is created by the exponential terms in Eq. (15). They signal that average network output grows exponentially as the difference between gain and loss parameters increases. This is delightful when attempting to detect weak signals against a steady background, but it is potentially disastrous as signal intensity increases because a deluge of activity might easily flood the network and paralyze it. Ideally we wish outputs to grow as a fractional power function of stimulus intensity. Because of the adaptation level property, even linear growth would be acceptable. In any event a transformation is needed in order to curb exponential growth at high intensities.

Log Transform

The average output of the mechanism in Fig. 4 and Eq (15) can be calculated without elaborate ritual. The birthrate at any stage is:

$$m = (1+p_2-p_0).$$

Each input either rides through, or is increased to two with probability p_2 , or is wiped out with probability p_0 . At the end of r stages, the birthrate through the entire network should be:

$$m_r = (1 + p_2 - p_0)^r.$$

Now apply the limiting process as $\Delta t \rightarrow 0$ and $r \rightarrow \infty$:

$$\lim(m_r) \rightarrow \left(1 + \frac{(\lambda - \mu) r \Delta t}{r}\right)^r, \\ m_t = e^{(\lambda - \mu) t}.$$

(16)

It is this exponential growth that concerns us. We might try to alter it by changing the pattern of probabilities in the Markov operator, but the form of the latter tends to enforce exponential growth whenever multiplication is allowed to happen. A more practical solution is to put stimulus intensity through a simple transformation prior to setting the network parameters λ and μ .

Once transformations are contemplated the term "stimulus intensity" becomes too vague. We need to say what is transformed into what.

Earlier in this paper we spoke of an energy exchange in the sense organ as if it were the primal event of the stimulus message. This is certainly a common view in sensory psychology and neurophysiology. Among vision researchers, for example, the

principle that light energy carries the key stimulus information is widely accepted. Generally, the same is true in audition although there are significant exceptions. Green (1960b) and Green and Swets (1988/1966) based ideal detection of sinusoids in gaussian noise on complete reconstruction of the stimulus waveform, ending up with amplitude (proportional to the square root of energy) as the key variable. Jeffress (1964,1968), and Laming (1986), also argued for amplitude as the conveyor of intensity information in auditory perception.

Readers encountering these arguments for the first time may be forgiven expressions of bafflement. Energy and amplitude are simple functions of each other. If it is known how information flows with one, it is then easy to calculate how it flows with the other. What is the argument all about?

Typically we look for standard outcomes in detection data: gaussian integrals, specific values for Weber fractions, I.T laws, sometimes even negative masking. Which version of the intensity variable produces them? If expectations on outcomes are confirmed with stimulus energy as the variable, amplitude will not work, and vice versa. So a real question does exist.

It is brought into sharp focus by the task of converting the generating function in Eq. (15) into a specific counting distribution determined not by the parameters λ , μ , and t but by

some still-unspecified stimulus property expressing intensity.

Linear growth can be achieved in the stimulus domain if we let:

$$\ln (1+ax) = (\lambda-\mu) t, \quad (17)$$

where now x is the intensity variable, a is a constant matching intensity units on the left hand side of Eq. (17) with values of the parameters on the right hand side. To establish linearity the log transform is introduced into Eq. (16):

$$\begin{aligned} m_t &= e^{(\lambda-\mu) t}, \\ m_t &= e^{\ln(1+ax)}, \\ m_t &= 1+ax. \end{aligned} \quad (18)$$

Mean network output now increases linearly with incremental intensity. Obviously other transforms are applicable and might be considered.

In Eq. (17) the time parameter t is the transit time for messages passing through the network, while $(\lambda - \mu)$ is the (average) net rate at which new message events are generated. In amplifier networks, the product $(\lambda - \mu) \cdot t$ turns out to be the log of the mean number of messages or counts observed at the network terminus t seconds following stimulation (see Eq. 18). But

individual events will rarely consume exactly the same amount of time in passing the network. An observer at the output sees a short discharge measuring the intensity of a brief stimulus above adaptation level. Evidently this discharge takes place over time. Some message events lead; others straggle. The observer must be able to accumulate a record of events over an interval long enough to pick up the bulk of the discharge.

The nub of this argument is that the right hand side of Eq. (17) is the log of a count of message events. If an energy exchange occurs at the sense organ, the size of the discharge at network output during an observation time t matched to the average transit time, is a measure of stimulus energy. So the average number of messages corresponding to an intensity x in Eq. (17) implies that x is measured in energy units.

This finding does not suggest that energy is always the key stimulus property. An amplifier network adapted to a background intensity considerably above absolute threshold will exhibit constant detectability for increments whenever the latter stand in fixed ratio to the amplitude of the background (see, later, Eq 32).

Linear Output

If the log transform in Eq. (17) is applied to the amplifier network generating function:

$$G_t(s) = \frac{\mu/\lambda - \left(\frac{\mu/\lambda - s}{1-s} \right) e^{-(\lambda-\mu)t}}{1 - \left(\frac{\mu/\lambda - s}{1-s} \right) e^{-(\lambda-\mu)t}}, \quad (15)$$

so as to limit exponential growth of information and make output linear with stimulus energy, we obtain:

$$G_x(s) = \frac{\pi(x) - \left(\frac{\pi(x) - s}{1-s} \right) \left(\frac{1}{1+ax} \right)}{1 - \left(\frac{\pi(x) - s}{1-s} \right) \left(\frac{1}{1+ax} \right)}. \quad (19)$$

Here $\pi(x) = \mu/\lambda$ is the extinction probability corresponding to an incremental intensity x above adaptation level.

The parameters μ and λ are then not fixed within the network but set by the driving force of a particular intensity. Extinction probability is at or very near unity in the vicinity of adaptation level, dropping off to zero as incremental intensity grows. Without difficulty we find:

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1-\gamma(x)(s-1)}. \quad (20)$$

Introduced here is a new variable:

$$\gamma(x) = \frac{ax}{1-\pi(x)},$$

a modified intensity, corrected for the probability of survival. The generating function in Eq (20) resembles our earlier expression for a balanced network (Eq (13)). There is a spike of probability at the zero state representing defacto extinction, messages failing to make it through the network, and a distribution of survivors extending across all the positive integers.

These points are readily verified when the network counting distribution (incorporating the log transform in Eq (17)) is extracted from the generating function (see Appendix). We find:

$$\begin{aligned} P_x(0) &= \left[\frac{\pi(x) \gamma(x)}{1+\gamma(x)} \right], \\ P_x(k) &= [1-P_x(0)] \left[\frac{1}{1+\gamma(x)} \right] \left[\frac{\gamma(x)}{1+\gamma(x)} \right]^{k-1}; \end{aligned} \tag{21}$$

$$\begin{aligned} \text{where: } & k = 1, 2, 3, \dots, \infty; \\ & \pi(x) = \mu_x / \lambda_x; \\ & \gamma(x) = ax / (1 - \pi(x)). \end{aligned}$$

The output probability, $p_x(k)$ is again geometric and weighted by the probability of survival. The geometric component of Eq (21) is sometimes called a Bose-Einstein distribution (see Feller, 1957, p.59) because it turns up as a limit of the Bose-Einstein statistics in statistical mechanics. The mean-value parameter (in this case $\gamma(x)$) is typically a measurement corresponding to some expected number of events. Hence, the Bose-Einstein distribution

is a first cousin of the Poisson distribution.

An important consequence of amplification in a transmission network is that it alters the internal noise. Branching generates larger variance than would be found in devices passing information without multiplication. These effects are not subtle. They produce modifications in network output away from Poisson form and toward Bose-Einstein (geometric) form. They also affect the detectability of signals. Hence, if amplifier networks do in fact mediate transmission between sense organ and brain, they are likely to have a powerful impact on the information passing through them. Psychophysical data would not then be simple depictions of receptor processes but something more complicated. The caveat is hardly surprising or unexpected. What is surprising is that our analysis suggests the complexities may be easily managed.

Ultra-Linear Output

To pursue manageability we consider the close resemblance between the generating function in Eq. (20) for a linear amplifier network and the analogous expression in Eq (13) at adaptation level. Start with Eq (20) involving amplification:

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1-\gamma(x)(s-1)}$$

(20)

Now let x , the incremental intensity, go to zero. This moves the generating function to adaptation level. We need:

$$\lim_{x \rightarrow 0} \gamma(x) = \frac{ax}{1-\pi(x)} = \gamma.$$

A limit develops because as x goes to zero, extinction probability approaches unity. We let γ symbolize this limit at adaptation level. Then:

$$\gamma(x) = \gamma + ax.$$

Accordingly, as x goes to zero we obtain:

$$G_x(s) = 1 + \frac{(s-1)}{1-\gamma(s-1)},$$

(22)

which, except for notation, is Eq. (13). Our so-called "ultra-linear" restriction defines an extinction probability at every intensity above adaptation level:

$$\begin{aligned} \gamma(x) &= \gamma + ax, \\ \gamma(x) &= ax / (1 - \pi(x)). \end{aligned}$$

It follows that the relation between extinction probability and incremental intensity must be:

(23)

$$\pi(x) = \frac{\gamma}{\gamma + ax}.$$

We are led at once to the most elementary solution for the network differential equation (Eq (12)) currently known to us. It is:

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1-(\gamma+ax)(s-1)}. \quad (24)$$

Extraction of the output counting distribution is now a simple matter (see Appendix):

$$\begin{aligned} p_x(0) &= \left(\frac{\gamma}{1+\gamma+ax} \right), \\ p_x(k) &= (1-p_x(0)) \left(\frac{1}{1+\gamma+ax} \right) \left(\frac{\gamma+ax}{1+\gamma+ax} \right)^{k-1}; \end{aligned} \quad (25)$$

where:

- $k = 1, 2, 3, \dots, \infty;$
- a is a units constant;
- x is an incremental intensity above adaptation level;
- γ is an intensity parameter corresponding to adaptation level.

The counting distribution in Eq (25) represents the response of a transmission network (such as the one diagramed in Fig. 4) when a transient stimulus is introduced as an increment above a steady background to which the network has been adapted.

At or near adaptation level ($x = 0$), the network is virtually silent. The chain of events leading back from the sense organ guarantees that very little information will get through. A transient increment in intensity generates a burst of activity (counting) at the output of the network during a fixed observation period following the stimulus. The size of this discharge tends to increase with the size of the increment, although the precise relation depends on the network. A proportional increase, in which average count increases in parallel with the energy of the increment, is implied by Eq (25). Our expression involves an intensity parameter γ , determined by the adaptation level. Evidently the magnitude of γ is also presumed linear with intensity, but the silent condition of the network at adaptation level does not force the assumption. We can accept γ as whatever it turns out to be while the output discharge remains proportional to a stimulus increment. Since the latter is typically quite small in detection experiments, our ultra-linear restriction is neither unrealistic nor disabling. In effect, an adaptation level reproduces conditions approximating absolute threshold wherever the adaptation level happens to be set.

Probable Structure of a Sensory Network

The number of events in an output discharge following a transient stimulus proves to have a Bose-Einstein (geometric) distribution over a wide range of assumptions on details of the

Markov operator and its relation with intensity. The robustness of this outcome must reflect the character of multiplicative internal noise generated by a chain-like structure. Such chains are conceived to be narrowband processes tracing back from minute regions in the sense organ. A stimulus would normally activate a number of these regions, generating an aggregate network output having a Pascal or negative binomial distribution if individual chains are independent. Evidently the tightly packed branching structure of a sensory pathway offers abundant opportunity for multiple occupancy and covariance. Independence is therefore both highly prized and very unlikely. Even so, a network system with the branching structure depicted in Fig. 5, replicated over a number of adjacent receptor fields, would tend to spread out on its journey up a sensory pathway.

FIG. 5

Although each chain penetrates parts of other chains serving nearby receptor units, if there were enough room to spread out, the chains might be kept nearly additive. Additivity is fostered by controlling overall activity level. This is one source of our concern expressed earlier about explosive growth of information in amplifier networks.

Aggregate output in such circumstances has a negative binomial distribution when stimulus intensity is precisely constant from trial to trial. This makes the internal noise of the process

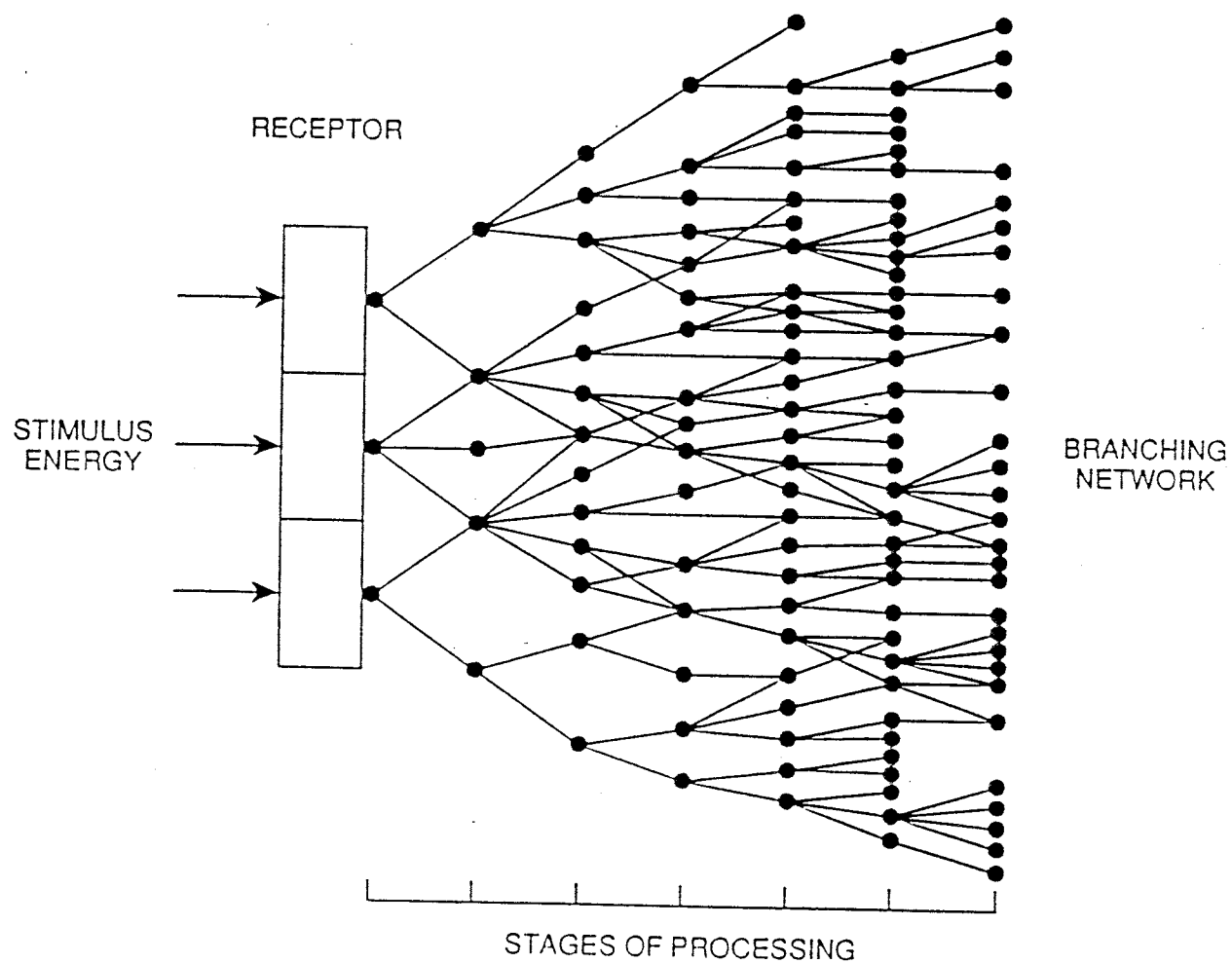


Figure 5: Interactions among nearby chains leading back from receptor surface and forming a branching network. If there is room to spread out and activity levels are not too high, chains may be treated as additive.. At successive stages volume of activity increases but spreads out over an increasingly larger region.

roughly gaussian but with a well-defined relation between mean and variance prescribed by the negative binomial. The latter then governs detectability of non-stochastic signals passing through the network.

Stochastic signals such as wide band acoustic noise or incandescent light are more difficult to study. The statistics of signals drive the counting distribution of the network, making its output a compound of signal variation and network multiplicative noise. Early research on signal detection ignored this compounding altogether.

Some time ago we attempted to introduce compounding via a Poisson flow process for the internal noise, but with results that were not entirely satisfactory. Now, influenced by initial skepticism of Laming's (1986) position, we have constructed a full-scale stochastic model of an amplifier network. We discover that it adapts to steady state backgrounds, and generates Bose-Einstein or negative binomial (not Poisson) counting statistics under transient pure tone stimulation. Moreover, we find unanticipated correspondences between Laming's analysis of differential coupling and our treatment of branching networks.

Importance of Pure Tones

This type of stochastic network analysis is formidable. There

are no known shortcuts outflanking the complex physiology of sensory transmission. Redefinitions of the stimulus based upon clever transformations of psychophysical data remain just as visionary today as they were fifty years ago. Had they existed, shrewd investigators would have discovered them long since. No doubt, pursuit of such easy solutions has deflected concentration from important problems: the phenomena of transmission, for example. This at its core is our critique of the current state of sensory psychology.

Unavoidable complexity ought to place a high premium on simplicity in experimental analysis. We should certainly shun complex experimental designs and complex stimuli. Yet, in still another paradox, the key experiments of auditory signal detection use sinusoids masked by wide band gaussian noise. Experimental work appears driven by a popular masker and an interesting mathematical detection problem, but not by any imperative of the auditory system. Important questions arise in the detectability of noise bursts and sinusoids in noise. Interpreting results is the critical matter. If we were entirely confident that sensory transmission is mediated by an amplifier network, interpretation would be straightforward. The problem would be to work out compounds of signal statistics with the negative binomial noise generated by such transmission. These compounds are not difficult to unravel, but our first priority must be to assure that the basic premises are correct.

If analytical simplicity is the goal, visual detection presents special problems because classical light is inevitably stochastic (see Teich and Saleh, 1988). It then becomes difficult to disentangle a stimulus from multiplicative noise generated in transmission. In many ways this conundrum has contributed to the longevity of Hecht et al.'s (1942) pure stimulus model of visual detection.

Auditory research is potentially less difficult because an analytically simple stimulus, the pure sinusoid, is relatively easy to deliver to the ear intact, provided that signal durations are not too short. Admittedly we do not know what the ear considers analytically simple, but would be well-advised to avoid getting too cute until we understand a great deal more than we do now about auditory transmission.

Pure tones provide excellent probes for stochastic phenomena associated with auditory transmission. The reasons are obvious. Stimuli can be fixed in frequency, phase-locked, and precisely timed. Variability is thus confined to the tone generator's error-level which, with today's equipment, can be made arbitrarily small. Fluctuations must then be attributable to noise created in the auditory system itself. This noise, especially its multiplicative properties, becomes the focus of investigation.

Essentially for these reasons, recent years have witnessed revived interest in one of the traditional problems of auditory psychophysics - pure tone intensity discrimination. We do not attempt to review the new literature here but as a final point in our discussion of transmission networks, we illustrate the interplay between theory and experiment that led to Markov operators.

Pure Tone Intensity Discrimination

The experimental relation between intensity of a pure tone masker and size of a just-detectable increment (or decrement) is shown in Fig. 6. Data are from McGill & Goldberg (1968b).

FIG. 6

They were obtained with pure tones at 1000 HZ and signal durations of 15 or 20 milliseconds. Also plotted are closely related results obtained by Campbell and Lasky (1967). The function is developed over eight log units of masker intensity (80db). For nearly 60db of this range, data fall along a straight line having a slope of roughly 0.9 in log coordinates. It is not Weber's law but a "near-miss" to Weber's law. Similar comparisons made with short bursts of wide band gaussian noise produce a slope of unity over the same range (see Miller (1947) and Green (1960a)).

What accounts for the difference? McGill and Goldberg sought

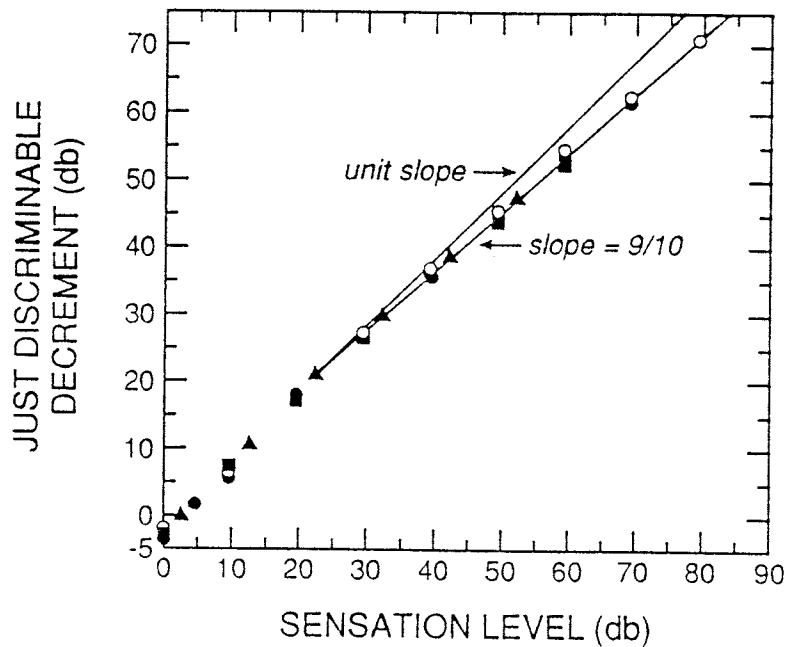


Figure 6: Pure tone intensity discrimination. Abscissa is sensation level (db. above threshold) of 1,000 Hz test tone. Ordinate is sensation level of tone just detectably weaker than test tone. (McGill and Goldberg (1968a,b)). Unit slope depicts Weber's law. Actual slope fitted by least squares is .905. Data of three listeners shown: dark squares (15 msc tones); dark circles (15 msc); dark triangles (20 msc). Open circles are data on 20 msc pure tones obtained by Campbell and Laskey (1967). Campbell-Lasky function is moved horizontally for best match to McGill-Goldberg data.

an explanation in a power function relating masker intensity to the Poisson rate parameter. They worked out the details, showing: 1) intensity discrimination should have a slope of unity with bursts of noise; 2) their value of .9 with pure tones implies a slope of .2 for the loudness function. Since this latter figure is not far off typical loudness functions, and since direct measurements arguably overestimate the loudness slope (Krueger (1989) p.263), McGill and Goldberg believed these indirect loudness measurements supported their Poisson transmission hypothesis.

Others were skeptical. Starting from the fact that the theory, even after a power-function transform, fails to predict intensity discrimination with wide band non-stochastic stimuli (such as "frozen" noise), Poisson transmission attracted few admirers. A conviction grew that something else must be depressing the slope and producing the near-miss to Weber's law.

Viemeister (1972) discovered that when a pure tone was centered in a slot or notch of a low-level wide band masking noise, Weber's law would reappear. The masking slope was restored to unity. Moore and Raab (1974) repeated Viemeister's work, concluding that restoration of Weber's law for pure tones surrounded by notched noise must mean that the barriers posed by the boundaries of the notch effectively block listening off center-frequency as a way to improve detection. Listening is confined to a narrow range of frequencies dominated by the masking stimulus.

If it could then be established that a given transmission network shows Weber's law detectability for individual chains stretching back from the receptor, and also shows the near-miss phenomenon when excitation spreads uncontrolled away from center-frequency with increasing intensity, that network would become a serious candidate for study.

In a recent chapter, McGill and Teich (1990) made such a demonstration using the amplifier network and counting distribution given here as Eq (25). Pure tone intensity discrimination follows Weber's law in these circumstances because of the relation between mean and variance in a Bose-Einstein distribution. If increasing intensity causes the signal to spread out across more and more remote transmission chains, the near-miss will develop. Pure tones embedded in notched noise encounter the edges of the notch as excitation spreads, preventing recruitment of remote chains and limiting bandwidth to the region of the notch. This maneuver effectively fixes the degrees of freedom of the network devoted to transmission, restoring Weber's law.

In their (1990) chapter, McGill and Teich were not fully aware of the adaptation level property of networks constructed as in Fig. 4. Their demonstration was developed for $\gamma = 1$, a low adaptation level such as might accompany a silent or nearly silent background. Hence, the proof was aimed at conditions in which pairs of pure tones are presented in sequence with a brief blank interval

separating the two members - in other words classical intensity discrimination.

This presentation format is quite different from one in which a series of tonal increments are detected as chirps against a steady state tonal background. Adaptation levels are different in the two cases and we should not expect them to yield identical results.

What happens when the two detection formats are compared? Does Weber's law develop in both or, better yet, is there a predictable relation between intensity discrimination and increment detection when identical pure tones are used to generate the comparisons. Laming (1986) posed the question originally in his chapter 13, formulating it as a general experimental issue. Our attention for the present is focused on audition, pure tones, and Markov operators.

The output counting distribution of an ultra-linear network is given in Eq (25). An ideal observer making paired comparisons on the basis of such counts, will generate a psychometric function given by:

$$P(c) = P(j > k) + 1/2 \cdot P(j = k), \quad (26)$$

where:

$P(c)$ is the probability of a correct response in two-alternative forced choice;

j is the count generated by the incrementally stronger pure tone;

k is the count generated by the weaker pure tone.

The probability of a correct response in 2AFC is composed of the probability that signal count exceeds background count, added to the probability of guessing correctly when the two counts happen to coincide.

If Eq (25) is inserted into this formula, we find:

$$P(C) = \left(\frac{\gamma}{1+\gamma+aE_o} \right) \left(\frac{2+\gamma+2aE_s}{2+2\gamma+2aE_s} \right) + \left(\frac{1+aE_o}{1+\gamma+aE_o} \right) \left(\frac{1+aE_s}{1+\gamma+aE_s} \right) \left(\frac{1+2\gamma+2aE_s}{2+4\gamma+2aE_o+2aE_s} \right). \quad (27)$$

This is the psychometric function of a single narrowband chain adapted to a background determined as weak or intense by the size of γ ; and stimulated by two intensity levels, x , above that background. In this instance $x = E_o$, a pure tone energy, or alternatively $x = E_s$, incrementally larger. The complete format is a steady background or pedestal tone with a pair of superimposed chirps differing slightly in intensity, with all components at the same frequency. The psychometric function describes detectability of E_s against E_o , each lying above an adaptation level, γ . Counting distributions are based on the ultra-linear network prescribed in Eq (25).

With everything simple, why should the psychometric function

be so formidable? There are two reasons. First, prolixity accompanies the spike of probability corresponding to a count of zero. The first term of Eq. (27) is generated when the count k attributable to the weaker tone is zero. The right hand term appears when $k > 0$. A second reason is that we are computing a three-parameter (E_0, E_s, γ) expression in order to encompass several different potential experimental comparisons. For example, when adaptation level is set for background silence ($\gamma = 0$), the psychometric function in Eq (27) reduces to:

$$P(c) = \frac{1+2aE_s}{2+2aE_0+2aE_s} \Rightarrow \frac{E_s}{E_0+E_s}$$

(28)

This is essentially the result obtained by McGill & Teich (1990). It is the performance of an ideal observer discriminating two different levels of narrowband Rayleigh noise (see Green and McGill, 1970, eq. (9a); and McGill and Teich, 1990, eq. (11)). We end up with Rayleigh noise because an ultra-linear network is a proportional counter, and the branching process composing the network is in effect a narrowband noise generator. Listening at the back end of the network's terminal stage, we would find it impossible to establish whether this noise came from inside or outside the transmission system. The branching process converts all deterministic inputs into stochastic outputs by infusing them with multiplicative noise.

Obviously, we have uncovered an important result in Eq (28). It is easy to see how gaussian noise required by Laming's differential coupler would develop. If a pure tone were spread out over a region that, in cross section, encompassed many individual branching chains, Rayleigh noise in Eq (28) would become wide band, provided the activity in individual chains could be kept additive. It would then follow that Weber's law must be fundamental in auditory intensity discrimination, an intrinsic part of the transmission mechanism. Before we can actually make such a claim, we need to know how robust Eqs. (25) and (28) really are. They have been developed here under highly simplified conditions stressing linearity and additivity. Nature is clearly not so simpleminded, but if the result is as solid as we suspect it to be, these additional restrictions may not make a great deal of difference.

Suppose we let $E_o = 0$ in Eq. (27), eliminating one member of the pair of comparison tones. The format is then a steady state tonal background and a series of transient chirps or burps at a single intensity detected against the background - pure tone increment detection. In that event:

$$\begin{aligned}
 P(c) &= \left(\frac{\gamma}{1+\gamma} \right) \left(\frac{2+\gamma+2aE_s}{2+2\gamma+2aE_s} \right) + \left(\frac{1}{1+\gamma} \right) \left(\frac{1+aE_s}{1+\gamma+aE_s} \right) \left(\frac{1+2\gamma+2aE_s}{2+4\gamma+2aE_s} \right), \\
 P(c) &\Rightarrow \frac{1+\gamma/2+aE_s}{1+\gamma+aE_s}, \\
 P(c) &\Rightarrow \frac{\gamma/2+aE_s}{\gamma+aE_s}.
 \end{aligned}$$

The second term of Eq (29) is negligible whenever adaptation level is modestly high. Since the second term develops only when background counts are greater than zero, it follows that the network must be virtually silent at adaptation level, a fact now established in several different ways.

Eq (29) again portrays an observer discriminating two levels of narrowband noise. In increment detection the average background count is near zero. This suggests that detectability is influenced chiefly by the variance of the background count. A pure tone increment of fixed size is less detectable against a strong background than against a weak one. We are in effect stating Weber's law for increment detection as it develops in the ultra-linear network.

To compare intensity discrimination and increment detection we need to test Eq (28) against Eq (29), but they do not measure intensities from the same origin. Accordingly, let:

$$\gamma = aE_0 .$$

This shifts the origin in Eq (29) from adaptation level down to zero intensity. the same shift applied to incremental intensity produces:

$$E_s = E_0 + \Delta E.$$

The increment ΔE is then equivalent to the value of E_s in Eq (29). With these notational changes, our psychometric function for

pure tone increment detection becomes:

$$P(c) \Rightarrow 1 - 1/2 \cdot \frac{E_o}{E_g}.$$
(30)

The psychometric function is now set up for direct comparison with its intensity discrimination counterpart in Eq (28). Both exhibit Weber's law; when the intensity ratio is fixed, probability of a correct response is also fixed. In addition, the intensity ratio in Eq (30) corresponds precisely to the extinction probability defined in Eq (23). If a network is long enough to be nearly silent at adaptation level, a strategy for increment detection becomes clear, and psychologically it is not a comparison of the counts associated with increment and background. The observer listens at the output. If anything at all comes through, even a single count, an increment is detected for the interval in which it appears. If nothing comes through, the observer guesses. Counts virtually never occur in both intervals. The strategy is almost exactly that described by Sackitt (1972) for absolute visual detection.

Of course, zero count probability is always less than extinction probability in networks of finite length. A criterion adjustment permitting more effective background suppression might improve detection substantially. Such adjustments would lead to adaptation levels some distance above background intensity. The point is that detection strategies involving an adaptation level

are likely to be different from those in pure intensity discrimination where direct comparisons of numerically large counts are almost mandatory. We know very little about the actual strategies followed by sensory systems.

Comparison of the two psychometric functions, Eq (28) and Eq (30) shows them to be separated by about 3db, with increment detection the more sensitive of the two procedures. Not much credence should be put in these numbers. They are based on properties of a narrowband process, whereas evidence exists (Moore and Raab, 1974), that excitation begins to spread over a range of transmitters when background intensity reaches a modest level 20 to 30db above absolute threshold.

Suppose then we carry out the same calculations assuming that many independent chains are involved. Exact solutions for the psychometric function are difficult but a normal approximation is easy. McGill and Teich (1990) showed that in the ultra-linear network (Eq 25), pure tone intensity discrimination is determined by the energy ratio of increment to background if the spread is uniform and degrees of freedom are fixed.

What happens in increment detection? Using paired comparisons to model the process, we begin with the mean and variance of the counting distribution in Eq (25) and set adaptation level directly on the background. There is, as we have noted, a possibility that

the ear may not use paired comparisons in increment detection. Sackitt's (1972) alternative improves threshold only slightly, even with 5 to 10 independent chains, but the slope of the resulting psychometric function is much steeper than its paired comparison counterpart. We now develop the latter:

$$\begin{aligned} \text{mean count} &= 1 + a\Delta E, \\ \text{variance} &= 2aE_0 + a\Delta E (1+2aE_0 + a\Delta E). \end{aligned} \tag{31}$$

These parameters characterize a single-chain output counting distribution when an increment ΔE is added to the tonal background. Unmodified backgrounds are given by Eq (31) with $\Delta E = 0$.

If the same energies are spread out uniformly across v independent transmission channels, the counting distribution is driven toward gaussian form and the parameters change:

$$\begin{aligned} \text{mean count} &= v + a\Delta E, \\ \text{variance} &= v [2aE_0/v + (a\Delta E/v) (1+2aE_0/v+a\Delta E/v)]. \end{aligned} \tag{31a}$$

The standard normal deviate describing a paired comparison is then:

$$\frac{a\Delta E}{[4aE_0 + (a\Delta E) (1+2aE_0/v + a\Delta E/v)]^{1/2}} = d.$$

Constant terms divided by the background intensity become very small and can be neglected. Hence:

$$\frac{v \cdot \Delta E / E_0}{2 + \Delta E / E_0} \approx d^2$$

Thus for v , the number of independent sensory channels, sufficiently large, and $\Delta E / E_0$ sufficiently small, the standard normal deviate characterizing increment detection in an ultra-linear network may be approximated by:

$$(v \Delta E / 2 E_0)^{1/2} \approx d.$$

(32)

We must remember that this process was said to have been initiated by an energy exchange in the receptor. Yet a spread-out amplifier network, configured as in Fig. 4 and operating in increment detection, will have its detectability determined by stimulus amplitude, not energy. It is unnecessary to define the stimulus as an amplitude. The result simply falls out of the adaptation-level properties of a transmission network operating in increment detection. Our findings confirm a long standing belief that fundamental differences exist between intensity discrimination and increment detection. The principal data are discussed in Green and Swets (1988/1966, Chapter 7). They seem to accord quite well with results presented here although formats of pedestal experiments typically include a wideband masking noise introducing other complexities cited in our development of Eq (27).

Summary

In this paper we have sought to show that precise, uncomplicated models of sensory transmission are not only possible, but capable of contributing deep meaning to psychophysical data, especially to the study of intensity effects in audition. Transmission systems as sophisticated as those nature installs behind the major senses probably include a mechanism for amplifying weak signals. Our version of the mechanism is a Markov operator that can increase its output on each pass through an amplifier stage. Repeated application of the operator generates a branching network that moves information through the system, increasing in volume and spreading out as it goes.

Nature has probably constructed its transmission networks so that a repetitive buildup of information via cycles of branching occurs early in transmission while later stages pass data along more or less as received. Building a primitive model, we do not provide for alterations of function from place to place in the network. We concentrate on a stochastic process operating in the same way everywhere. If the ideas are sound, there will be opportunity for fine tuning later on.

One can argue that introducing a log transform early in transmission eliminates most of the advantages of amplification. The transform squeezes information and the network then amplifies it. Counterbalanced processes generate a linear result leaving us

more or less where we began.

Not quite. This is a subtle topic, and there may be more than just mathematical convenience in the log transform. Rushton's (1961, pp. 171-181) treatment of peripheral coding in the nervous system uses an identical transform. He bases his analysis on experimental evidence from the visual system of the horseshoe crab.

A unique property of a birth-death stochastic process is its adaptation level. Networks with information counts that go both up and down at random are capable of adapting to steady state backgrounds so that output is virtually wiped out in such conditions. This property causes a network to detect changes in nearly the same way at all background levels. Linear transmission might be unrealistic as nature's way, but if the task is to detect small changes in a steady background, a logarithmic relation between increment intensity and network output rarely produces counts large enough to invoke the squeezing action of the log transform. At these low levels there is little difference between log and linear scales. The main effect is likely to be in numbers associated with adaption levels. We treat these as estimated parameters.

The change in perspective contributed by an adaptation level is spectacular. A stimulus is not seen as controlling the number of triggering events transmitted by a network with fixed parameters.

Instead stimulus intensity alters the information rate via a control grid action that inhibits or facilitates passage of information everywhere in the network. When a network is balanced in this way as illustrated in Fig.4, negligible transmission results unless a transient of some kind occurs.

What creates this control grid action? Is it opposed processes of multiplication and inhibition, somehow built into sensory networks by nature? Or does it happen that some parts, perhaps even all, of the multiplied data are lost through wastage and random dropouts. We do not know. If the phenomenon itself is established as necessary for sensory transmission, we can concentrate on identifying neurophysiology that produces it.

An ultimate puzzle in research on sensory thresholds is that we must often wait until the entire conceptual structure is erected before we can figure out how to begin to build its foundation.

Appendix

A counting process is a stochastic mechanism whose observed output consists of discrete events or counts. The process runs for a fixed time producing a counting record that varies from trial to trial in a characteristic probability distribution. Typically we start with some idea of the basic rules and deduce from them the expected form of the counting distribution.

Inference is materially aided by an analytical device known as a probability generating function (p.g.f.). It is a transform of the counting distribution constructed as follows:

$$G(s) = p_0 + p_1s + p_2s^2 + p_3s^3 + \dots + p_k s^k + \dots \quad (\text{A.1})$$

The values of p_k are probabilities of the various counts (including zero), and s is an operational variable used to generate results. Generating functions are especially effective when the counting probabilities combine with s to form a power series that sums into a closed expression for $G(s)$.

If there are two such generating functions $G_x(s)$ and $G_y(s)$ based on independent counting distributions, the product of the p.g.f.s proves to be the generating function of the random variable $x + y$. Now if we can expand this new p.g.f. as a power series in s , it will generate new coefficients of s^k . Evidently these are the counting probabilities in the distribution of $x + y$.

Suppose as an alternative we wish to form a linear combination of two counting processes:

$$G(s) = p_x G_x(s) + p_y G_y(s),$$

where $1 = p_x + p_y$

(A.2)

A combination such as this would be needed when one or the other of the two outcomes occurred at random. The probability distribution of the composite would then be found by expanding $G(s)$ in power series.

In the text of the paper we use both properties to construct the argument leading to Eq. (9). A counting process moves from stage $r-1$ to stage r by applying an operator diagramed in Fig. 4. One of three things can happen to each message event emerging from stage $r-1$. It may be wiped out at random with probability p_0 or reproduced with probability p_1 ; or doubled with probability p_2 . In the latter circumstance, we must deal with the sum of two independent chains formed by each of the progeny. Accordingly, even though we do not know what the generating function at stage r is, we know it must obey the rule:

$$G_r(s) = p_0 + p_1 G_{r-1}(s) + p_2 G_{r-1}^2(s).$$

This is Eq. (9) in the text. It is an iterative restriction governing all transitions between successive stages.

Our text shows in Eqs. (9-12) how the rule is converted into a time-based differential equation. Eq (15) displays the solution of the differential equation. The point of this maneuver is that the solution is free of the iterative restriction. If the resulting generating function can be expanded as a power series in s , it will yield the output counting distribution of the process at time t , corresponding to an arbitrary number r of processing stages.

For example, when $p_0 = p_2$ (i.e. $\mu = \lambda$) the counting process is said to be at its adaptation level. Our argument leading to Eq (13) shows the solution of the differential equation to be:

$$G_t(s) = 1 + \frac{s-1}{1-\lambda t(s-1)}. \quad (\text{A.3})$$

To expand this expression as a series in s , divide through the second term on the right by $(1 + \lambda t)$:

$$\begin{aligned} G_t(s) &= 1 + \left(\frac{s-1}{1+\lambda t} \right) \cdot \left(\frac{1}{1 - \left(\frac{\lambda t}{1+\lambda t} \right) s} \right), \\ G_t(s) &= 1 + \left(\frac{s-1}{1+\lambda t} \right) \left(1 + \left(\frac{\lambda t}{1+\lambda t} \right) s + \left(\frac{\lambda t}{1+\lambda t} \right)^2 s^2 + \dots + \left(\frac{\lambda t}{1+\lambda t} \right)^k s^k + \dots \right), \\ G_t(s) &= \left(\frac{\lambda t}{1+\lambda t} \right) + \frac{1}{(1+\lambda t)^2} s + \frac{\lambda t}{(1+\lambda t)^3} s^2 + \dots + \frac{(\lambda t)^{k-1}}{(1+\lambda t)^k} \left(1 - \frac{\lambda t}{1+\lambda t} \right) s^k + \dots, \\ G_t(s) &= \frac{\lambda t}{1+\lambda t} s^0 + \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(1+\lambda t)^{k+1}} \cdot s^k. \end{aligned} \quad (\text{A.4})$$

Eq (A.4) for the generating function displays the values of $p_t(k)$ in Eq (14) as coefficients of s^k . They form the output counting distribution of the process at adaptation level.

Similarly, when a log transform in Eq (17) is applied to Eq (15) in the text, the latter becomes:

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1-\gamma(x)(s-1)}.$$

(A.5)

This is Eq (20) in the text and we expand it in exactly the same way to produce a Bose-Einstein distribution given as Eq (21). Finally when

$$\gamma(x) = \gamma + ax,$$

and the extinction probability,

$$\pi(x) = \frac{\gamma}{\gamma + ax},$$

then Eq(A.5) for the generating function simplifies to:

$$G_x(s) = 1 + \frac{(1+ax)(s-1)}{1-(\gamma+ax)(s-1)}.$$

(A.6)

This is a so-called "ultra-linear" version of the generating function. To expand $G_x(s)$, observe that dividing through the second term on the right by $(1+\gamma+ax)$ produces:

$$G_x(s) = 1 + \left(\frac{1+ax}{1+\gamma+ax} \right) (s-1) \cdot \left(\frac{1}{1 - \frac{\gamma+ax}{1+\gamma+ax} s} \right),$$

$$G_x(s) = \frac{\gamma}{1+\gamma+ax} s^0 + \left(\frac{1+ax}{1+\gamma+ax} \right) \cdot \sum_{k=1}^{k=\infty} \frac{(\gamma+ax)^{k-1}}{(1+\gamma+ax)^k} s^k.$$

(A.7)

The coefficients of s^k now generate the ultra-linear counting distribution given as Eq(25) in the text.

Moments of each of these counting distributions can be found via their generating functions. Differentiation of $G(s)$ with respect to s in Eq (A.1) generates the mean value of the counting distribution when $s=1$. A second differentiation produces the second factorial moment from which the variance of the counting distribution may be easily obtained. Specifically:

$$m = G'(1),$$

$$\sigma^2 = G''(1) + G'(1) - (G'(1))^2.$$

The generating functions developed here are usually easier to expand as functions of $(s-1)$. Accordingly let:

$$z=s-1$$

and note that:

$$\frac{dG(s)}{ds} = \frac{dG(z)}{dz},$$

where $z=0$ when $s=1$.

Then the generating function in Eq (A.3) can be rewritten as:

$$\begin{aligned} G_t(z) &= 1 + \frac{z}{1 - (\lambda t) z}, \\ &= 1 + z + (\lambda t) z^2 + (\lambda t)^2 z^3 + \dots \end{aligned}$$

Differentiating this expansion once shows the mean to be unity when z is set equal to zero. A second differentiation produces $2\lambda t$ for the second factorial moment. Hence:

$$\begin{aligned} m &= 1, \\ \sigma^2 &= 2\lambda t, \end{aligned}$$

(A.8)

in the adaptation level counting distribution given as Eq (14) in the text.

Finally, the ultra-linear generating function in Eq (A.6) converted to the z variable is:

$$\begin{aligned} G_x(z) &= 1 + \frac{(1+ax) z}{1 - (\gamma+ax) z} \\ &= 1 + (1+ax) z + (1+ax) (\gamma+ax) z^2 + (1+ax) (\gamma+ax)^2 z^3 + \dots \end{aligned}$$

Differentiation shows:

$$m = 1 + ax, \quad (A.9)$$

after z is set equal to zero. A second differentiation yields:

$$\begin{aligned} \sigma^2 &= 2(1+ax) (\gamma+ax) + (1+ax) - (1+ax)^2, \\ &= (1+ax) (2\gamma+ax). \end{aligned}$$

(A.10)

When x , the incremental intensity, is reduced to zero, the mean and variance regress to their respective adaptation level values given by Eq (A.8).

If we now adjust our notation to correspond with Eq (30) in the text:

$$\begin{aligned}\gamma &= aE_o, \\ \gamma + ax &= aE_s, \\ ax &= a\Delta E,\end{aligned}$$

we can rewrite Eqs (A.9-A.10) as follows:

$$\begin{aligned}m &= 1 + a\Delta E, \\ \sigma^2 &= (1 + a\Delta E) (aE_o + aE_s).\end{aligned}$$

(A.11)

The latter expression for the variance can then be reworked as:

$$\begin{aligned}\sigma^2 &= (1 + a\Delta E) (2aE_o + a\Delta E), \\ &= 2aE_o + a\Delta E(1 + 2aE_o + a\Delta E),\end{aligned}$$

and these are the mean and variance of the ultra-linear counting distribution (i.e. Eq (25)), given by Eq (35) in the text.

REFERENCES

- Barlow, H. (1956). Retinal noise and absolute threshold. *Journal of the Optical Society of America*, 46, 634-639.
- Békésy, G.v., & Rosenblith, W. (1951). The mechanical properties of the ear. Chapter 27 in S.S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley.
- Bharucha-Reid, A. (1960). *Elements of the theory of Markov processes and their applications*. New York: McGraw-Hill.
- Brown, J. (1965). The structure of the visual system. In C.H. Graham (Ed.), *Vision and visual perception*. New York: Wiley.
- Campbell, R., & Lasky, E. (1967). Masker level and sinusoidal-signal detection. *Journal of the Acoustical Society of America*, 42, 972-976.
- Ehret, G., & Merzenich, M. (1985). Auditory midbrain responses parallel spectral integration phenomena. *Science*, 229, 1245-1247.
- Feller, W. (1957). *An introduction to probability theory and its applications V.1*. New York: Wiley.
- Geisler, W. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, 96, 267-314.
- Graham, C. (1965). Color mixture and color systems. In C.H. Graham (Ed.), *Vision and visual perception*. New York: Wiley.
- Green, D. (1960a). Detection of a noise signal. *Journal of the Acoustical Society of America*, 32, 121-131.
- Green, D. (1960b). Psychoacoustics and detection theory. *Journal of the Acoustical Society of America*, 32, 1189-1203.
- Green, D., & McGill, W. (1970). On the equivalence of detection probabilities and well-known statistical quantities. *Psychological Review*, 77, 294-301.
- Green, D., & Swets, J. (1988/1966). *Signal detection theory and psychophysics*. Originally published, New York: Wiley (1966). Revised and reprinted, Los Altos, CA: Peninsula Publishing (1988).

Hecht, S., Shlaer, S., & Pirenne, M. (1942). Energy, quanta, and vision. *Journal of General Physiology*, 25, 819-840.

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology*, 60, 1-29.

Hodgman, C., Selby, S., & Weast, R. (Eds.). (1954). *Mathematical Tables from Handbook of Chemistry and Physics*. Cleveland, OH: Chemical Rubber Publishing Co.

Jeffress, L. (1964). Stimulus-oriented approach to detection. *Journal of the Acoustical Society of America*, 36, 766-774.

Jeffress, L. (1968). Mathematical and electrical models of auditory detection. *Journal of the Acoustical Society of America*, 44, 187-203.

Krueger, L. (1989). Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behavioral and Brain Sciences*, 12, 251-320.

Kuhn, T. (1957). *The Copernican revolution*. Cambridge, MA: Harvard University Press.

Laming, D. (1986). *Sensory analysis*. London: Academic Press.

Margolis, H. (1987). *Patterns, thinking, and cognition*. Chicago: University of Chicago Press.

McGill, W. (1967). Neural counting mechanisms and energy detection in audition. *Journal of Mathematical Psychology*, 4, 351-376.

McGill, W., & Goldberg, J. (1968a). Pure tone intensity discrimination and energy detection. *Journal of the Acoustical Society of America*, 44, 576-581.

McGill, W., & Goldberg, J. (1968b). A study of the near-miss involving Weber's law and pure-tone intensity discrimination. *Perception and Psychophysics*, 4, 105-109.

McGill, W., & Teich, M. (1989). A unique approach to stimulus detection theory in psychophysics based upon the properties of zero-mean gaussian noise. (Review of Laming's *Sensory Analysis*). *Journal of Mathematical Psychology*, 33, 99-108.

McGill, W. & Teich, M. (1990). Auditory signal detection and amplification in a neural transmission network. In M. Commons, & J. Nevin (Eds.), *Signal Detection*. (Chapter 1). Hillsdale, NJ: Erlbaum.

Miller, G. (1947). Sensitivity to changes in the intensity of white noise and its relation to masking and loudness. *Journal of the Acoustical Society of America*, 19, 609-619.

Moore, B., & Raab, D. (1974). Pure tone intensity discrimination: Some experiments relating to the "near miss" to Weber's law. *Journal of the Acoustical Society of America*, 55, 1049-1054.

Peřina, J. Superposition of coherent and incoherent fields, *Physics Letters*, 1967, 24A, 333-334.

Rushton, W. (1961). Peripheral coding in the nervous system. In W. Rosenblith (Ed.), *Sensory communication*. New York: Wiley.

Sackitt, B. (1972). Counting every quantum. *Journal of Physiology*, 223, 131-150.

Stevens, S., & Davis, H. (1938). *Hearing: Its psychology and physiology*. New York: Wiley.

Tanner, W., & Swets, J. (1954). The human use of information, I: Signal detection for the case of the signal known exactly. *Transactions of the IRE professional Group on Information Theory*, PGIT-4, 213-221.

Teich, M., & McGill, W. (1976). Neural counting and photon counting in the presence of dead time. *Physical Review Letters*, 36, 754-758.

Teich, M., Prucual, P., Vanucci, G., Breton, M., & McGill, W. (1982). Multiplication noise in human visual system at threshold: 1. Quantum fluctuations and minimum detectable energy. *Journal of the Optical Society of America*, 72, 419-431.

Teich, M., & Saleh, B. (1988). Photon bunching and antibunching. In E. Wolf (Ed.), *Progress in Optics XXVI*, Amsterdam: Elsevier.

Viemeister, N. (1971). Intensity discrimination of pulsed sinusoids. The effect of filtered noise. *Journal of the Acoustical Society of America*, 51, 1265-1269.

CHIP Technical Report List

1. David M. Green and William J. McGill. *On the equivalence of detection probabilities and well known statistical quantities.* October 1969.
2. Donald A. Norman. *Comments on the information structure of memory.* October 1969.
3. Norman H. Anderson. *Functional measurement and psychophysical judgment.* October 1969.
4. James C. Shanteau. *An additive decision-making model for sequential estimation and interference judgments.* October 1969.
5. Norman H. Anderson. *Averaging model applied to the size-weight illusion.* October 1969.
6. Norman H. Anderson and James C. Shanteau. *Information integration in risky decision making.* November 1969.
7. George Mandler, Richard H. Meltzer, and Zena Pearlstone. *The structure of recognition. Effects of list tags and of acoustic and semantic confusion.* November 1969.
8. Dominic Massaro. *Perceptual processes and forgetting in memory tasks.* January 1970.
9. Daniel Graboi. *Searching for targets: The effects of specific practice.* February 1970.
10. James H. Patterson and David M. Green. *Discrimination of transient signals having identical energy spectra.* February 1970.
11. Donald A. Norman. *Remembrance of things past.* June 1970.
12. Norman H. Anderson. *Integration theory and attitude change.* August 1970.
13. A.D. Baddeley and J.R. Ecob. *Reaction time and short-term memory: A trace strength alternative to the high-speed exhaustive scanning hypothesis.* November 1970.
14. A.D. Baddeley. *Retrieval rules and semantic coding in short-term memory.* December 1970.
15. Roy D. Patterson. *Residue pitch as a function of the number and relative phase of the component sinusoids.* March 1971.
16. George Mandler and Marilyn A. Borges. *Effects of list differentiation, category membership and prior recall on recognition.* May 1971.
17. David E. Rumelhart, Peter H. Lindsay, and Donald A. Norman. *A process model for long-term memory.* May 1971.
18. David E. Rumelhart and Adele A. Abrahamsen. *Toward a theory of analogical reasoning.* July 1971.
19. Martin F. Kaplan. *How response dispositions integrate with stimulus information.* August 1971.
20. Martin F. Kaplan and Norman H. Anderson. *Comparison of information integration and reinforcement models for interpersonal attraction.* August 1971.
21. David M. Green and R. Duncan Luce. *Speed-accuracy tradeoff in auditory detection.* September 1971.
22. David E. Rumelhart. *A multicomponent theory of confusion among briefly exposed alphabetic characters.* November 1971.
23. Norman H. Anderson and Arthur J. Farkas. *New light on order effects in attitude change.* March 1972.
24. Norman H. Anderson. *Information integration theory: A brief survey.* April 1972.
25. Donald A. Norman. *Memory, knowledge, and the answering of questions.* May 1972.
26. David J. Weiss. *Averaging: An empirical validity criterion for magnitude estimation.* Norman H. Anderson. *Cross-task validation of functional measurement.* June 1972.
27. David E. Rumelhart and Patricia Siple. *The process of recognizing tachistoscopically presented words.* August 1972.
28. Ebbe B. Ebbesen and Richard Bowers. *The effects of proportion of risky to conservative arguments in a group discussion on risky shift.* September 1972.

29. Ebbe B. Ebbesen and Michael Haney. *Flirting with death: Variables affecting risk taking on our nation's highways*. September 1972.
30. Norman H. Anderson. *Algebraic models in perception*. November 1972.
31. Norman H. Anderson. *Cognitive algebra: Information integration applied to social attribution*. December 1972.
32. Jean M. Mandler and Nancy L. Stein. *Recall recognition of pictures by children as a function of organization and of distractor similarity*. January 1973.
33. David E. Rumelhart and Donald A. Norman. *Active semantic networks as a model of human memory*. Marc Eisenstadt and Yaakov Kareev. *Towards a model of human game playing*. June 1973.
34. George Mandler. *Memory storage and retrieval: Some limits on the reach of attention and consciousness*. July 1973.
35. Kent L. Norman. *A method of maximum likelihood estimation for stimulus integration theory*. August 1973.
36. Yaakov Kareev. *A model of human game playing*. August 1973.
37. Donald A. Norman. *Cognitive organization and learning*. August 1973.
38. The Center for Human Information Processing: A Five Year Report - 1968-73.
39. Larry D. Rosen and J. Edward Russo. *Binary processing in multi-alternative choice*. October 1973.
40. Samuel Himmelfarb and Norman H. Anderson. *Integration theory analysis of opinion attribution*. December 1973.
41. George Mandler. *Consciousness: Respectable, useful, and probably necessary*. March 1974.
42. Norman H. Anderson. *The problem of change-of-meaning*. June 1974.
43. Norman H. Anderson. *Methods for studying information integration*. June 1974.
44. Norman H. Anderson. *Basic experiments in person perception*. June 1974.
45. Norman H. Anderson. *Algebraic models for information integration*. June 1974.
46. Ebbe B. Ebbesen and Vladimir J. Konecni. *Cognitive algebra in legal decision making*. September 1974.
47. Norman H. Anderson. *Equity judgments as information integration*.
Arthur J. Farkas and Norman H. Anderson. *Input summation and equity summation in multi-cue equity judgments*. December 1974.
48. George Mandler and Arthur Graesser II. *Dimensional analysis and the locus of organization*. January 1975.
49. James L. McClelland. *Preliminary letter identification in the perception of words and nonwords*. April 1975.
50. Donald A. Norman and Daniel G. Bobrow. *On the role of active memory processes in perception and cognition*. May 1973.
51. J. Edward Russo. *The value of unit price information. An information processing analysis of point-of-purchase decisions*. June 1975.
52. Elissa L. Newport. *Motherese: The speech of mothers to young children*. August 1975.
53. Norman H. Anderson and Cheryl C. Graesser. *An information integration analysis of attitude change in group discussion*. September 1975.
54. Lynn A. Cooper. *Demonstration of a mental analog of an external rotation*.
Lynn A. Cooper and Peter Podgorny. *Mental transformations and visual comparison processes: Effects of complexity and similarity*. October 1975.
55. David E. Rumelhart and Andrew Ortony. *The representation of knowledge in memory*. January 1976.
56. David E. Rumelhart. *Toward an interactive model of reading*. March 1976.
57. Jean M. Mandler, Nancy S. Johnson, and Marsha DeForest. *A structural analysis of stories and their recall: From "Once upon a time" to "Happily ever after"*. March 1976.
58. David E. Rumelhart. *Understanding and summarizing brief stories*. April 1976.
59. Lynn A. Cooper and Roger N. Shepard. *Transformations on representations of objects in space*. April 1976.
60. George Mandler. *Some attempts to study the rotation and reversal of integrated motor patterns*. May 1976.
61. Norman H. Anderson. *Problems in using analysis of variance in balance theory*. June 1976.
62. Norman H. Anderson. *Social perception and cognition*. July 1976.
63. David E. Rumelhart and Donald A. Norman. *Accretion, tuning and restructuring: Three modes of learning*. August 1976.
64. George Mandler. *Memory research reconsidered: A critical view of traditional methods and distinction*. September 1976.

65. Norman H. Anderson and Michael D. Klitzner. *Measurement of motivation*.
Michael D. Klitzner and Norman H. Anderson. *Motivation x expectancy x value: A functional measurement approach*. November 1976.
66. Vladimir J. Konecni. *Some social, emotional, and cognitive determinants of aesthetic preference for melodies in complexity*. December 1976.
67. Hugh Mehan, Courtney B. Cazden, LaDonna Coles, Sue Fisher, Nick Maroules. *The social organization of classroom lessons*. December 1976.
- 67a. Hugh Mehan, Courtney B. Cazden, LaDonna Coles, Sue Fisher, Nick Maroules. *Appendices to the social organization of classroom lessons*. December 1976.
68. Norman H. Anderson. *Integration theory applied to cognitive responses and attitudes*. December 1976.
69. Norman H. Anderson and Diane O. Cuneo. *The height + width rule in children's judgments of quantity*.
Norman H. Anderson and Clifford H. Butzin. *Children's judgments of equity*. June 1977.
70. Donald R. Gentner and Donald A. Norman. *The FLOW tutor: Schemas for tutoring*. June 1977.
71. George Mandler. *Organization and repetition: An extension of organizational principles with special reference to rote learning*. May 1977.
72. Manuel Leon. *Coordination of intent and consequence information in children's moral judgments*. August 1977.
73. Ted Supalla and Elissa L. Newport. *How many seats in a chair? The derivation of nouns and verbs in American Sign Language*. November 1977.
74. Donald A. Norman and Daniel G. Bobrow. *Descriptions: A basis for memory acquisition and retrieval*. November 1977.
75. Michael D. Williams. *The process of retrieval from very long term memory*. September 1978.
76. Jean M. Mandler. *Categorical and schematic organization in memory*. October 1978.
77. James L. McClelland. *On time relations of mental processes: A framework for analyzing processes in cascade*. October 1978.
78. Jean M. Mandler, Sylvia Scribner, Michael Cole, and Marsha DeForest. *Cross-cultural invariance in story recall*.
Jean M. Mandler and Marsha DeForest. *Developmental invariance in story recall*. November 1978.
79. David E. Rumelhart. *Schemata: The building blocks of cognition*. December 1978.
80. Nancy S. Johnson and Jean M. Mandler. *A tale of two structures: Underlying and surface forms in stories*. January 1979.
81. David E. Rumelhart. *Analogical processes and procedural representations*. February 1979.
82. Ross A. Bott. *A study of complex learning: Theory and methodologies*. March 1979.
83. Laboratory of Comparative Human Cognition. *Toward a unified approach to problems of culture and cognition*. May 1979.
84. George Mandler and Lawrence W. Barsalou. *Steady state memory: What does the one-shot experiment assess?* May 1979.
85. Norman H. Anderson. *Introduction to cognitive algebra*. June 1979.
86. Michael Cole, Edwin Hutchins, James Levin, and Naomi Miyake, Editors. *Naturalistic problem solving and microcomputers. Report of a Conference*. June 1979.
87. Donald A. Norman. *Twelve issues for cognitive science*. October 1979.
88. Donald A. Norman. *Slips of the mind and an outline for a theory of action*. November 1979.
89. The Center for Human Information Processing: *A Description and a Five-Year Report (1974-79)*. November 1979.
90. Michael Cole and Peg Griffin. *Cultural amplifiers reconsidered*. December 1979.
91. James L. McClelland and David E. Rumelhart. *An interactive activation model of the effect of context in perception. Part I*. April 1980.
92. James L. McClelland and J.K. O'Regan. *The role of expectations in the use of peripheral visual information in reading*. February 1980.
93. Edwin Hutchins. *Conceptual structures of Caroline Island navigation*. May 1980.
94. Friedrich Wilkening and Norman H. Anderson. *Comparison of two rule assessment methodologies for studying cognitive development*. June 1980.

95. David E. Rumelhart and James L. McClelland. *An interactive activation model of the effect of context in perception. Part II.* August 1980.
96. Jean M. Mandler. *Structural invariants in development.* September 1980.
97. David E. Rumelhart and Donald A. Norman. *Analogical processes in learning.* October 1980.
98. James A. Levin and Yaakov Kareev. *Personal computers and education: The challenge to schools.* November 1980.
99. Donald A. Norman and Tim Shallice. *Attention to action: Willed and automatic control of behavior.* December 1980.
100. David E. Rumelhart. *Understanding understanding.* January 1981.
101. George Mandler. *The structure of value: Accounting for taste.* May 1981.
102. David E. Rumelhart and Donald A. Norman. *Simulating a skilled typist: A study of skilled cognitive-motor performance.* May 1981.
103. Jean M. Mandler. *Representation.* June 1981.
104. Donald R. Gentner. *Skilled finger movements in typing.* July 1981.
105. Edwin L. Hutchins and James A. Levin. *Point of view in problem solving.* August 1981.
106. Michael Cole. *Society, mind and development.*
Michael Cole. *The zone of proximal development: Where culture and cognition create each other.* September 1981.
107. Laboratory of Comparative Human Cognition. *Culture and cognitive development.* November 1981.
108. Donald R. Gentner. *Evidence against a central control model of timing in typing.* December 1981.
109. Robert M. Boynton and Allen L. Nagy. *The La Jolla analytic colorimeter: Optics, calibrations, procedures and control experiments.* December 1981.
110. Jonathan T. Grudin and Serge Larochelle. *Digraph frequency effects in skilled typing.* February 1982.
111. Dimensional analysis and the locus of organization. Donald R. Gentner, Jonathan T. Grudin, Serge Larochelle, Donald A. Norman, David E. Rumelhart. *Studies of typing from the LNR Research Group.* September 1982.
112. Donald A. Norman. *Five papers on human-machine interaction.* May 1982.
113. Naomi Miyake. *Constructive interaction.* June 1982.
114. Donald R. Gentner. *The development of typewriting skill.* September 1982.
115. Margaret M. Riel. *Investigating the system of development: The skills and abilities of dysphasic children.* February 1983.
116. David E. Rumelhart and Donald A. Norman. *Representation in memory.* June 1983.
117. George Mandler. *Consciousness: Its function and construction.* June 1983.
118. Norman H. Anderson. *Schemas in person cognition.* September 1983.
119. Norman H. Anderson. *A theory of stereotypes.* December 1983.
120. Norman H. Anderson. *Psychodynamics of everyday life: Blaming and avoiding blame.* December 1983.
121. Donald R. Gentner. *Expertise in typewriting.* May 1984.
122. Alessandro Duranti. *Intentions, self, and local theories of meaning: Words and social action in a Samoan context.* June 1984.
123. Geoffrey Hinton and Paul Smolensky. *Parallel computation and the mass-spring model of motor control.* June 1984.
124. Donald R. Gentner. *Skilled motor performance at variable rates: A composite view of motor control.* October 1985.
125. George Mandler, Carmen Overson, Peter Graf, and Dolores Kraft. *Activation and elaboration: Indirect sources and effects.* November 1985.
126. Geoffrey M. Boynton. *Documentation of computer programs for experiments on categorical color perception.* September 1987.
127. Conrad X. Olson. *A shortcut method for naming OSA color specimens.* March 1988. January 1990.
128. John S. Houston and Pamela A. Sample. *Programmable automated perimetry: UCSD user's guide.* April 1988.
129. William J. McGill and Malvin C. Teich. *Auditory signal detection and amplification in a neural transmission network.* May 1989.

130. Yrjö Engeström. *Developing thinking at the changing workplace: Toward a redefinition of expertise*. October 1989.
131. Michael Cole. *Cultural psychology: A once and future discipline?* January 1990.
132. William J. McGill and Malvin C. Teich. *Simple models of sensory transmission*. January 1991.