A correlational model applied to metacontrast: Reply to Weisstein, Ozog, and Szoc

BRUCE BRIDGEMAN

Psychology Board of Studies, University of California, Santa Cruz, California 95064

A model based on Hartline-Ratliff lateral inhibition (LI) with a time delay shows that LI distributes information across a parallel network, makes output noise lower than input noise, and can simulate psychophysical results in several variations of the metacontrast-masking paradigm. This model answers many of the Weisstein, Ozog, and Szoc (1975) criticisms of a previous version of the Bridgeman lateral inhibitory model: Temporal oscillations in the masking function are eliminated, and differentiation of homophotic from metaphotic activity by the modeled system is shown to be unnecessary. Experimental results consistent with the LI theory but inconsistent with two-channel theories are discussed.

"Metacontrast" masking (Stigler, 1910) is of interest because the effect seems to work backward in time, a later mask (M) obscuring an earlier target (T). It suggests that the internal representation of a stimulus remains vulnerable to interference for a short time, and thus provides a tool for dissecting temporal aspects of visual coding. The promise of this tool has resulted in several neuronal models of metacontrast (Breitmeyer & Ganz, 1976; Bridgeman, 1971; Weisstein, 1968, 1972). The extensive literature on metacontrast has been reviewed by Lefton (1973) and by Weisstein (1972).

THE LATERAL INHIBITORY MODEL

The present model, an extension of an earlier one (Bridgeman, 1971), is based on the presence of lateral inhibition (LI) in the visual cortex (Benevento, Creutzfeldt, & Kuhnt, 1972). Details of the model and assignment of parameters have been given previously (Bridgeman, 1971). Briefly, a lateral inhibitory nerve network is adapted from the Hartline-Ratliff equations (Ratliff, 1965) and modified for application to mammals. Each cell in an array of n cells is modified in its firing by the equation

$$r_p(t) = e_p(t) - \sum_{j=1}^{n} k_{p,j}[r_j(t - |p - j|) - r_{p,j}^{\circ}] \quad j \neq p, \quad (1)$$

where r_p is the firing rate of neuron p, r_j is the rate of neuron j, e_p is the excitatory input to p, $k_{p,j}$ is an inhibitory coefficient expressing the effect of j on p, and $r_{p,j}^{\circ}$ is the corresponding inhibitory threshold. An iteration in the simulation represents 30 msec of real time; this is approximately the latency of reciprocal LI interactions in cat lateral geniculate (Bridgeman, 1972; Singer & Creutzfeldt, 1970), and represents a physiologically reasonable latency for inhibitory effects. Each "unit" in the model might represent a small group or column of cortical neurons.

It is assumed here that all receptive fields are identical; Weisstein, Ozog, and Szoc (1975) point out that this assumption

does not necessarily contradict the evidence that single units in the visual system have widely different characteristics, since all units could contribute to a composite size and response rate. This is an example of the technique of modeling only part of a system; only simplified sustained-cell-like neurons are modeled.

A "Linking Hypothesis"

The process of recognition (re-cognition) requires a comparison of the present input with some past input or genetically programmed state. This process has been modeled by performing a simulation under one set of conditions, storing the network activity for all iterations, and repeating the entire simulation under different stimulus conditions. Activity generated by the network in the second pass is then compared to that stored from the first pass, iteration by iteration. A high average correlation r between the two sets of outputs means that activity similar to previous activity is being generated in the network. Recognition then becomes a decision process based on the average correlation of present activity with previously stored information. Uttal (1973) has proposed a simple nerve network which can perform autocorrelations; a similar one could mediate cross-correlations.

Revisions of the Model

There is evidence that a simple r is not the best similarity measure, for a rectification process may take place in a visual system. Spinelli and Barrett (1969), for instance, mapped single cells in the visual cortex of the cat under two conditions, one with a black stimulus on a white background and the other with white on black; some cells were insensitive to reversal of contrast. Measuring human occipital visual evoked potentials, Spekreijse, Van der Tweel, and Zuidema (1973) found that response is determined by the instantaneous relative contrast whether it is reached by an increase or a decrease in luminance. Thus, the evoked response was related to change in contrast rather than luminance. Spekreijse et al. (1973) also found that the electroretinogram had the opposite characteristic, implying that the rectification takes place at a higher anatomical level than the retina and that it is related to extraction of contrast information. To simulate an anatomically late rectification, the rs were squared; this measure can then be interpreted as the proportion of the variance in network activity which can be attributed to the presence of the stimulus. A squared measure has the further advantage that the stimulus-specific information contained in negative correlations is not lost.

Independent Gaussian noise was added to each unit's input at each iteration of LI. The noise simulates the variability in background activity which is observed physiologically in single cells in intact preparations.

The responses of the cells in the lateral inhibitory model (LIM) are similar to those found physiologically in cortical cells with late peaks (Bridgeman, 1975): Both simulated and observed neurons show a large burst of firing followed by a period of inhibition and a subsequent lower peak of firing.

SIMULATION OF METACONTRAST

Figure 1 shows the results of presenting metacontrast stimuli to the network. Each point in the figure represents an average squared correlation between two simulated runs, the first simulation done with T alone and the second with the T-M latency indicated on the horizontal axis. The "early" interval consists of averages of the first three iterations, while the "late" interval is calculated by averaging iterations 1-9. The two r^2 values at each Δt thus consist of measures of early or late activity in the same pair of simulations. Figure 1 is essentially a replication of earlier simulations (Bridgeman, 1971) with a larger network, more Δt values, the addition of Gaussian noise, and a display format similar to that of psychophysical studies of metacontrast.

The simulated curve should be compared to dichoptic psychophysical masking functions, for the LIM allows no interactions of T and M before the cortical level. The results of Schiller and Smith (1968) can be used as comparison functions offering a close approximation to the conditions used here; the stimuli are equal in area, brightness, and energy, are spatially adjacent, and are presented dichoptically for equal durations. The real and simulated masking curves show the following similarities: (1) All curves are U shaped, with a minimum at 40-80 msec; (2) masking begins before $\Delta t = 0$ and the function decreases smoothly to the minimum; (3) the function increases slowly after the minimum, asymptoting at about 200 msec. (The LIM, operating in a noisy environment, would have to be satisfied with $r^2 s < 1.0$. Otherwise, recognition could never occur. In the context

of the model, a r^2 of more than about .6 should be taken to indicate good recognition.)

The model reproduces the U-shaped masking function when a long time is allowed for network reverberation, and a weaker monotonic function when only a short time is allowed. In addition, the model yields modifications of the simulated masking curve in the psychophysically appropriate directions for four other spatial manipulations of masking: identical T and M, large M, large T-M spatial separation, and masking with overlapping disks. In summary, the performance of the model on spatial manipulations of stimuli is six out of six: the two basic effects plus four modifications of the stimuli.

Of five other conditions simulated, the model performs well on three (one- vs. two-iteration stimulus duration, repeated T-M sequence, and "black" T and M), responds in the correct direction but with distortions on one more (T/M energy ratio), and fails to predict the psychophysically obtained result in one (the two-mask paradigm). Details of these simulations are in preparation.

REPLY TO WEISSTEIN, OZOG, AND SZOC

Using the specifications of the earlier version of this model (Bridgeman, 1971), Weisstein et al. (1975) have repeated and extended the lateral inhibitory simulations, producing a critique of the model. This section of the paper reviews the comments of Weisstein et al. in view of the second-generation simulations reviewed above.

The first point concerns minor discrepancies between the network oscillations obtained by Bridgeman (1971) and by Weisstein et al. (1975). To double check the accuracy of oscillations, the LI simulation program was rewritten in a different format, with a different computer language, and run on a larger computer (PDP-11). Network activity was identical to that described in 1971. Despite small differences in modeled responses, however, all of the simulation results which Weisstein et al. bring up are valid for the 1971 model.

Another minor point is that in 1971 I displayed only the peaks of the cross-correlation functions for the points F(O), G(O); as Weisstein et al. point out, this reduces to a Pearson r correlation. The same convention is used here.

Though Weisstein et al. (1975, p. 331) assert that the LIM requires the central levels of the visual system to be able to distinguish metaphotic (after the stimulus) from homophotic (during the stimulus) activity, the only assumption required is that perception takes time, and that the visual system must use metaphotic activity only if homophotic activity is not available due to the brevity of a stimulus. Homophotic and metaphotic activity need not be distinguished.

The problem of temporal oscillation in the masking function is addressed by Weisstein et al. (1975, p. 333 and Figure 4). Figure 1 shows that the current LIM does not
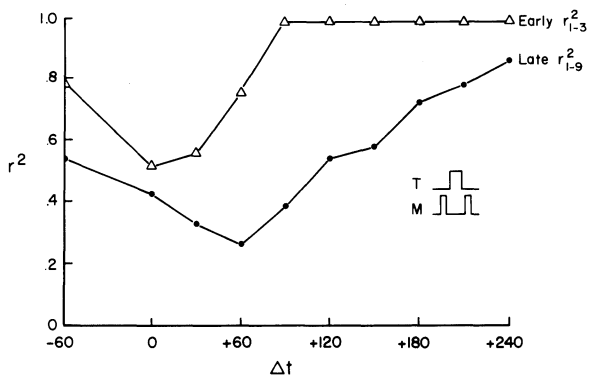

Figure 1. Backward masking in response to the target (T) and mask (M) stimuli shown in the inset. T consists of 4 units in the center of the network; M is compounded from 2 groups of 2 units each, separated by 1 unit from the edge of T. Stimulus intensity is 50% greater than ongoing background intensity. T is presented at 0 and 30 msec, and M (also presented for 2 iterations) begins at the Δt indicated. The stimuli represent long bars whose cross-sections are shown in the inset.

oscillate even at large positive Δts, and other simulated conditions show stability over a large number of stimulus conditions. Though oscillation was a serious shortcoming of the 1971 version of the LIM, the problem has been overcome with the use of a r^2 measure, a summing of iterations 1-9 and the addition of noise. The use of a broad spread of k_is also reduces the tendency for temporal oscillation.

Though Weisstein et al. (p. 334) observe that the present LIM cannot simulate stimulus durations of less than 30 msec, this interval is within the luminance-duration reciprocity of Bloch's law, so that briefer durations can be simulated by reducing stimulus intensity. This was the technique used in simulation of modifications in T/M energy ratio. Further subdivisions of time would not add to the value of the model given the variability of psychophysical masking functions. There is no theoretical limit to the temporal resolution of the LIM, however, for modeling techniques described by Lange, Hartline, and Ratliff (1966) can increase the temporal resolution of the model without changes in parameters. Weisstein et al. also point out that the 1971 version of the LIM predicts different masking functions for stimulus durations of 30 and 60 msec. The present model yields similar masking functions for the two stimulus durations with no changes in other parameters.

Spatial oscillation of the masking function with increasing T-M separation, however, is still a problem. As Weisstein et al. note, there is more masking at a T-M separation of three than at a separation of two, though the reverse effect is obtained psychophysically. A possible modification of the model would be a random variation in the k_i distribution, which is now identical in every cell, though cells even in simple animals show wide variations in k_i distribution in a single individual (Ratliff, 1965). Though every recurrent system shows oscillations under some conditions, this modification might reduce them because oscillations would not be so strongly in phase at some spatial separations and out of phase at others.

The most important point of discussion concerns the linking hypothesis, the hypothesized correspondence of the model's responses with perceptual data. Weisstein et al. note that the model will produce some masking no matter what stimulus is present in the network other than the target, though this characteristic is made less important because only a sample of visual space is represented; beyond the limits of the representation no interactions can occur. Again, a recognition system operating in a noisy environment would have to be satisfied with rs of less than 1.0.

Every model must assume a linking hypothesis, however. For Weisstein et al. "the measure of target apparent brightness is proportional to the frequency of firing of the *unit designated* as carrying information about the target" (1975, p. 337, italics added); the unanswered question is who designates the unit, and on what grounds. Breitmeyer and Ganz (1976, Figure 12) use an identical linking hypothesis, with the output of a single central

neuron representing the visibility of T. Neither model can use any other hypothesis, for each has a single output neuron and must assume that frequency of firing in a single cell codes information adequate for preception. The theoretical problem remains: How does the visual system use a cell's excitation for further processing? Evidence against the idea that a "feature-detector" cell can code a perceptually meaningful stimulus includes the ambiguity of the information contained in a single cell's response (Sekuler, 1974) and the extensive evidence for distributed coding of sensory information (reviewed by John, 1967, 1972; Pribram, 1971).

Both kinds of models make the assumption that some further neurological machinery, using the coded output of the model, makes a decision about the apparent brightness of T. This mechanism must be given a "psychological set" to recognize T; the nature of the mechanism is beyond the scope of present models and is partially represented in the LIM by the correlation process. The task of this higher mechanism is similar for both single-cell and distributed processing models, for there are many cells in the single-cell models which could potentially code the brightness of T, and the higher mechanism has no a priori way of distinguishing among them. The mechanism must look at many cells in either case.

Comparisons of the LIM and the fast-inhibition model on spatial manipulations cannot be made because the present configuration of the fast-inhibition model using two input neurons is insensitive to spatial variations in T and M, as Weisstein (1972) notes. The Weisstein et al. (1975) modification of the fast-inhibition model addresses this problem indirectly by assuming that T and M interact in activity generated by channels sensitive to different spatial frequencies. Breitmeyer and Ganz (1976) have extended this model by assuming that monotonic masking is mediated by inhibition among sustained cells, while U-shaped masking is induced by inhibition of sustained cells by transient cells. While this theory is promising, it must be considered a preliminary proposal until mathematical modeling and simulation have proved that all of the many experimental results discussed by Breitmeyer and Ganz can be accounted for with a single set of model parameters. The simulations given here show that LI among sustained cells might mediate many more of the properties of masking than Breitmeyer and Ganz assume. The units in the present LIM have many of the characteristics of sustained cells: They have small centers, large and strong inhibitory surrounds, spatial-frequency sensitivity, and sustained responses to a continuous stimulus on the receptive field center.

The LIM of Bridgeman and the sustained-transient theories of Breitmeyer-Ganz-Weisstein make different predictions in some situations; one of them is the subject of an experiment by White and Lorber (1976). Instead of using adjacent filled rectangles for T and M, White and Lorber used patches of square-wave gratings. Under this condition the sustained-transient theory presumably predicts that masking should be maximal when the

spatial frequency of the T (sustained activity) is higher than that of the M (transient activity). The LIM, however, predicts maximum masking when the spatial frequencies of T and M are the same. The experimental results support the latter prediction. Though White and Lorber interpret their results as contradictory to the Weisstein and Bridgeman models because all of the stimuli have similar edges at the T-M interface, a spatial-frequency interpretation can also follow, because the neurons in the LIM are tuned by inhibition to be maximally sensitive to one spatial frequency.

A second indirect test of the models comes from Growney (1976), who investigated metacontrast with stimuli having sine, rectangular, or edge-enhanced luminance distributions. According to the sustained-transient theory a sine stimulus, with its low-spatial-frequency components intact, should be an effective M while the LI theory predicts that sines should be less effective as T or M because cells near the edges of the stimuli would receive less stimulation. Again the results support the LI theory; the effects of the sines as Ms were small in magnitude. Breitmeyer and Ganz cite many results, however, which are consistent with their theory, and the issue must remain open until more data are available.

REFERENCES

BENEVENTO, L. A., CREUTZFELDT, O. -D., & KUHNT, U. Significance of intracortical inhibition in the visual cortex. *Nature (London)*, 1972, **238**, 124-126.

BREITMEYER, B. G., & GANZ, L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, 1976, **83**, 1-36.

BRIDGEMAN, B. Metacontrast and lateral inhibition. *Psychological Review*, 1971, **78**, 528-539.

BRIDGEMAN, B. Lateral interactions and visual coding. Unpublished doctoral dissertation, Stanford University, 1971. *Dissertation Abstracts*, 1972, **326**, 7333.

BRIDGEMAN, B. Correlates of metacontrast in single cells of the cat visual system. *Vision Research*, 1975, **15**, 91-99.

GROWNEY, R. The function of contour in metacontrast. *Vision Research*, 1976, **16**, 253-261.

JOHN, E. R. *Mechanisms of memory*. New York: Academic Press, 1967.

JOHN, E. R. Switchboard versus statistical theories of learning and memory. *Science*, 1972, **177**, 850-864.

LANGE, D., HARTLINE, H. K., & RATLIFF, F. Dynamics of lateral inhibition in the compound eye of *Limulus*: II. In C. G. Bernhard (Ed.), *The functional organization of the compound eye*. Oxford: Pergamon Press, 1966.

LEFTON, L. A. Metacontrast: A review. *Perception & Psychophysics*, 1973, **13**, 161-171.

PRIBRAM, K. H. *Languages of the brain*. Englewood Cliffs, N.J: Prentice-Hall, 1971.

RATLIFF, F. *Mach bands: Quantitative studies on neural networks in the retina*. San Francisco: Holden-Day, 1965.

SCHILLER, P. H., & SMITH, M. Monoptic and dichoptic metacontrast. *Perception & Psychophysics*, 1968, **3**, 237-239.

SEKULER, R. Spatial vision. In M. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 25). Palo Alto: Annual Reviews, 1974.

SINGER, W., & CREUTZFELDT, O. -D. Reciprocal lateral inhibition of on- and off-center neurones in the lateral geniculate body of the cat. *Experimental Brain Research*, 1970, **10**, 311-330.

SPEKREIJSE, H., VAN DER TWEEL, L. H., & ZUIDEMA, TH. Contrast evoked responses in man. *Vision Research*, 1973, **13**, 1577-1601.

SPINELLI, D. N., & BARRETT, T. W. Visual receptive field organization of single units in the cat's visual cortex. *Experimental Neurology*, 1969, **24**, 76-98.

STIGLER, R. Chronophotische Studien über den Umgebungskontrast. *Pflügers Archiv für die gesamte Physiologie*, 1910, **134**, 365-435.

UTTAL, W. R. *The psychobiology of sensory coding*. New York: Harper & Row, 1973.

WEISSTEIN, N. A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychological Review*, 1968, **75**, 494-521.

WEISSTEIN, N. Metacontrast. In D. Jameson & L. M. Hurvich (Eds.), *Handbook of sensory physiology* (Vol. 7/4, *Visual psychophysics*). New York: Springer-Verlag, 1972.

WEISSTEIN, N., OZOG, G., & SZOC, R. A comparison and elaboration of two models of metacontrast. *Psychological Review*, 1975, **82**, 325-343.

WHITE, C. W., & LORBER, C. Spatial-frequency specificity in visual masking. *Perception & Psychophysics*, 1976, **19**, 281-284.