Understanding the Efficiency of Human Perception

William Bialek[1,2] and A. Zee[1]
[1] *Institute for Theoretical Physics, University of California, Santa Barbara, Santa Barbara, California 93106*
[2] *Departments of Physics and Biophysics, University of California, Berkeley, California 94720*
(Received 14 September 1987)

We compare the high perceptual efficiency exhibited by humans at a number of tasks in visual perception with that predicted by models of visual information processing, in particular the broad class of "feature detector" models, which we represent as an Ising spin system responding to a quasirandom magnetic field. The observed excitatory-center, inhibitory-surround organization of the neural receptive fields is seen to be desirable.

PACS numbers: 87.10.+e, 05.50.+q, 42.66.Si, 89.70.+c

The sensory systems of higher animals receive an enormous amount of information from the environment. It seems natural to assume that much of this "information" is discarded by the nervous system as it extracts the "meaningful" bits. This hypothesis is supported by the qualitative description of sensory neurons as encoding discrete features,[1] features which are presumably not sufficient to fully reconstruct the input signal. The idea of feature extraction has also led to practical algorithms for electronic processing of images and sound. Several recent experiments, however, suggest that there are several tasks in visual perception where humans use all or nearly all of the information available.[2,3] We believe that this evidence of optimal performance, while still somewhat scattered, is providing an important clue about the computational structure[4] of human vision. In this note we make some first attempts at testing the classical notion of feature detection against these more recent observations.

First recall what optimal performance means in the simple task of detecting a known signal against a background of Gaussian noise. For definiteness consider only black-and-white (including gray) pictures, described by some scalar field in two dimensions $\phi(x)$. A signal $\phi_0(x)$ in a background of white noise defines an ensemble of signals with probability distribution

$$P[\phi(x)\,|\,\phi_0(x)] = Z^{-1}\exp\left(-\tfrac{1}{2}\gamma\int d^2x[\phi(x)-\phi_0(x)]^2\right), \tag{1}$$

where γ^{-1} measures the strength of the noise and Z is a normalization constant. In a typical experiment an observer will be asked to decide if the signal $\phi_0(x)$ is actually present in a particular image $\phi(x)$. More precisely,[5] on each trial of the experiment we randomly choose an image $\phi(x)$ either from the "signal" distribution $P[\phi(x)\,|\,\phi_0(x)]$ or from the "noise" distribution $P[\phi(x)\,|\,\phi_0(x)=0]$; the choice of signal versus noise is also random and for simplicity we consider the case of equal probabilities for these two alternatives. The observer, having seen $\phi(x)$, must choose between signal and noise. The fraction of correct choices is maximized if the observer bases the decision on the discriminant

$$\lambda_{\rm opt}[\phi(x)] = \ln\left\{\frac{P[\phi(x)\,|\,\phi_0(x)]}{P[\phi(x)\,|\,\phi_0(x)=0]}\right\}, \tag{2}$$

assigning $\phi(x)$ to $\phi_0(x)$ if $\lambda[\phi(x)]>0$ and to $\phi_0(x)=0$ if $\lambda[\phi(x)]<0$. This particular decision rule is termed maximum likelihood and clearly makes use of all the available information.

In signal detection theory,[5] a generalized signal-to-noise ratio (SNR) is defined by

$$(d')^2 = \frac{(\langle\lambda[\phi]\rangle_{\phi_0} - \langle\lambda[\phi]\rangle_{\phi_0=0})^2}{\{\langle(\delta\lambda)^2\rangle_{\phi_0} + \langle(\delta\lambda)^2\rangle_{\phi_0=0}\}/2}; \tag{3}$$

$$\langle\lambda^n[\phi]\rangle_{\phi_0} = \int \mathcal{D}\phi\,\lambda^n[\phi(x)]P[\phi(x)\,|\,\phi_0(x)]$$

and $\langle(\delta\lambda)^2\rangle = \langle\lambda^2\rangle - \langle\lambda\rangle^2$, with $\lambda[\phi(x)]$ as defined in Eq. (2). This definition evidently measures how far apart the two probability distributions

$$\mathcal{P}(\lambda;\phi_0) \equiv \int \mathcal{D}\phi\,\delta(\lambda-\lambda[\phi(x)])P[\phi(x)\,|\,\phi_0(x)];$$

$$\mathcal{P}(\lambda;\phi_0=0) \equiv \int \mathcal{D}\phi\,\delta(\lambda-\lambda[\phi(x)])P[\phi(x)\,|\,\phi_0=0]$$

are relative to the average of their two widths. In simple situations, such as the signal in Gaussian white-noise problem as described by Eq. (1), it is a standard exercise to show that the maximum fraction of correct signal versus noise discriminations is

$$P_c^{\max} = \tfrac{1}{2}[1+\Phi(d'/2)], \tag{4}$$

where $\Phi(z) = (2/\pi)^{1/2}\int_0^z dx\,\exp(-x^2/2)$ and

$$(d')^2_{\rm opt} = \gamma\int d^2x\,\phi_0^2(x). \tag{5}$$

$(d')^2$ is optimal in the sense that all the information contained in $\phi(x)$ is utilized. (For this simple task, obviously, only the "power" in ϕ_0 enters.) It is clear that d' corresponds to the usual notion of SNR. In more complex problems, at least for small SNR, $(d')^2$ gives us a good estimate of P_c and hence the discrimination performance. It is thus conventional to report the results of experiments on human observers in terms of an apparent

d', inverting Eq. (3). This may be used to define the experimental detection efficiency $\epsilon_{\rm expt} \equiv (d')^2_{\rm expt}/(d')^2_{\rm opt}$ with $(d')^2_{\rm opt}$ as defined in Eq. (5). It is this quantity which has been reported[2,3] to range from 0.5 to 0.95 in tasks as varied as simple signal-noise discrimination, detection of symmetry in weakly correlated random-dot patterns, and discrimination of density variations in random dots.

In general we do not expect that $\phi(x)$ itself is represented in the nervous system. Instead there is some set of variables $\{\sigma_\mu\}$ which are related to the stimulus $\phi(x)$ by a distribution $\tilde{P}(\sigma|\phi)$ and hence to the true signal by

$$P(\sigma|\phi_0) = \int \mathcal{D}\phi\, \tilde{P}(\sigma|\phi) P[\phi(x)|\phi_0(x)].$$

Both the character of the internal representation $\{\sigma\}$—e.g., discrete or continuous variables—as well as the rule for transforming $\phi \rightarrow \sigma$ [which is just $\tilde{P}(\sigma|\phi)$] must be specified in any concrete model of the perceptual process. Once we have this model, however, we can predict its limiting performance at signal-noise discrimination, since by the the arguments above the best that one can do is to form

$$\lambda_{\rm approx}[\{\sigma_\mu\}] = \ln\left\{\frac{P[\{\sigma_\mu\}|\phi_0(x)]}{P[\{\sigma_\mu\}|\phi_0(x)=0]}\right\} \qquad (6)$$

and decide that the signal is present if $\lambda_{\rm approx} > 0$. This procedure allows us to characterize any approximate model of perception by a value of d' computed with $\lambda_{\rm approx}$ and hence by an efficiency ϵ.

In general the calculation of perceptual efficiency can only be done by rather tedious Monte Carlo simulations. In the limit of small SNR, however, we can make some analytical progress, and we know that this limit is (a) accessible experimentally and (b) exact for all SNR in the optimal case of Eq. (4).

To reach this limit we expand

$$P(\sigma|\phi_0) = P| + \frac{\delta P}{\delta \phi_0}\bigg|\phi_0 + \tfrac{1}{2}\phi_0 \frac{\delta^2 P}{\delta\phi_0\delta\phi_0}\bigg|\phi_0 + \cdots;$$

here and in what follows, P always stands for $P(\sigma|\phi_0)$. We find easily that

$$(d')^2 = \int d^2y\, d^2z\, \phi_0(y)\phi_0(z) \mathrm{Tr}_\sigma\{P[\delta \ln P/\delta\phi_0(y)]\delta \ln P/\delta\phi_0(z)\}|_{\phi_0=0}, \qquad (7)$$

provided that $\mathrm{Tr}_\sigma(\delta P/\delta\phi_0)| = 0$ as in the examples below. The occurrence of $\ln P$ may be handled by the replica trick, $\ln P = \lim_{n\rightarrow 0} n^{-1}(P^n - 1)$, familiar from the theory of spin-glasses. Alternatively and more simply, we carry out the variation of $\ln P$ against ϕ_0 and analytically continue $1/P$ as $\lim_{n\rightarrow 0} P^{n-1}$,

$$(d')^2 = \lim_{n\rightarrow 0} \int d^2y\, d^2z\, \phi_0(y)\phi_0(z) \mathrm{Tr}_\sigma\{P^{n-1}(\delta P/\delta\phi_0)\delta P/\delta\phi_0\}|_{\phi_0=0}. \qquad (8)$$

To proceed further we must give some specific examples for the representation $\phi \rightarrow \sigma$. Consider the feature-detector theory[1] which originated in the neurophysiological experiments of the 1950's. Neurons in the visual system are assumed to compute nonlinear functionals of the image intensity and thus signal the presence of features in the image. We attempt to capture the essence of this theory by taking the simplest possibility for the feature tokens: Ising spins σ_μ located at x_μ, $\mu = 1, 2, \ldots, N$, with $\sigma_\mu = \pm 1$, which track the sign of the filtered image, so that

$$\tilde{P}[\{\sigma\}|\phi(x)] = \prod_\mu \theta\left[\sigma_\mu \int d^2x\, f(x - x_\mu)\phi(x)\right], \qquad (9)$$

with $\theta(x)$ the unit step function. Thus, the Ising spins at x_μ respond to a quasirandom "magnetic" field $\phi_\mu \equiv \int d^2x\, f(x-x_\mu)\phi(x)$ fluctuating around some mean $\phi_{0\mu} \equiv \int d^2x\, f(x-x_\mu)\phi_0(x)$ determined by the image $\phi_0(x)$. Here $f(x)$ represents the response function of a feature-detector neuron located at x_μ. The response function, with its typical excitatory center and inhibitory surround, is often modeled[6] as the Laplacian of a Gaussian $\nabla^2 G$ or as the difference of two Gaussians.

A classic feature-detector idea concerns the extraction of edges,[1] contours of maximal gradients of ϕ. This concept was formalized by Marr and others as the location of contours where some appropriately filtered version of the image vanishes.[7] Since the "domain walls" between spin-up and spin-down regions mark the zero-crossing contours, this $\{\sigma\}$ representation contains more information (but perhaps only a bit more) than a "sketch" based on zero-crossing contours alone. We believe that this Ising-type model is prototypical of a large family of models which replace the continuous image $\phi(x)$ by discrete and local "feature tokens." [More generally, we can incorporate noise in the visual system by writing

$$\tilde{P}(\sigma|\phi) = \prod_\mu e^{\beta\sigma_\mu\phi_\mu - \ln 2\cosh\beta\phi_\mu},$$

where β, a measure of noise in the visual system, may be thought of as an inverse temperature. The ln cosh term required by normalization makes it difficult to evaluate $(d')^2$. Thus, we go to the zero-temperature or deterministic limit. Of course, a nonzero temperature will only reduce $(d')^2$ further.]

We can directly combine Eqs. (8) and (9) to find

$$(d')^2 = \lim_{n\rightarrow 0} \mathrm{Tr}_\sigma \prod_{\alpha=1}^{n+1} \frac{1}{Z}\int \mathcal{D}\phi^\alpha \prod_\mu \theta(\sigma_\mu\phi_\mu^\alpha) e^{-(\gamma/2)\int(\phi^\alpha)^2} \times \left[\gamma\int \phi^1\phi_0\right]\left[\gamma\int\phi^2\phi_0\right]. \qquad (10)$$

Note the replica index α which runs from 1 to $n+1$.

Thanks to the replica trick, the trace over spin states can now be taken. Obviously,

$$\mathrm{Tr}_\sigma \prod_\mu \prod_a \theta(\sigma_\mu \phi_\mu^a) = \prod_\mu \left[\prod_a \theta(\phi_\mu^a) + \prod_a \theta(-\phi_\mu^a) \right]. \quad (11)$$

Multiplying this expression out, we find a sum of terms of the form $\prod_a \prod_\mu \theta(\epsilon_\mu \phi_\mu^a)$, where we have introduced a set of signs $\epsilon_\mu = \pm 1$. The signs at different sites are now coupled while the ϕ's with different replica indices factor nicely. Thus, we have

$$(d')^2 = \lim_{n\to 0} \sum B^{n-1} C^2 = \sum C^2/B, \quad (12)$$

where the $\sum$ symbol ranges over the 2^N ways in which the signs ϵ_μ can be assigned. We have written

$$B = Z^{-1} \int \mathcal{D}\phi \prod_\mu \theta(\epsilon_\mu \phi_\mu) e^{-(\gamma/2)\int \phi^2} \quad (13a)$$

and

$$C = Z^{-1} \int \mathcal{D}\phi \prod_\mu \theta(\epsilon_\mu \phi_\mu) e^{-(\gamma/2)\int \phi^2} \gamma \int \phi \phi_0. \quad (13b)$$

In this case the replica limit can evidently be taken without any fuss. Using an integral representation of the step function we find

$$B = Z^{-1} \int \mathcal{D}\phi \prod_\mu \int_0^\infty dw_\mu \int_{-\infty}^{\infty} \frac{dt_\mu}{2\pi} \exp\left[i \sum_\mu t_\mu (w_\mu - \epsilon_\mu \phi_\mu) \right] \exp\left[-\frac{\gamma}{2} \int \phi^2 \right]$$

$$= \prod_\mu \int_0^\infty dw_\mu \int_{-\infty}^{\infty} \frac{dt_\mu}{2\pi} \exp\left[i \sum_\mu t_\mu w_\mu - \frac{1}{2\gamma} \sum_{\mu\nu} \epsilon_\mu t_\mu J_{\mu\nu} \epsilon_\nu t_\nu \right] = \left(\frac{\gamma}{2\pi} \right)^{N/2} (\det K)^{1/2} \prod_\mu \int_0^\infty dw_\mu \exp\left[-\frac{\gamma}{2} \sum_{\mu\nu} \epsilon_\mu w_\mu K_{\mu\nu} \epsilon_\nu w_\nu \right]. \quad (14)$$

Here $J_{\mu\nu} = J(x_\mu - x_\nu) = \int d^2x\, f(x - x_\mu) f(x - x_\nu)$ and K is the inverse of J. This describes a rather peculiar system in which fields w_μ whose values are restricted to be positive interact with each other via an interaction $\epsilon_\mu K_{\mu\nu} \epsilon_\nu$, with the ϵ's alternately making each interaction ferromagnetic or antiferromagnetic. C is given similarly. We finally have

$$(d')^2 = \sum \frac{\{D \prod_\mu \int_0^\infty dw_\mu \exp[-(\gamma/2) \sum \epsilon w K \epsilon w] \gamma \sum_{\lambda\lambda'} \phi_{0\lambda} K_{\lambda\lambda'} \epsilon_{\lambda'} w_{\lambda'}\}^2}{D \prod_\mu \int_0^\infty dw_\mu \exp[-(\gamma/2) \sum \epsilon w K \epsilon w]}, \quad (15)$$

where $D = (\gamma/2\pi)^{N/2} (\det K)^{1/2}$.

Note that $(d')^2$ has the form $(d')^2 = \gamma \int d^2x \times d^2y\, \phi_0(x) G(x,y) \phi_0(y)$ as is required by dimensional analysis. Comparing with the optimal d' in Eq. (5), we see that the efficacy of feature detection is measured by how closely $G(x,y)$ reproduces the delta function. More precisely, considering $G(x,y)$ as a kernel with eigenvalues λ and eigenfunctions $\psi_\lambda(x)$ we can write

$$(d')^2 = \gamma \int d\lambda\, \rho(\lambda) \lambda \left[\int d^2x\, \phi_0(x) \psi_\lambda(x) \right]^2 \quad (16)$$

as an integral over the eigenvalue spectrum (described by the density). If the integral is dominated by the largest eigenvalue, we see that roughly it is the overlap of ϕ_0 with a "canonical" image $\psi_\lambda(x)$ that counts. In general, however, the integral will not be so dominated.

We see no way of doing the integral over w_μ in general. The following approximation helps to indicate what may be going on. With the normalization $\int d^2x f^2(x) = 1$ we can write $J_{\mu\nu} = \delta_{\mu\nu} + A_{\mu\nu}$, where $A_{\mu\nu}$ is a symmetric matrix with diagonal elements zero. Evaluating $(d')^2$ perturbatively in A, we find the kernel

$$G(x,y) = \frac{2}{\pi} \sum_{\lambda\sigma} f(x - x_\lambda) f(y - x_\sigma) \left[\delta_{\lambda\sigma} - \frac{2}{\pi} A_{\lambda\sigma} + \cdots \right]. \quad (17)$$

In the "local" limit $A \to 0$, we have $G(x,y) = (2/\pi) \times \sum_\lambda f(x - x_\lambda) f(y - x_\lambda)$. If the f's form a complete set then this most closely approaches a delta function, maximizing d'. Thus in this limit the maximum efficiency is $2/\pi = 0.64\ldots$, which is excluded experimentally.[2,3] We conclude that off-diagonal terms in $J_{\mu\nu}$, corresponding to overlaps of the "receptive fields" of neighboring cells, are essential for understanding the observed efficiency of human perception. Further, these overlaps must be negative to enhance d', which necessitates an antagonistic center-surround type of organization as found experimentally at various levels of the visual system.[1]

Detailed analysis of Eq. (17) to determine the conditions for $\epsilon \to 1$ seems unprofitable, since this result is only valid at small A. An alternative approach is to assume that the decision variable $\lambda_{\mathrm{approx}}[\{\sigma\}]$ is some general linear function $\lambda_{\mathrm{approx}} = \sum_\mu g_\mu \sigma_\mu$, with the coefficients g_μ chosen to maximize d'. This corresponds to a literal notion of feature *detection* as a strategy in signal detection—the presence or absence of the signal is judged by a tallying of the presence or absence of the features coded by $\{\sigma_\mu\}$. With this approach we can make considerable progress. In the limit of small ϕ_0 as above we find the maximum d' (optimizing g_μ) to be

$$(d')^2 = \gamma \int d^2x \int d^2y\, \phi_0(x) \phi_0(y) \frac{2}{\pi} \sum_{\mu\nu} f(x - x_\mu) (S^{-1})_{\mu\nu} f(y - x_\nu), \quad (18)$$

where the matrix $S_{\mu\nu} = \mathrm{Tr}_\sigma (\sigma_\mu \sigma_\nu P[\{\sigma_\mu\} \,|\, \phi_0 = 0])$ can be expressed in terms of formulas like Eq. (14), but with integrals over only two variables w_μ since we need to keep track only of one pair of spins. As a result the problem can be solved exactly, to give

$$S_{\mu\nu}=\frac{4}{\pi}\tan^{-1}\left[\left(\frac{1+J_{\mu\nu}}{1-J_{\mu\nu}}\right)^{1/2}\right]-1. \tag{19}$$

In the limit $J_{\mu\nu}=\delta_{\mu\nu}+A_{\mu\nu}$, $A\ll 1$, we recover Eq. (17) exactly, thus strongly suggesting that simple linear combinations of spins provide *all* of the information for signal detection available in the set $\{\sigma_\mu\}$ in the local limit.

We can also study the performance of the "linear combination" model in the continuum limit, where $S_{\mu\nu}\to S(x_\mu - x_\nu)$ and sums over lattice sites become integrals. After some calculation we find

$$(d')^2=\frac{2}{\pi}\gamma\int\frac{d^2k}{(2\pi)^2}|\tilde{\phi}_0(k)|^2\frac{\tilde{J}(k)}{\tilde{S}(k)}, \tag{20}$$

where $\tilde{J}(k)=|\tilde{f}(k)|^2$ is positive definite and we have the transform convention $\tilde{f}(k)=\int d^2x\, e^{-ikx}f(x)$. Expanding Eq. (19), we have

$$\tilde{S}(k)=\frac{2}{\pi}\left[\tilde{J}(k)+\frac{1}{6}\int\frac{d^2q}{(2\pi)^2}\int\frac{d^2q'}{(2\pi)^2}\tilde{J}(q)\tilde{J}(q')\tilde{J}(k-q-q')+\cdots\right]. \tag{21}$$

If $\tilde{J}(k)$ were small for all k, we would apparently reach optimal performance, $(d')^2_{\rm opt}=\gamma\int d^2k(2\pi)^{-2}|\tilde{\phi}_0(k)|^2$ [Eq. (4)]. The normalization condition $\int d^2k(2\pi)^{-2}\times\tilde{J}(k)=\int d^2x f^2(x)=1$ forces $\tilde{S}(k)$ to be larger than $(2/\pi)\tilde{J}(k)$ and hence the detection efficiency is less than unity, as it must be.

What is remarkable is how close we can come to optimality. As an example we consider approximating $\tilde{J}(k)$ by Gaussians of width $(\Delta k)^2$ centered at $\pm k_0$,

$$\tilde{J}(k)=A[e^{-(k-k_0)^2/2(\Delta k)^2}+e^{-(k+k_0)^2/2(\Delta k)^2}],$$

as would be the case if the response function is the Laplacian of a Gaussian.[8] The constant A is determined by the normalization $\int d^2k(2\pi)^{-2}\tilde{J}(k)=1$. For $\Delta k \ll k_0$, we have $A\sim\pi(\Delta k)^{-2}$, and after some algebra we find $\tilde{S}(k_0)\sim(2/\pi)A(1+\frac{1}{24}+\cdots)$. We conclude that for signals whose power is concentrated near k_0, this system can achieve a detection efficiency $\epsilon\sim(1+\frac{1}{24}+\cdots)^{-1}$ within a few percent of unity for some optimal ϕ_0. Recall that the coefficients g_μ were chosen to optimize performance. This result corresponds well with expectations from theorems on the information content of zero crossings in narrow-band signals.[9]

To summarize, we have shown explicitly how feature-detector models can be tested against the observed near optimality of human performance at specific perceptual tasks. To the extent that perceptual efficiency is understandable in these models it appears that an antagonistic center-surround organization of the receptive fields is essential. It appears possible to get the most out of a feature-detector scenario even with relatively simple linear processing of the detector outputs, and within this scheme we have found expressions for the perceptual efficiency which are easily evaluated for specific models. In particular, the limit of "spatial frequency detection"—very-narrow-bandwidth receptive fields—seems to allow efficiencies within a few percent of unity, in marked contrast to the untenable broadband ($J_{\mu\nu}\sim\delta_{\mu\nu}$) result of $\epsilon=2/\pi$. It is tempting to conclude that a realistic feature-detector model of the visual system exists somewhere between these limits. It may be safer to conclude that simple signal-detection tasks do not provide a definitive test of the feature-detector concept. We are currently studying the performance of these models at more complex tasks.

This work was supported by the National Science Foundation under Grant No. PHY82-17853, supplemented by funds from NASA, by the Miller Institute for Basic Research in Science, and by a Presidential Young Investigator Award (to W.B.) from the NSF.

[1]S. W. Kuffler, J. G. Nicholls, and A. R. Martin, *From Neuron to Brain* (Sinauer, Sunderland, MA, 1984), 2nd ed.; M. W. Levine and J. M. Shefner, *Fundamentals of Sensation and Perception* (Random House, New York, 1981); *Recognition of Pattern and Form,* edited by D. G. Albrecht (Springer-Verlag, New York, 1982).

[2]For a general discussion of these experiments see H. B. Barlow, Philos. Trans. Roy. Soc. London, Ser. B **290**, 71 (1980); and a special issue of J. Opt. Soc. Am. A, "The Statistical Efficiency of Natural and Artificial Vision" (to be published).

[3]A. E. Burgess, R. F. Wagner, R. J. Jennings, and H. B. Barlow, Science **214**, 93 (1981).

[4]W. Bialek and A. Zee, Phys. Rev. Lett. **58**, 741 (1987); see also W. Bialek, Annu. Rev. Biophys. Chem. **16**, 455 (1987); A. Zee, Institute of Theoretical Physics, University of California, Santa Barbara, Report No. NSF-ITP-87-122, 1987 (to be published).

[5]D. M. Green and J. A. Swets, *Signal Detection Theory in Psychophysics* (Krieger, New York, 1966).

[6]For example, Levine and Shefner, Ref. 1, p. 116ff; A. Parker and M. Hawker, J. Opt. Soc. Am. **2**, 1101 (1985).

[7]D. Marr, *Vision* (Freeman, San Francisco, 1982); T. Poggio, in *Recognition of Pattern and Form,* edited by D. G. Albrecht (Springer-Verlag, New York, 1982).

[8]The function $\tilde{f}(k)$, being equal to k^2 times a Gaussian, has two peaks and vanishes at the origin. We approximate its square by the sum of two Gaussians.

[9]B. F. Logan, Jr., Bell Syst. Tech. J. **56**, 487 (1977).