# Detection of noisy visual targets: Models for the effects of spatial uncertainty and signal-to-noise ratio

RICHARD G. SWENSSON and PHILIP F. JUDY
*Harvard Medical School and the Brigham and Women's Hospital, Boston, Massachusetts 02115*

An "extreme-detector" model for detecting spatially uncertain targets in noisy backgrounds predicts how both detection and localization abilities are degraded by increasing the number of possible target locations. Experiments 1 and 2 show that the model accurately predicts detection and localization performance in tasks with two, four, and eight locations from d' estimates of the observer's ability to detect the target in a known spatial location. These predictions can be linked to the physical stimuli by combining the extreme-detector model with a "psychophysical" model that specifies how stimulus measures determine the target's detectability in a given location. Single-parameter fits of four such combined models were compared with estimates of detection and localization performance in Experiment 3, which manipulated the target's physical signal-to-noise ratio across various conditions of an eight-location task.

Situations that require the detection of visual targets often include considerable uncertainty about the spatial location in which the target might appear. For example, a radiologist who interprets a radiographic image must decide whether or not a lesion is present in any one of many possible anatomic locations. An air-traffic controller who monitors a visual display may need to determine when an aircraft has entered the radar field from any direction. Many of these situations include sources of physical noise that would limit the detectability of the visual signal, even if there were no uncertainty about precisely where the target might appear on the visual display. In such noise-limited situations, any increase in the target's spatial uncertainty will degrade an observer's visual performance, even if there are no time constraints that prevent adequate attention to the information from all relevant spatial locations. Intuitively, the degradation in performance occurs because each additional spatial location that the observer must consider increases his opportunity to encounter a sample of noise extreme enough to resemble an actual target.

For a target that may appear in any one of $m$ distinct spatial locations, an increase in $m$ will reduce both the observer's ability to distinguish cases when the target is present from cases when it is not (target detectability) and his ability to identify the correct location of the target. The following section develops an "extreme-detector" model for performance under spatial uncertainty, which assumes no limita-

tion on the observer's ability to process visual information from additional spatial locations. This model predicts a specific relation between measures of target-detection and target-localization performance and also predicts how these measures should vary as a function of $m$, the number of distinct spatial locations in which the target might appear.

The three experiments applied the extreme-detector model to observers' performance with noise-limited visual images viewed without any time pressure. Experiments 1 and 2 presented detection and localization tasks, separately and in combination, varying the number of possible locations for the visual target from one to eight. Experiment 3 used several types of stimulus manipulations to vary the physical detectability of the target, as measured by the "signal-to-noise" ratio (S/N) calculated for a matched filter. Detection and localization performances in 10 conditions of an eight-location task were compared with predictions the extreme-detector model made when combined with four separate "psychophysical" models. Each of these combined models used a single fitted parameter to specify the assumed relation between measures defined on the physical stimuli and the target's detectability ($d'$) in a known location on the image.

## The Extreme-Detector Model

The task of detecting and locating a target that can appear at any one of $m$ distinct spatial locations may be treated as a particular type of combined detection and identification task, in which both the detection and identification (localization) decisions are based upon the same visual information. One formal model for such a task represents the observer's stimulus in-

0031-5117/81/060521-14$01.65/0

formation as a set of m orthogonal random variables $(x_i, i = 1, \ldots, m)$ that are defined by the output of m filters, each separately tuned to detect the target in one of the possible locations (Green & Swets, 1974; Green & Birdsall, 1978). The distribution of $x_i$ is assumed to be identical for all m locations $(i = 1, \ldots, m)$ and independent of the value of m. The model assumes that the observer's detection and location decisions depend only upon the most-extreme sample of x obtained for the m random variables. This extreme value determines both the location he identifies as "most likely" to contain the target and his rating of the absolute likelihood that a target is present. Whether this "extreme detector" considers the maximum or the minimum sampled value depends upon whether the target tends to increase or decrease $x_i$ when it appears in the ith location.

## Target Detection

The performance of an extreme detector in an m-location task can be expressed in terms of the detection performance obtained when the target appears in a known location (m = 1). Consider the extreme detector's ability to differentiate cases when the target is present in one of the m locations from cases when it is absent, which ignores the accuracy of the location decisions. This detection performance can be described by the receiver operating characteristic (ROC) curve for the m-location task. The m-location ROC curve is defined by the covariation in $1 - \beta_m$ and $\alpha_m$, the conditional probabilities of true-positive and false-positive decisions (given the presence and absence of a target) as a function of the criterion for making a positive decision. However, the performance of an extreme detector is easier to calculate in terms of $1 - \alpha_m$ and $\beta_m$, the complementary probabilities of true and false *negative* decisions.

The extreme detector will make a negative decision only if all m of the observed variables fail to reach the (minimum or maximum) cutoff for a positive response, which is assumed to be constant for all m locations. Since the m variables are identically distributed and mutually orthogonal, this means that the respective conditional probabilities of the true-negative and false-negative decisions in an m-location task will be given by:

$$1 - \alpha_m = (1 - \alpha_1)^m \qquad (1a)$$

and

$$\beta_m = \beta_1 (1 - \alpha_1)^{m-1}, \qquad (1b)$$

where $1 - \alpha_1$ and $\beta_1$ are the respective conditional probabilities of negative decisions given the absence and the presence of a target in any particular individual location (m = 1).

Equations 1a and 1b can be used to predict the ROC curve for detection in any m-location task from the

ROC curve for detection performance in a known-location task (m = 1). This prediction does not depend upon any specific assumptions about the form of the known-location ROC curve or the distributions of the underlying random variables. In certain cases, however, it will closely approximate the detection performance achieved by an optimal procedure—which would be based upon the joint likelihood ratio for the entire set of m sampled variables, rather than simply the most-extreme value (Green & Birdsall, 1978).

If $x_i$ has a normal distribution with mean $\mu_t$ or $\mu_n$ and standard deviation $\sigma_t$ or $\sigma_n$, depending upon whether a target is or is not present in the ith location, then the ROC curve for detection in any single location is given by:

$$z^{-1}(1 - \beta_1) = Bz^{-1}(\alpha_1) + A, \qquad (2)$$

where $B = \sigma_n/\sigma_t$ and $A = (|\mu_n - \mu_t|)/\sigma_t$, and where $z^{-1}(1 - \beta_1)$ and $z^{-1}(\alpha_1)$ are the deviates of the standard normal distribution whose (lower-tail) areas are equal to $1 - \beta_1$ and $\alpha_1$, respectively (Green & Swets, 1974). In the case that $\sigma_n = \sigma_t \equiv \sigma$ (i.e., B = 1), Equation 2 becomes:

$$z^{-1}(1 - \beta_1) = z^{-1}(\alpha_1) + d', \qquad (3)$$

and the ROC curve is specified by the single parameter $d' \equiv (|\mu_n - \mu_t|)/\sigma$. In this latter case, Nolte and Jaarsma (1967) have shown that the extreme detector's ROC curve very closely approximates the ROC curve for the optimal (joint likelihood-ratio) procedure in an m-location task. Over a wide range in both $d'$ and m, these m-location ROC curves are almost linear in normal-deviate coordinates—that is, when the values of $1 - \beta_m$ and $\alpha_m$ are transformed into $z^{-1}(1 - \beta_m)$ and $z^{-1}(\alpha_m)$, their equivalent deviates of the standard normal distribution.

An index of the target's detectability in an m-location task can be defined by $d_\varepsilon \equiv 2z^{-1}(1 - \varepsilon_m)$, which characterizes the ROC curve by its point of intersection with the negative diagonal, when both types of detection errors are equiprobable ($\alpha_m = \beta_m \equiv \varepsilon_m$). This index can be interpreted as a perceptual "signal-to-noise ratio" for the m-location detection process. For detection in a known location (m = 1), setting $\alpha_1 = \beta_1 \equiv \varepsilon_1$ in Equation 2 shows that $d_\varepsilon = 2A/(B + 1) = 2(|\mu_n - \mu_t|)/(\sigma_n + \sigma_t)$. Thus, the mean of $\sigma_n$ and $\sigma_t$ represents the "noise" in the known-location detection process for the index $d_\varepsilon$, which reduces to $d'$ when $\sigma_n = \sigma_t$. An analogous interpretation of $d_\varepsilon$ can be made for the extreme-detection process in an m-location task. Once the known-location ROC curve is specified, the extreme detector's $d_\varepsilon$ can be calculated for any m-location task by iterating Equations 1a and 1b until $\alpha_m = \beta_m$. Each trial value of $\alpha_m$ in the iteration implies a value for $\alpha_1$ (Equation 1a), which then provides $\beta_1$ (e.g., by Equation 2) and the corresponding value of $\beta_m$ (Equation 1b). Figure 1a shows how $d_\varepsilon$
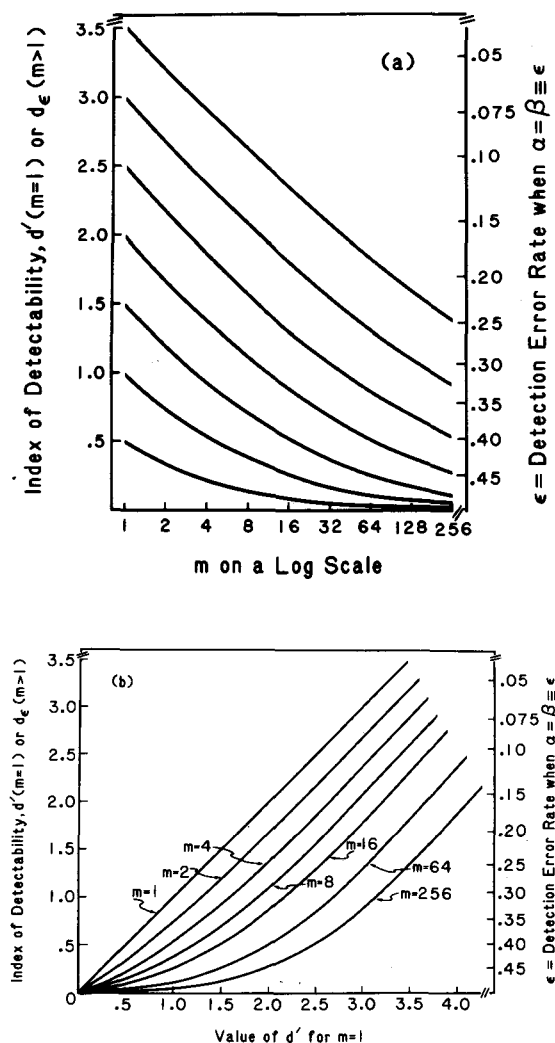
Figure 1. Values of the $d_\varepsilon$ index of detectability for the extreme-detector model in m-location tasks: (a) plotted as a function of m on a log scale for various values of d' in the m = 1 task, and (b) plotted as a function of d' in the m = 1 task for various values of m. The right ordinate gives values of the detection error rates that correspond to each value of $d_\varepsilon$ at the point on the m-location ROC curve where $\alpha = \beta \equiv \varepsilon$.

decreases as a function of m for various values of d' when the known location ROC curve (m = 1) is given by Equation 3; Figure 1b shows how $d_\varepsilon$ increases with d' for various values of m.

## Target Localization

The ROC curve for an m-location task ignores which locations were identified as the target and simply describes how well the observer's detection decisions separate the cases with a target in one of the m locations from cases with no target in any location. Similarly, the observer's ability to locate the target can be described independently of his detection decisions by P(CL), the probability of correctly identifying the target's location when it is present. This would be

the obvious measure of performance in a "pure localization" (forced choice) task that simply required the observer to identify which of m locations contained the target. For an extreme detector, however, detection and localization capabilities are perfectly related and the accuracy of location decisions can be specified in terms of the detection ROC curve for a given m-location task.

When both detection and location decisions are considered in an m-location task, the true-positive decisions can be separated into two mutually-exclusive categories: those that correctly identify the target's location and those that misidentify the target. An extreme detector misidentifies the target whenever one of the m − 1 nontarget locations gives rise to a sample of "noise" more extreme than the sample from the target's location. If $\gamma_m$ is the joint probability of making both a true-positive detection decision *and* a correct identification of the target's location, then the (complementary) joint probability of a true-positive decision and an erroneous localization is given by $(1 - \beta_m) - \gamma_m$. Since $\gamma_m$ (like $1 - \beta_m$ and $\alpha_m$) depends upon the criterion for a positive detection response, it is possible to define an "identification operating characteristic" (IOC) curve, which is analogous to the ROC curve. This IOC curve describes the functional relation between $\gamma_m$ and $\alpha_m$ over changes in the criterion for making a positive detection decision. The IOC curve lies below the corresponding ROC (since $\gamma_m \leqslant 1 - \beta_m$ for any given value of $\alpha_m$) and it ends at the point $\gamma_m \equiv$ P(CL), when $\alpha_m = 1.0$, that is, when all detection decisions are treated as "positive" and $1 - \beta_m = 1.0$.

Let $\alpha_m(\lambda)$ be the extreme detector's false-positive rate for a particular detection criterion $\lambda$, and let $1 - \beta_m(\lambda)$ and $\gamma_m(\lambda)$ be the corresponding points on the respective ROC and IOC curves. Starr, Metz, Lusted, and Goodenough (1975) showed that $\gamma_m(\lambda)$ can be specified by integrating a function of the m-location ROC curve over the values of $\alpha_m \leqslant \alpha_m(\lambda)$:

$$\gamma_m(\lambda) = 1 - \beta_m(\lambda) - \left(\frac{m-1}{m}\right) \int_0^{\alpha_m}{}^{(\lambda)} \left(\frac{\beta_m}{1-\alpha_m}\right) d\alpha_m.$$

(4)

The endpoint of the IOC curve is defined by $\gamma_m(\lambda) \equiv$ P(CL), when the integration of Equation 4 is taken over the entire ROC curve—that is, for $\alpha_m(\lambda) = 1.0$. Green, Weber, and Duncan (1977) proved that Equation 4 then reduces to $\gamma_m =$ P(CL) in the form of the "generalized area theorem":

$$P(CL) = \int_0^1 \alpha_1{}^{m-1} d(1 - \beta_1),$$

(5)

which is equal to the area below the known-location ROC curve for the special case of m = 2 (Green & Birdsall, 1978; Green & Swets, 1974).

Once the m-location ROC curve has been specified, either as a continuous function or as a discrete set of points connected by line segments, a relatively simple computer program can approximate the integral in Equation 4 to generate any desired number of points on the m-location IOC curve, including P(CL). Because the known-location ROC curve specifies the extreme detector's ROC curve in an m-location task (by Equation 1), it also determines the IOC curve and P(CL) in that task.

When the target is present in an m-location task, the particular location selected by the extreme detector does not depend upon whether or not the task also requires a detection decision. In either case, the problem of identifying the target's location is formally equivalent to an m-alternative "forced-choice" task. Under the assumption that the known-location ROC curve is given by Equation 3, there exist tables that relate P(CL) in an m-alternative task to $d'$, for various values of m (Elliot, 1964; Hacker & Ratcliff, 1979). The value $z^{-1}[P(CL)]$, the normal deviate that corresponds to P(CL), increases virtually linearly with $d'$ for all values of m (Elliot, 1964).

## EXPERIMENTS 1 AND 2

The first two experiments applied the extreme-detector model to observers' visual detection and localization performance in tasks that varied m, the number of possible target locations. In Experiment 1, one- and two-parameter fits of observers' ROC curves for target detection in a known location were used to predict the ROC and IOC curves obtained in an eight-location task, as well as estimates of P(CL) obtained in both the combined task and a separate eight-location forced-choice task. In Experiment 2, the changes in separate indices of target detection localization performance, $d_e$ and P(CL), were compared with predictions of the extreme-detector model for tasks with m = 1, 2, 4, and 8.

## Method

### Visual Stimuli

The visual stimuli were photographs of gray-scale displays produced on a cathode ray tube by a computer program that was designed to stimulate the images generated by computed tomographic scanners used in diagnostic radiology (Judy, Swensson, & Szulc, 1981). The visual display was a 64 by 64 matrix of square "pixel" elements which could assume one of 16 gray-scale values, equally spaced in luminance. The display showed an elliptical ring, 5 pixels thick, on a darker background. The target was a darker circular area, 2.5 pixels in diameter, that could be superimposed at any one of eight equally spaced positions within the ring. The "noise" in these displays represented random variation in the brightness values assigned to individual pixels; some "texture" in this noise was produced by a positive correlation of .11 between the brightness values of adjacent pixels. Figure 2 shows an example of one of the images with the target location in the bottom ("southernmost") position.
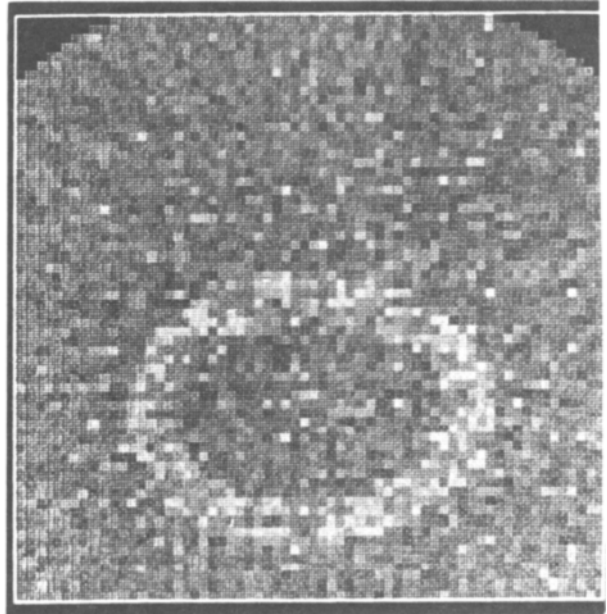


**Figure 2. An example of one of the stimulus images used in Experiments 1 and 2. The target is located at the bottom (southernmost) position in the elliptical ring.**

There were a total of 80 stimulus images, 40 images with no target in the ring and 40 images with a single target located at one of the eight possible positions (five images with the target in each of the eight positions). The images were photographed, using Polaroid 107 film, to produce hard-copy images for viewing by observers. The hard-copy images were glossy prints, 5.25 cm square, with each individual pixel measuring .82 mm.

The physical detectability of the target in these images was characterized in terms of the detection performance achievable by a "matched filter" detector, assumed to know the target's precise size and location. This was obtained by calculating the output distributions of a matched filter that integrated the brightness values of pixels at the target's location, given the presence and absence of the target (Judy, Swensson, & Szulc, 1981). The filter's output was a weighted sum of the random variables for individual pixels; it was approximately normally distributed, given the presence and absence of the target, with means $\mu_t$ and $\mu_n$ and standard deviations $\sigma_t$ and $\sigma_n$, respectively. Because of asymmetries in the transformation that mapped the numerical values into gray-scale levels, the two standard deviations were not equal; the linear ROC curve for a detector based on the matched filter's output was given by Equation 2, with parameters $B = \sigma_n/\sigma_t = .858$ and $A = (\mu_n - \mu_t)/\sigma_t = 2.452$. The target's physical "signal-to-noise ratio" was defined by S/N, with $S \equiv \mu_n - \mu_t$ and $N \equiv (\sigma_n + \sigma_t)/2$, which is equivalent to defining S/N as the value of $d_e = 2A/(B + 1)$ for the (known-location) ROC curve of the matched-filter detector. The value of S/N = $d_e$ was 2.64 for these images.

### Procedure

**Experiment 1.** The six observers were given some familiarity with these types of images before performing four separate tasks, using the same set of images. The first task presented all 80 images; it specified the location that would contain the target if it was present on that image (m = 1), and required the observer to give a 6-category rating of the likelihood that the target was present ("Yes" or "No" with a high, medium, or low level of confidence). The second task also presented all 80 images, but did not specify which of the eight locations would contain the target; the observer indicated the most likely location for the target and used a 6-category scale to rate

his confidence that the target was actually present. The third task was an eight-location forced-choice task that presented only the 40 images with targets and simply required the observer to indicate the location of the target. The fourth task replicated the first known-location task as a check on the stability of the detection performance over the experimental sessions.

The observers were aware that the target and nontarget images were equiprobable in the detection tasks, and they tried to adjust their detection criteria within each task to produce about the same number of judgments in all six rating categories. Reference diagrams showed the precise locations in the ring at which the target might appear. The observer also had a reducing lens, which extended the effective viewing distance for the image to 150 cm and facilitated the visual integration of the discrete pixels in the image. This reducing lens was deliberately used to vary the viewing distance over a large range during the inspection of each image. At the maximum effective viewing distance, the individual pixels subtended only about .03 deg of visual angle and were not separately resolvable.

**Experiment 2.** Three experienced observers from Experiment 1 were given two additional detection tasks, which specified either two or four possible locations for the target on each of the 80 images. As in Experiment 1, they indicated the most likely location for the target on each image (within $m = 2$ or $m = 4$) and rated their confidence that it was actually present on a 6-category scale. After becoming familiar with these images and targets, two additional observers performed the four detection tasks ($m = 8, 4, 2,$ and $1,$ performed in that order). The observers gave a 6-category rating of the likelihood that the target was present on each image and (except for $m = 1$) indicated which location was most likely to contain the target. Other procedural details were identical to those in Experiment 1.

## Results

### Experiment 1

**Detection at a known location.** The observers performed the detection task with no uncertainty about the target's location ($m = 1$) on two separate occasions, their first (early) and their final (late) experimental tasks. The 6-category confidence ratings about the presence of the target on each image in the early and late known-location tasks generated two known-location ROC curves for each observer, which were separately fit by both Equation 2 and Equation 3. Equation 2 was fit to each rating ROC curve by a maximum likelihood procedure (Dorfman & Alf, 1969; Grey & Morgan, 1972; Sandor & Swensson, 1978) that provided estimates of the linear slope of the ROC curve in normal-deviate coordinates (B) and its detectability index $d_\epsilon = 2A/(B + 1)$. A reduced set of likelihood equations was used with the same pro-

cedure to fit Equation 3 and to estimate d' for each ROC curve under the assumption that $B = 1.0$.

Table 1 presents values of $d_\epsilon$, B, and d' for all six observers in both their early and late known-location tasks, together with mean values of these estimates across tasks and observers. All statistical operations on estimates of B used arctan transformations, in order to correct for the asymmetry in scale between slope estimates above and below 1.0. Estimates of the separate detectability indices, $d_\epsilon$ and d', were almost identical in every case. Neither they not the estimates of B changed systematically between the early and late known-locations tasks, demonstrating that the observers' performance remained stable over the course of the experiment. The mean estimate of B for all observers (.867) was very close to .858, the slope of the linear ROC curve calculated for the performance of the matched-filter detector. However, the mean estimates of $d_\epsilon$ and d' (1.54) were considerably lower than the matched filter's value of $S/N = d_\epsilon = 2.64$.

The filled and open circles in each panel of Figure 3 show the five points on the ROC curves for each observer's early and late known-location tasks. The smooth curve drawn through both sets of points is the ROC curve defined by Equation 3, with d' specified by the mean of the two estimated values for that observer (from his early and late tasks).

**Detection with eight possible locations.** In the eight-location task, an ROC curve was generated from the confidence ratings that each observer assigned to the target and nontarget images, ignoring whether he located the target correctly or not. Estimates of $d_\epsilon$ and the linear slope of the eight-location ROC curve were provided by a maximum-likelihood fitting procedure that assumed that the true functional relation was linear in normal-deviate coordinates. According to the extreme-detector model, the eight-location ROC curve should be completely determined by the observer's ROC curve in the known-location task.

The one-parameter (d') and two-parameter ($d_\epsilon$, B) fits of each observer's known-location ROC curve provided two separate predicted ROC curves for his eight-location task. The predicted ROC curves were generated by applying Equations 1a and 1b to the values of $\alpha_1$ and $\beta_1$ obtained from either Equations 3

Table 1
Estimates of d' and of $d_\epsilon$ and B From Each Observer's ROC Curves in the Early and Late Known-Location Tasks in Experiment 1

| Observer | Estimate of d' | | | Estimate of $d_\epsilon$ | | | Estimate of B | | |
|---|---|---|---|---|---|---|---|---|---|
| | Early | Late | Mean | Early | Late | Mean | Early | Late | Mean |
| P.F.J. | 1.458 | 1.382 | 1.420 | 1.459 | 1.381 | 1.420 | .855 | .827 | .841 |
| R.G.S. | 1.753 | 2.217 | 1.985 | 1.790 | 2.228 | 2.009 | .592 | .847 | .712 |
| R.G. | 1.816 | 2.065 | 1.938 | 1.810 | 2.062 | 1.936 | .965 | .978 | .971 |
| K.P.D. | 1.343 | 1.342 | 1.343 | 1.343 | 1.347 | 1.345 | 1.012 | .838 | .921 |
| E.K. | 1.498 | 1.329 | 1.413 | 1.499 | 1.311 | 1.405 | 1.024 | .778 | .893 |
| F.B. | 1.309 | .935 | 1.122 | 1.307 | .925 | 1.116 | .925 | .847 | .885 |
| Mean | 1.530 | 1.545 | 1.537 | 1.535 | 1.542 | 1.538 | .884 | .851 | .867 |
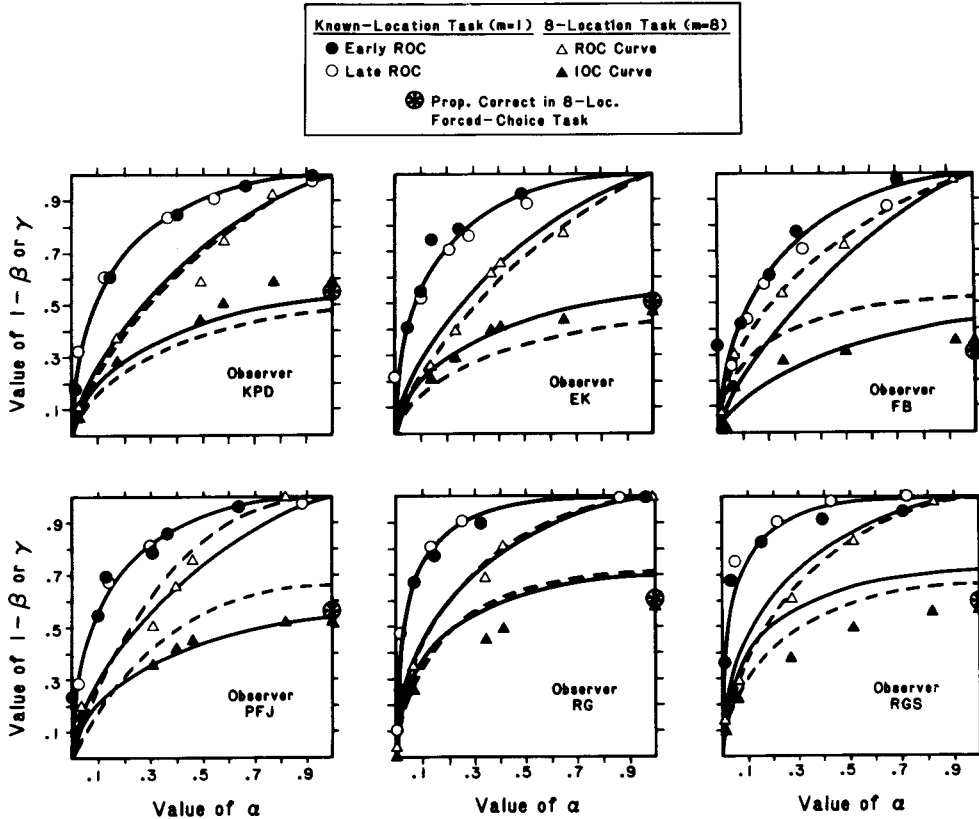
Figure 3. Known-location ROC curves and eight-location ROC and IOC curves generated from the confidence ratings of each observer in Experiment 1. The ROC curves plot $1-\beta$, the true-detection rate, as a function of $\alpha$, the false-positive rate; the solid and open circles indicate data from the early and late known-location tasks, respectively, while the open triangular points indicate the eight-location ROC curve. The IOC curves (solid triangles) plot $\gamma$, the joint proportion of targets both detected and correctly located at each rating category, as a function of $\alpha$. The solid triangular point on the right ordinate of each panel gives the value of $\gamma = P(CL)$ when $\alpha = 1 - \beta = 1.0$, the proportion of targets correctly located on all target stimuli; the circled asterisk gives the value of $P(CL)$ in the observer's eight-location forced-choice task. The three solid curves in each panel show the fitted ROC curve for the known-location task and the ROC and IOC curves predicted by the extreme-detector model for the eight-location task, based upon the mean estimate of $d'$ from the observer's early and late known-location tasks. The two broken curves in each panel show the fitted ROC curve for the eight-location task and the IOC curve predicted by the extreme-detector model from that fitted ROC curve.

or Equation 2, whose parameters were given by that observer's mean estimates (of $d'$ or of $d_\varepsilon$ and B) in the known-location tasks. Predicted values of $d_\varepsilon$ for the eight-location task were obtained by iterating Equations 1a and 1b until $\alpha_m = \beta_m = \varepsilon_m$ and calculating $d_\varepsilon = 2Z^{-1}(1 - \varepsilon_m)$. Because the predicted ROC curves were not perfectly linear in normal-deviate coordinates, the predicted slope for the ROC curve was taken to be the slope of the line connecting the "X" and "Y" normal-deviate intercepts—that is, the points on the predicted eight-location ROC curve that correspond to $1 - \beta_m = .5$ and $\alpha_m = .5$, respectively.

In Figure 3, the open triangular points in each panel show the ROC curve for each observer in the eight-location task. The broken curve through these data points is the ROC curve fitted to these data points by the maximum-likelihood procedure. The

solid curve through these data points is the eight-location ROC curve predicted by the extreme-detector model from the single-parameter ($d'$) ROC curve in the known-location task (i.e., the highest solid curve in each panel).

Table 2 presents the direct estimates of $d_\varepsilon$ and linear slope of each observer's fitted eight-location ROC curve, together with the values predicted by the extreme-detector model from the single-parameter ($d'$) and two-parameter ($d_\varepsilon$, B) fits to his known-location performance. The mean values in Table 2 show that the predictions based upon $d'$ were closer to the actual estimates of $d_\varepsilon$ and slope than were the predictions based upon the two parameter fits of $d_\varepsilon$ and B in the known-location task. However, none of the (correlated) differences between the predicted and estimated values of $d_\varepsilon$ or slope was consistent enough

**Table 2**
Experiment 1 Estimates of $d_\epsilon$ and Linear Slope of Each
Observer's ROC Curve in the Eight-Location Task and
Values Predicted by the Extreme-Detector Model
From Performance in the Known-Location Task

| | Value of $d_\epsilon$ | | | Linear Slope | | |
|---|---|---|---|---|---|---|
| | Esti- | Predicted From | | Esti- | Predicted From | |
| Observer | mate | d' | $d_\epsilon$,B | mate | d' | $d_\epsilon$,B |
| P.F.J. | .836 | .640 | .750 | 1.255 | .950 | .860 |
| R.G.S. | .934 | 1.090 | 1.320 | .990 | .890 | .690 |
| R.G. | 1.084 | 1.050 | 1.070 | .996 | .900 | .880 |
| K.P.D. | .544 | .590 | .640 | .977 | .960 | .920 |
| E.K. | .471 | .640 | .700 | .947 | .950 | .890 |
| F.B. | .712 | .450 | .510 | .848 | .980 | .920 |
| Mean | .764 | .743 | .832 | .990 | .938 | .857 |

across observers to be statistically significant at the .10 level (two-tailed t tests, df = 5).

**Detection and localization.** The identification operating characteristic (IOC) curve was defined as the functional relation between $\gamma_m$ and $\alpha_m$ over variations in the observer's detection criterion, where $\gamma_m$ is the joint probability of a positive detection response and a correct identification of the target's location when it is present. Each observer's detection and location judgments in the eight-location task were used to estimate six points on his IOC curve, one value of $\gamma_8$ for each of the six detection rating categories. The endpoint of the IOC curve is defined as P(CL), the marginal probability of a correct localization decision, estimated by the proportion of correct location-judgments across all the target images. Estimates of P(CL) were obtained from both the observer's eight-location detection task and his separate forced-choice task that presented the 40 target images and only required him to indicate the target's location on each image.

In Figure 3, the filled triangular points within each panel show the six estimated points on that observer's IOC curve in the eight-location detection task, with the estimate of P(CL) on the right ordinate (where $\alpha_8 = 1 - \beta_8 = 1.0$). The circled asterisk plotted on the right ordinate of each panel in Figure 3 is the estimate of P(CL) from the observer's forced-choice task. The two separate estimates of P(CL) were nearly identical for all six observers.

Equation 4 for the extreme-detector model expresses the complete IOC curve in terms of the ROC curve within the eight-location task, which, in turn, is predicted by the known-location ROC curve. Three predicted IOC curves were generated for each observer, using Equation 4 with three separate continuous ROC curves for the eight-location task: the ROC curve fit directly in his performance in the eight-location task and the two ROC curves predicted from the single parameter (d') and two-parameter $(d_\epsilon,B)$

summaries of his performance in the known-location tasks. The broken curve through the filled triangular points in each panel of Figure 3 is the IOC curve predicted from that observer's fitted ROC in the same eight-location task. The solid curve through these data points is the predicted IOC curve based upon the observer's mean estimate of d' in the known-location task.

Table 3 presents the estimates of P(CL) for each observer in the eight-location detection and forced-choice tasks, together with the three predicted values of P(CL) (specified by the endpoints of each separate predicted IOC curve). The extreme-detector model tended to overestimate the observers' ability to locate the target in this sample of images. The mean predicted value of P(CL) in Table 3 exceeded the mean P(CL) estimate, whether the prediction was based upon detection performance in the known-location or the eight-location task. The mean predictions based upon the observers' d' estimates in the known-location task came closest to the mean estimate of P(CL).

## Experiment 2

The analysis of the observers' detection performance followed the procedures used in Experiment 1. Known-location (m = 1) ROC curves characterized by separate maximum-likelihood estimates of d' and of $d_\epsilon$ and B were used to generate the extreme-detector model's predictions for observers' $d_\epsilon$ measures of detectability in the tasks with m = 2, 4, and 8. The direct estimates of $d_\epsilon$ in these tasks came from maximum-likelihood fits of each m-location ROC curve, under the assumption that the true ROC curve was linear in normal-deviate coordinates. The analysis of observers' location judgments focused upon the measure P(CL), the endpoint of the m-location IOC curve. The direct estimate of P(CL) in each m-location task was compared with three predicted values for the extreme-detector model, based upon: (1) the fitted ROC curve in that m-location task, (2) the known-location ROC curve defined by the estimate of d',

**Table 3**
Experiment 1 Estimates of P(CL) for Each Observer's Detection
and Forced-Choice Tasks and Values Predicted by the
Extreme-Detector Model From the Eight-Location
(8L) and Known-Location (KL) ROC Curves

| | Estimated From | | | Predicted From | |
| | | | | | KL | |
| Observer | Detection | Choice | 8L | d' | $d_\epsilon$,B |
|---|---|---|---|---|---|
| P.F.J. | .525 | .538 | .654 | .530 | .565 |
| R.G.S. | .562 | .575 | .660 | .710 | .759 |
| R.G. | .575 | .588 | .711 | .697 | .700 |
| E.K. | .475 | .488 | .425 | .527 | .548 |
| K.P.D. | .588 | .562 | .476 | .503 | .522 |
| Mean | .515 | .515 | .580 | .571 | .597 |

and (3) the known-location ROC curve defined by estimates of $d_\epsilon$ and B.

Table 4 presents the individual estimates of d' and $d_\epsilon$ in the known-location task, the individual fitted and predicted values of $d_\epsilon$ for m = 2, 4, and 8 and the means of these values across the five observers. The extreme-detector model tended to overestimate $d_\epsilon$, particularly for the predictions based upon $d_\epsilon$ and B in the known-location task. Two-tailed t tests of the differences between the 15 individual estimated and predicted values of $d_\epsilon$ (df = 14) were statistically significant for predictions based upon $d_\epsilon$ and B in the known-location task (t = 2.47, p < .05), but not for predictions based upon d' (t = 1.85, p > .10).

Table 5 presents the estimates and three predicted values of P(CL) for each observer in the two, four, and eight-location tasks together with their mean values across observers. The extreme-detector model tended to underestimate P(CL) in the two-location and four-location tasks and to overestimate these values in the eight-location task, whether the predictions were based upon the observers' detection performance in the same m-location task or on their detection performance in the known-location task.

Values of the chi-square statistic, calculated from the actual and predicted frequencies of each observer's correct location judgments for the 40 target images,

were summed over the five observers to obtain a "total chi-square" value for each type of prediction (df = 5). These total chi-square values are presented in the final row of Table 5. The predictions based upon d' in the known-location task were about as good as the predictions from detection performance in the same m-location task. For all three types of predictions, the null hypothesis of an exact fit to the true correct-location frequencies could be rejected in the eight-location task, but not in the two- or four-location task.

## Discussion

The extreme-detector model provides a theoretical account of the effects of signal uncertainty on detection and identification performance. The model links the detectability of a signal to the probability that it will be correctly identified and specifies how these measures change as a function of the observer's degree of uncertainty about the signal. As discussed earlier, in certain important cases this model closely approximates the performance of an ideal detection procedure (Nolte & Jaarmsa, 1967). Experiments 1 and 2 applied the extreme-detector model to observers' detection and localization of visual targets presented in one of m distinct locations on a noisy stimulus field.

Table 4
Estimates of $d_\epsilon$ for Each Observer in the Four Detection Tasks of Experiments 1 and 2
and Values Predicted by the Extreme-Detector Model for m = 2, 4, and 8

| Observer | Estimate for m = 1 | | $d_\epsilon$ for m = 2 | | | $d_\epsilon$ for m = 4 | | | $d_\epsilon$ for m = 8 | | |
| | | | | Predicted From | | | Predicted From | | | Predicted From | |
| | d' | $d_\epsilon$ | ROC Estimate | d' | $d_\epsilon$,B | ROC Estimate | d' | $d_\epsilon$,B | ROC Estimate | d' | $d_\epsilon$,B |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P.F.J. | 1.42 | 1.42 | .99 | 1.12 | 1.15 | .74 | .86 | .93 | .84 | .64 | .75 |
| R.G.S. | 1.98 | 2.01 | 1.32 | 1.66 | 1.76 | 1.08 | 1.36 | 1.53 | .93 | 1.09 | 1.32 |
| R.G. | 1.94 | 1.94 | 1.23 | 1.61 | 1.62 | 1.24 | 1.31 | 1.32 | 1.08 | 1.05 | 1.07 |
| M.S. | 1.34 | 1.34 | 1.12 | 1.05 | 1.02 | .90 | .80 | .76 | .51 | .59 | .55 |
| R.Z. | 1.53 | 1.53 | 1.24 | 1.22 | 1.26 | 1.03 | .95 | .99 | .59 | .72 | .76 |
| Mean | 1.64 | 1.65 | 1.18 | 1.33 | 1.36 | 1.00 | 1.06 | 1.11 | .79 | .82 | .89 |

Table 5
Estimates of P(CL) in Experiments 1 and 2 and Values Predicted by the Extreme-Detector Model for Each Observer in
Tasks With m = 2, 4, and 8, Together With the Total Chi-Square for Each Type of Prediction

| Observer | m = 2 Task | | | | m = 4 Task | | | | m = 8 Task | | | |
| | | | Predicted From | | | | Predicted From | | | | Predicted From | |
| | | | m = 1 Task | | | | m = 1 Task | | | | m = 1 Task | |
| | Estimate | m = 2 ROC | d' | $d_\epsilon$,B | Estimate | m = 4 ROC | d' | $d_\epsilon$,B | Estimate | m = 8 ROC | d' | $d_\epsilon$,B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P.F.J. | .900 | .008 | .844 | .840 | .712 | .613 | .682 | .692 | .515 | .654 | .530 | .565 |
| R.G.S. | .888 | .882 | .922 | .921 | .738 | .752 | .823 | .841 | .562 | .660 | .710 | .759 |
| R.G. | .875 | .832 | .917 | .917 | .850 | .802 | .813 | .814 | .575 | .711 | .697 | .700 |
| M.S. | .962 | .840 | .831 | .826 | .750 | .695 | .660 | .650 | .625 | .471 | .503 | .489 |
| R.Z. | .825 | .864 | .863 | .867 | .700 | .771 | .713 | .723 | .325 | .445 | .567 | .582 |
| Mean | .895 | .846 | .878 | .877 | .753 | .727 | .742 | .749 | .526 | .595 | .605 | .624 |
| Total $\chi^2$ | | .767 | 7.820 | 8.360 | | 3.990 | 3.990 | 5.460 | | 14.800* | 19.000* | 25.500* |

*The null hypothesis of an exact fit can be rejected with p < .05.

The results of Experiments 1 and 2 demonstrate that the extreme-detector model can describe an observer's ability to both detect and locate a spatially uncertain target, predicting how these capabilities change with the number of possible target positions. The probability that an observer would correctly locate the target could be predicted from the ROC curve generated by his ability to distinguish between stimuli that contained a target and those that did not. In addition, both detection and localization performances in tasks with two, four, and eight separate target locations could be predicted from the observer's ability to detect the target when its spatial location was specified on the stimulus image. For the visual stimuli used in these experiments, the extreme-detector model generated accurate predictions of the observer's detection and localization performance in the m-location tasks from a single-parameter (d') specification of his known-location (m = 1) ROC curve.

Various recent studies have compared observers' performance with predictions generated by theoretical assumptions that are equivalent to the present extreme-detector model. Most of these studies concentrated on the relation between the ROC curve in an m-alternative task and what we have termed the identification operating characteristic (IOC) curve. The IOC curve is defined as the joint probability of a positive detection response and a correct identification of the signal, plotted as a function of the false-positive rate. In agreement with the present results, these experiments demonstrate that Equation 4 accurately predicts the relation between observers' ROC and IOC curves, provided that the m stimuli are designed to satisfy the extreme-detector model's assumption that they provide a set of mutually orthogonal perceptual variables.

Starr, Metz, Lusted, and Goodenough (1975) first showed that Equation 4 could predict observers' IOC curves from their ROC curves when the task required observers to identify which quadrant of a noisy visual stimulus contained the target (m = 4). Swets, Green, Getty, and Swets (1978) obtained similarly good predictions from Equation 4 for decisions that observers made at five successive stages in an eight-location task, which progressively increased the amount of visual information available at each stage. Sandor and Swensson (1978) showed that Equation 4 could be used with observers' fitted ROC curves to predict P(CL), the overall probability of correctly locating the target, for visual targets presented at various levels of physical detectability in a five-location task. In a combined auditory-detection and frequency-identification task, Green, Weber, and Duncan (1977) found that Equation 4 successfully predicted the joint probability of detecting and identifying the stimulus tone in tasks with both two and four different frequencies, provided that the differences in stimulus-frequency were sufficiently large.

Aside from the present results, however, there is less empirical evidence that the extreme-detector model can predict the changes in detectability produced by variations in m, the number of alternative signals. Swets and Birdsall (1978) presented observers with repeated independent observations of an auditory stimulus in detection tasks that either kept the frequency of the tone constant (m = 1) or selected the stimulus randomly from a set of eight well-spaced frequencies. For two of their three observers, the values of d' obtained in the eight-frequency task agreed with values predicted by the extreme-detector model from estimates of d' based upon the same number of observations in the single-frequency task. The third observer's estimates of d' in the eight-frequency task increased with d' in the single-frequency task, but were below the values predicted by the extreme-detector model for m = 8. Figure 1b of the present paper suggests that such a pattern of performance could be explained by the extreme-detector model if m were assumed to be greater than eight—that is, if this observer considered more possible stimulus frequencies than the eight that were actually presented.

The results of Experiments 1 and 2 indicate that the extreme-detector model, which should approximate the optimal detection procedure for m locations, adequately describes the changes in detectability of these visual targets as the number of possible target locations increases from one to eight. The absolute levels of detection performance, however, were considerably less than could be achieved by a physical detector, as indicated by the target's signal-to-noise ratio (S/N) calculated for a matched filter at a known location on the image. The matched-filter detector would have obtained a detectability index ($d_\varepsilon = S/N$) of 2.64, while the mean estimate of $d_\varepsilon$ (and d') in the known-location task was only 1.54. This deviation from the theoretically achievable level of performance must reflect observers' inability to extract all the relevant physical information about the presence of a target from a given location on these images.

## EXPERIMENT 3

This experiment used the physical description of the visual stimulus to predict the observers' performance in an eight-location task. For the task of detecting these visual targets, the physical stimulus can be characterized by the target's signal-to-noise ratio (S/N), calculated from the output of a matched filter that is tuned to detect the target in a specified location on the image. As discussed earlier, this value of S/N is equal to the detectability index $d_\varepsilon$ for a matched filter when m = 1. Experiments 1 and 2 found a considerable discrepancy between observers' detection performance in the known-location task and the predicted performance of a matched filter. However, given an estimate of the target's detectability

in a known location, the detection and localization performance in m-location tasks could be predicted by modeling the observer as an extreme detector. This suggests that the extreme-detector model could predict an observer's absolute level of performance in an m-location task from the target's S/N ratio if it included an adequate model of how this S/N ratio determines the target's "effective detectability" in a specified location on the image.

The experiment obtained measures of observers' eight-location detection and localization performance in 10 different conditions, which used various stimulus manipulations to vary the target's S/N ratio over a large range. Since the details of this experiment have been reported elsewhere (Judy, Swensson, & Szulc, 1981), we present only selected results here. These data are used to compare four different "psychophysical" models that represent alternative formulations of the relation between the physical characteristics of these visual stimuli (e.g., the S/N ratio) and the target's detectability in a specified location.

## The Psychophysical Models

The four models described below relate the detectability of a target that appears in a known location on the image to the "signal" (S) and "noise" (N) parameters of the physical stimuli: $S \equiv \mu_n - \mu_t$ and $N \equiv (\sigma_n + \sigma_t)/2$, where $\mu_t, \mu_n$ and $\sigma_t, \sigma_n$ are the respective means and standard deviations of the output of a matched filter that considers the brightness levels of the relevant set of image pixels, given the presence and absence of a target. These models assume that an observer's detection ROC curve in a known-location task can be characterized by the single parameter $d'_{ef}$, his "effective value of $d'$" for that stimulus condition. Each model uses a single fitted parameter to specify a unique value of $d'_{ef}$ for a given combination of the physical variables S and N. Models 1 and 2 are simply two different formal mappings between the target's S/N ratio and $d'_{ef}$. Models 3 and 4 represent different possible interpretations of the reduction in $d'_{ef}$ from the physical upper bound set by the S/N ratio.

### Model 1

This model assumes that the target's detectability is reduced by a constant amount, independent of the value of S/N for the matched filter: $d'_{ef} = (S/N) - k_1$, where $k_1 > 0$ and $d'_{ef} = 0$ whenever $S/N \leqslant k_1$.

### Model 2

This model assumes that the reduction in target detectability is proportional to the value of S/N for the matched filter: $d'_{ef} = k_2(S/N)$, where $0 < k_2 < 1$.

### Model 3

This model attributes the reduction in target detectability to some source of "noise" ($k_3 > 0$), contributed by the observer, which is assumed to be independent of N, the physical noise in the output of the matched filter. Examples of such noise sources would include inappropriate tuning of the observer's "perceptual filter" and any variability in his detection procedures (e.g., the specific decision criteria) between experimental trials. These independent sources of noise would combine additively with N to yield $N_{ef}$, the "effective level of noise" for the human observer: $N_{ef}^2 = (N^2 + k_3^2)$. This model assumes that $d'_{ef} = S/N_{ef}$, which depends upon the particular values of S and N for each stimulus condition (rather than the ratio S/N). A change in S/N produced by varying N has less effect upon $d'_{ef}$ than the same change when produced by varying S; this asymmetry increases with the size of the change in S/N.

### Model 4

This model assumes that the observer has some residual spatial uncertainty about each discrete location for the target, because he is unable to fix its precise position on the image. This region of uncertainty about each discrete location degrades the target's detectability, since it renders the observer susceptible to noise variations that would be ignored by the matched filter. The degree of this spatial uncertainty can be represented by a parameter $k_4 > 1$, which has an interpretation as the number of discrete orthogonal spatial locations that would provide the equivalent reduction in target detectability for an extreme detector. Under this interpretation of the model, the effect of any assumed degree of spatial uncertainty ($k_4$) can be calculated by substituting $m = k_4$ into Equations 1a and 1b for the extreme-detector model, where values of $\alpha_1$ and $1 - \beta_1$ represent points on the matched filter's ROC curve (specified by Equation 3 with $d' = S/N$). The value of $d'_{ef}$ for this model would then be obtained by iterating Equations 1a and 1b until $\alpha_k = \beta_k = \epsilon_k$ and calculating $d'_{ef} = 2Z^{-1}(1 - \epsilon_k)$.

Figure 1b shows that, except for small values of $d'$, the effect of increasing the number of target locations for an extreme detector is essentially to shift the function that relates the m-location $d_\epsilon$ to $d'$ in the $m = 1$ situation. Thus, the effect of $k_4 > 1$ would be to reduce $d'_{ef}$ by an approximately constant amount for larger values of S/N. For this reason, the predictions of Model 4 should be similar to those of Model 1, given appropriate choices of $k_1$ and $k_4$.

## Method

The visual stimuli were similar to those used in Experiments 1 and 2. Ten different stimulus conditions varied the target's S/N ratio for separate sets of 80 images, 40 images with a single target in one of the eight locations and 40 images with no target. Two sets of 80 images were generated for a baseline condition, in which the stimulus parameters were identical to those for the images used in Experiments 1 and 2. The remaining nine conditions changed one physical parameter of the stimuli, while the

others remained fixed at their baseline levels. These conditions included three levels of increased S/N ratio for each of three different manipulations of the stimulus parameters: (1) increasing the contrast between the target and the surrounding ring (S), (2) decreasing the level of the image noise (N), and (3) increasing the target's size (from 4.95 sq pixels in the baseline condition to 5.94, 6.93, and 8.92 sq pixels). All other characteristics of the stimuli were the same as in Experiments 1 and 2.

The observers considered one set of images per session. On each image, the observer indicated which of the eight possible locations he considered most likely to contain the target and rated the likelihood that the target was present on a 6-category scale. There were 31 sessions, including four replications of the baseline condition (Sessions 1, 11, 21, and 31) and three replications of the other nine conditions. In order to minimize the apparent changes in the image sets between sessions, the three conditions produced by a given type of stimulus manipulation always appeared together, in order of increasing S/N ratio. The order of these three types of conditions was counterbalanced across the three replications for each observer. All other procedural details were identical to those in Experiments 1 and 2.

## Results

Detailed results for all three observers in this experiment are available elsewhere (Judy, Swensson, & Szulc, 1981); we present only mean data for two observers (P.F.J. and R.G.S.), who also performed in Experiments 1 and 2. There were two measures of performance for each of the 31 individual sessions. One measure was the $d_\varepsilon$ index of detectability, estimated from the ROC curve generated by the ob-
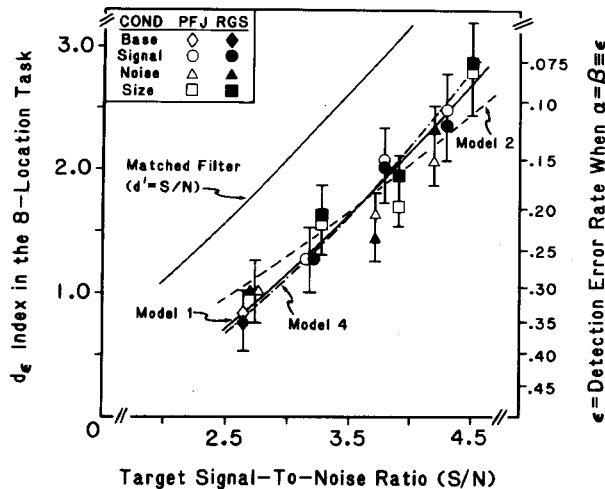


**Figure 4.** Mean estimates of the $d_\varepsilon$ index of detectability for each observer in the 10 stimulus conditions of Experiment 3, plotted as a function of the target's S/N ratio. The right ordinate gives values of the detection error rates that correspond to each value of $d_\varepsilon$ at the point on the ROC curve where $\alpha = \beta \equiv \varepsilon$. The shape of each data point indicates the type of stimulus manipulation used to increase the S/N ratio in that condition. Vertical bars around the mean $d_\varepsilon$ for both observers indicate plus and minus one standard error of the $d_\varepsilon$ estimate in each of the 10 stimulus conditions, obtained from the maximum likelihood fitting procedure. The continuous curves show the value of $d_\varepsilon$ predicted by the extreme-detector model for the matched filter (with $d' = S/N$) and for three of the fitted psychophysical models.
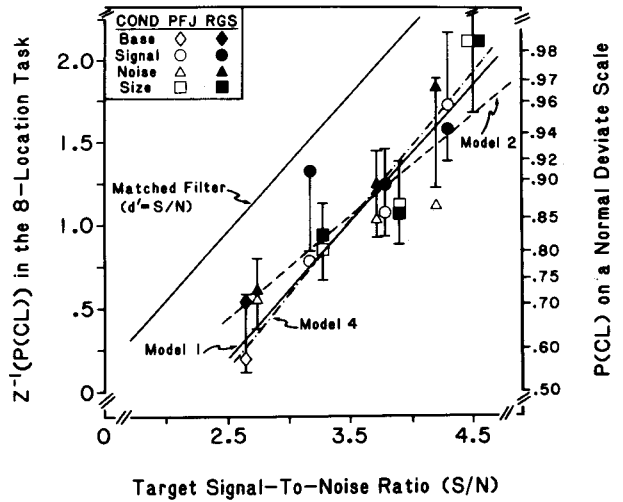


**Figure 5.** Mean estimates of P(CL), the proportion of correctly located targets, for each observer in the 10 stimulus conditions of Experiment 3, plotted (on a normal-deviate scale) as a function of the target's S/N ratio. Vertical bars around the mean value of P(CL) for the two observers indicate plus and minus the binomial standard error of that estimate for a sample of 40 target images. The continuous curves show the values of P(CL) predicted by the extreme-detector model for the matched-filter (with $d' = S/N$) and for three of the fitted psychophysical models.

server's confidence ratings, using the same maximum-likelihood procedure as in Experiments 1 and 2. The second measure was P(CL), the proportion of the 40 target images on which the observer correctly identified the target's location. The three individual estimates were averaged over replications to obtain the observer's mean estimate of $d_\varepsilon$ and P(CL) in each of the 10 stimulus conditions.

The observed changes in both $d_\varepsilon$ and P(CL) were closely related to the changes in the target's S/N ratio. Figures 4 and 5 present values of $d_\varepsilon$ (Figure 4) and P(CL) on a normal-deviate scale (Figure 5) for each observer's 10 stimulus conditions, plotted as a function of the target's S/N ratio. The shape of each data point in Figures 4 and 5 indicates the type of stimulus manipulation used to obtain that level of S/N ratio. Vertical bars indicate plus or minus one standard error of the estimates of $d_\varepsilon$ and P(CL) around the mean value for the two observers in each stimulus condition. The standard errors of the $d_\varepsilon$ estimates were obtained from the maximum-likelihood fitting procedure; they reflect the expected variability in estimates of $d_\varepsilon$ among independent samples of 40 target and 40 nontarget stimuli generated from the same physical parameters. The standard errors of P(CL) represent the binomial variability in probability estimates based upon a sample of 40 cases (target images).

The smooth curves in Figures 4 and 5 show the relations between S/N and measures of $d_\varepsilon$ and P(CL) in the eight-location task, predicted by the matched

filter and by single-parameter fits of three of the previously defined psychophysical models (Models 1, 2, and 4). These curves are the predicted values given by the extreme-detector model for the eight-location task, generated from values of d′ for the known-location situation under the assumption that d′ = S/N (for the matched filter) or d′ = $d'_{ef}$ (for the fitted models).

The four psychophysical models were fitted to the combined data from both observers by finding the values of k that would equate the means of the observed and predicted values of $d_\varepsilon$ for the 10 stimulus conditions. The fitted k for each model reduced the mean S/N ratio of 3.618 to $d'_{ef} = 2.68$, the value of d′ for which the extreme-detector model predicted the observed mean $d_\varepsilon$ of 1.748 in the eight-location task. Each fitted psychophysical model specified a unique value of $d'_{ef}$ in each of the 10 stimulus conditions from the physical variables of S and N in that condition. The extreme-detector model then used these values of $d'_{ef}$ to predict the 10 values of $d_\varepsilon$ and P(CL) for each model in the eight-location task. The ability of each fitted model (in combination with the extreme-detector model) to account for the observed variations in $d_\varepsilon$ of $Z^{-1}[P(CL)]$ can be measured by the proportion of the total variance in these estimates that the model explains. This proportion of explained variance was calculated by subtracting the mean squared error in the 10 predicted values of $d_\varepsilon$ or $Z^{-1}[P(CL)]$—the unexplained variance for each model—from the total variance in the 10 mean estimates (averaged over both observers) and dividing by this total variance.

Table 6 presents the fitted parameters for each of the four psychophysical models and the proportion of variance in estimates of $d_\varepsilon$ and $Z^{-1}[P(CL)]$ explained by that model. As a rough estimate of the amount of systematic variation in these data, Table 6 also gives the proportion of variance ($r^2$) in each set of estimates explained by its two-parameter linear fit to the values of S/N. The four models are ordered in nearly the same way by their ability to account for variations in both $d_\varepsilon$ and P(CL). Model 1 provided the best fits to both performance measures, ac-

counting for almost all of the systematic (linear) variation in the estimates of $d_\varepsilon$ and $Z^{-1}[P(CL)]$, while Model 3 yielded the worst fits. As seen in Figures 4 and 5, the predicted functions for Models 1 and 4 were virtually indistinguishable. Model 2 tended to underestimate the rate of improvement in observers' performance as the S/N ratio increased.

The predictions of Model 3 were not functions of the ratio S/N and could not be plotted in Figures 4 and 5, but they were clearly inferior to those of the other models in terms of the proportions of variance they explained in the estimates of performance. In order to match the predicted and observed mean levels of $d_\varepsilon$, Model 3 had to assume that the amount of additive noise contributed by the observer (1.004 units) was roughly the same in magnitude as N, the noise in the physical stimuli, which ranged between .69 and 1.77 units over the 10 stimulus conditions with a mean of 1.20. This level of assumed additive noise made $d'_{ef}$ relatively insensitive to decreases in N. As a result, Model 3 considerably underestimated the improvement in performance produced by large decreases in the physical image noise and overestimated the improvement for increases in target size and contrast (S). For the stimulus condition with the largest decrease in N, the predicted and observed values of $d_\varepsilon$ were 1.45 and 2.19, respectively; the corresponding values of P(CL) were .81 and .93. The shortcomings of Model 3 reflect the fact that observers' performance appeared to depend only upon the ratio between S and N for these visual stimuli, rather than on their separate individual values.

## Discussion

The combined models tested in Experiment 3 all assume that the observer functions as an extreme detector in an m-location task, but is less able to detect the target in any individual location than a physical detector that uses the output of a matched filter, whose value of d′ is given by the target's physical signal-to-noise (S/N) ratio. The four combined models differ in the particular psychophysical model that links measures of the physical stimuli to the target's detectability, its effective value of d′ in a known spatial location. Once the single parameter of each psychophysical model had been fit to the observer's detection data, the model specified the effective d′ in each stimulus condition from calculations of S and N for these physical stimuli. The extreme detector model was then used to predict indices of the observer's detection and localization performance, $d_\varepsilon$ and P(CL), in the eight-location task. Thus, the predictions of each combined model in Experiment 3 tested the extreme-detector model, as well as the particular psychophysical model to which it was appended.

Model 1, the simplest psychophysical model, assumed that the physical S/N ratio was reduced by a

### Table 6
Fitted Parameters for the Four Psychological Models in Experiment 3 and the Proportion of Variance in Observers' Mean Estimates of $d_\varepsilon$ and $Z^{-1}[P(CL)]$ Explained by Each Model and by the Linear Fit to S/N

| Model | Fitted Value of K | Proportion of Variance Explained in Estimates of | |
|---|---|---|---|
| | | $d_\varepsilon$ | $Z^{-1}[P(CL)]$ |
| Linear Fit | | .930 | .886 |
| Model 1 | $k_1$ = .938 | .933 | .864 |
| Model 2 | $k_2$ = .7409 | .885 | .853 |
| Model 3 | $k_3$ = 1.0042 | .694 | .566 |
| Model 4 | $k_4$ = 9.4 | .920 | .844 |

constant to yield the observer's d' for a known-location target in each condition. This produced the most successful combined model; it accounted for 93% of the variation in mean estimates of $d_\xi$ and 86% of the variation in mean estimates of $Z^{-1}[P(CL)]$ over the 10 independent stimulus conditions. Model 2, which assumed that the observer's d' was proportional to the physical S/N ratio in each condition, generated predictions that tended to underestimate how rapidly detection and localization performance improved with increases in the target's S/N ratio.

Model 3 attributed the degradation from ideal detection performance to additive noise, contributed by the observer, which was independent of the noise in the physical stimuli (N). This was the least successful of the four psychophysical models for predicting detection and localization performance in the eight-location task, primarily because the model greatly underestimated the degree of improvement in performance produced by decreases in N. Since this model had to assume that the additive noise from the observer was about the same magnitude as the stimulus noise, it predicted that performance would be insensitive to variations in N. In fact, however, the observers' performance was closely related to the physical S/N ratio, regardless of the type of stimulus manipulation used to achieve that level of S/N ratio.

Model 4 attributed observers' suboptimal detection performance to the presence of some residual uncertainty about the precise position of each possible location on the visual images. The model approximated this region of spatial uncertainty about each possible location by an "equivalent" number of discrete orthogonal observations, assuming that each additional noise sample was distributed the same as a sample of noise from the target's actual position. The observer was assumed to fix the spatial position of his "target candidate" (in each possible location) by the area corresponding to the most extreme sample within that region of uncertainty. When Model 4 was combined with the extreme-detector model for the eight-location task and fitted to the mean performance in Experiment 3, its predictions were essentially indistinguishable from those generated by Model 1. This is because the effect of considering additional noise samples is to reduce an extreme-detector's ability to detect the target—that is, d' for a given target-location in this case—by an approximately constant amount, over the range in S/N values used in Experiment 3.

Because of the similar predictions made by Models 1 and 4, both of which closely fit the observed data, it is tempting to regard Model 4 as a theoretical interpretation of the process that underlies the formal approximation specified by Model 1. However, there are difficulties with such an interpretation. Model 4 had to assume that the residual spatial uncertainty was equivalent to a mean of 9.4 orthogonal areas

around each possible target location on the image. This estimate seems excessive for these highly experienced observers, even though it need not correspond to a physical region as large as 9.4 times the actual target area. A more serious problem for Model 4 is the fact that manipulations of the target's size were equivalent to the other stimulus manipulations, in that comparable increases in the physical S/N ratio all produced the same degree of improvement in observers' performance. An increase in the target's size would have reduced the detrimental effects of any fixed amount of residual spatial uncertainty, which should have caused Model 4 to underestimate the observed rate of improvement in performance as the target became larger. To interpret the effects of manipulating the target's size, Model 4 seems forced to the unlikely assumption that the observer's region of spatial uncertainty about each target location expanded in proportion to the target's area.

## CONCLUSIONS

The results of these three experiments demonstrate that an extreme-detector model can describe the relation between target detection and target localization, as well as the changes in these capabilities with variations in: (1) the number of possible target locations and (2) the signal-to-noise (S/N) ratio of the physical target stimuli. The extreme-detector model assumes that the observer selects the location that gives rise to the most extreme value of some underlying decision variable, basing his detection decision only upon the value of this extreme sample. For relatively simple types of visual targets and noise backgrounds, like those used in these experiments, estimates of a single parameter may be sufficient to predict the changes in detection and localization performance produced by manipulations of both the number of possible target locations and the target's physical S/N ratio. From estimates of d', the observer's ability to detect the target in a known location on the image, the extreme-detector model closely predicted detection and localization in tasks with two, four, and eight possible locations, which varied the target's detectability by a factor of two. From calculations of the target's physical S/N ratio (degraded by a fitted constant) the extreme-detector model predicted a wide range of detection and localization performances in an eight-location task that manipulated the target's S/N ratio in various ways.

### REFERENCES

DORFMAN, D. D., & ALF, E., JR. Maximum-likelihood estimation of parameters of signal-detection theory and determination of confidence intervals—Rating-method data. *Journal of Mathematical Psychology*, 1969, 6, 487-496.

ELLIOT, P. B. Tables of d'. In J. A. Swets (Ed.), *Signal detection and recognition by human observers*. New York: Wiley, 1964.

GREEN, D. M., & BIRDSALL, T. G. Detection and recognition. *Psychological Review*, 1978, **85**, 192-206.

GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966. (Reprinted: Huntington, N.Y: Krieger, 1974.)

GREEN, D. M., WEBER, D. L., & DUNCAN, J. E. Detection and recognition of pure tones in noise. *Journal of the Acoustical Society of America*, 1977, **62**, 948-954.

GREY, D. R., & MORGAN, B. J. T. Some aspects of ROC curve-fitting: Normal and logistic models. *Journal of Mathematical Psychology*, 1972, **9**, 128-139.

HACKER, M. J., & RATCLIFF, R. A revised table of d' for M-alternative forced choice. *Perception & Psychophysics*, 1979, **26**, 168-170.

JUDY, P. F., SWENSSON, R. G., & SZULC, M. Lesion detection and signal-to-noise ratio in CT images. *Medical Physics*, 1981, **8**, 13-23.

NOLTE, L. W., & JAARSMA, D. More on the detection of one of M orthogonal signals. *Journal of the Acoustical Society of America*, 1967, **41**, 497-505.

SANDOR, T., & SWENSSON, R. G. Evaluation of observer performance in detecting blood vessels on simulated angiographic images. *Medical Physics*, 1978, **5**, 380-386.

STARR, S. J., METZ, C. E., LUSTED, L. B., & GOODENOUGH, D. J. Visual detection and localization of radiographic images. *Radiology*, 1975, **116**, 533-538.

SWETS, J. A., & BIRDSALL, T. G. Repeated observation of an uncertain signal. *Perception & Psychophysics*, 1978, **23**, 269-274.

SWETS, J. A., GREEN, D. M., GETTY, D. J., & SWETS, J. B. Signal detection and identification at successive stages of observation. *Perception & Psychophysics*, 1978, **23**, 275-289.