Distraction of attention and the slope of the psychometric function

Leonid L. Kontsevich and Christopher W. Tyler

Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, California 94115

Received March 23, 1998; revised manuscript received August 17, 1998; accepted September 28, 1998

The influential uncertainty model [J. Opt. Soc. Am. A **2**, 1508 (1985)] attributes nonlinear contrast sensitivity near threshold to the inability of the observer to discriminate between the signal from stimulated locations and the noise from nonstimulated locations. We introduce an alternative interpretation, the distraction model, to describe the behavior of an observer who knows exactly which location was stimulated but may miss the test stimulus because attention has been distracted by irrelevant (noise) signals. For any stimulus sample, the observer is assumed to be certain of whether this sample is relevant or irrelevant to the stimulus. The nonlinear effects predicted by the distraction model without uncertainty are similar to those predicted by the uncertainty model. [S0740-3232(99)00402-0]

OCIS codes: 330.4060. 330.5510, 000.5490.

1. INTRODUCTION

One of the most important recent developments in signal detection theory came with the realization that the observer may be uncertain about the key parameters of the stimulus, such as time, position, spatial frequency, and phase.[1,2] The uncertainty model assumes that the visual system monitors a set of analyzers and is completely uncertain about which of them carries the signal. A fuller development of the uncertainty idea was provided by Pelli,[3] who proposed some approximations to the original uncertainty model to achieve analytic results. (The optimal maximum-likelihood behavior of such an observer requires summation of the exponentiated likelihood ratios across all analyzers monitored by the observer,[4] which is hard to formalize or explain in biological terms.) According to Pelli's model, the visual system makes a judgment based on the strongest signal present across all monitored local analyzers, not just those stimulated by the test stimulus. Reaching the decision stage, therefore, requires that the signal evoked by the stimulus exceed the strongest noise signal from a potentially large number of nonstimulated locations, depending on how many are monitored by the observer. The distribution of the maxima of the noise samples is biased in the positive direction, which acts as a soft threshold. This bias steepens the slope (or exponent) of the psychometric function of d' versus stimulus strength and produces effects similar to those for systems with nonlinear signal transducers.

Uncertainty is a factor that can, indeed, affect the slope of psychometric functions. It is not clear, however, how significant this factor is in most psychophysical experiments. For example, the quadratic transducer acceleration that is typical for contrast-detection tasks[5,6] implies that the observer mistakenly attributes approximately 20 times as many local signals to the test stimulus as the actual number of signals (see Table 1 of Ref. 3). For a nonlinearity exponent of 4, which is also quite common in psychophysical studies,[7,8] this ratio increases to up to 10,000 times as many channels monitored as stimulated. It seems at least questionable that a practiced observer is so inefficient in using the prior information about the stimulus.

The concept of uncertainty is closely related to spatial attention.[9] The detection scheme of the uncertainty model can be presented as if attention monitors a certain pool of local analyzers and then gets attracted by and reads the maximum signal in this pool. Under this concept the visual system does not distinguish among the signals from the pool; any signal coming from anywhere in the pool is equally attributed to the stimulus. As mentioned, to produce nonlinear effects consistent with the experimental data, the pool must be many times larger than the set of local analyzers relevant to the stimulus detection.

We introduce an alternative scheme that assumes that the observer is perfectly certain both where and when the detection event will appear but that the observer's attention may nevertheless be distracted by the signals produced in irrelevant (unstimulated) locations, reducing sensitivity to the test stimulus. The analysis shows that this distraction model predicts nonlinear effects similar to those predicted by the uncertainty model. We argue that for many experimental conditions, distraction rather than uncertainty may be the predominant factor in defining the shape of the transducer nonlinearity.

2. DISTRACTION MODEL

A. Definition

In the distraction model the observer monitors the activity of M independent local analyzers, of which K get incremented by the test pattern (see Appendix A for a complete list of notation). It is assumed that the observer can identify the K analyzers that are relevant to the task. The analyzer with maximum activity in the monitored pool attracts the observer's attention. If that analyzer is one of the relevant ones, then the activity level of that analyzer is remembered. If the analyzer is irrelevant (the distraction), then observer knows to ignore its signal, and zero information is obtained from the presentation.

The noise is assumed to be additive, Gaussian, identically distributed, and stochastically independent in all M analyzers. The test stimulus evokes equal signals in all K relevant analyzers before the noise infusion. These two assumptions are made for simplicity of the model derivation.

A diagram representing the distraction model is shown in Fig. 1(b); for comparison a diagram of the uncertainty model is provided in Fig. 1(a). The front end up to the "choose-largest" stage in the distraction model is identical to the maximum operator of the uncertainty model. In the uncertainty model this stage is followed by a standard decision maker. In the distraction model, operation of the decision maker is controlled by a process that checks for relevance of the selected sample to the stimulus.

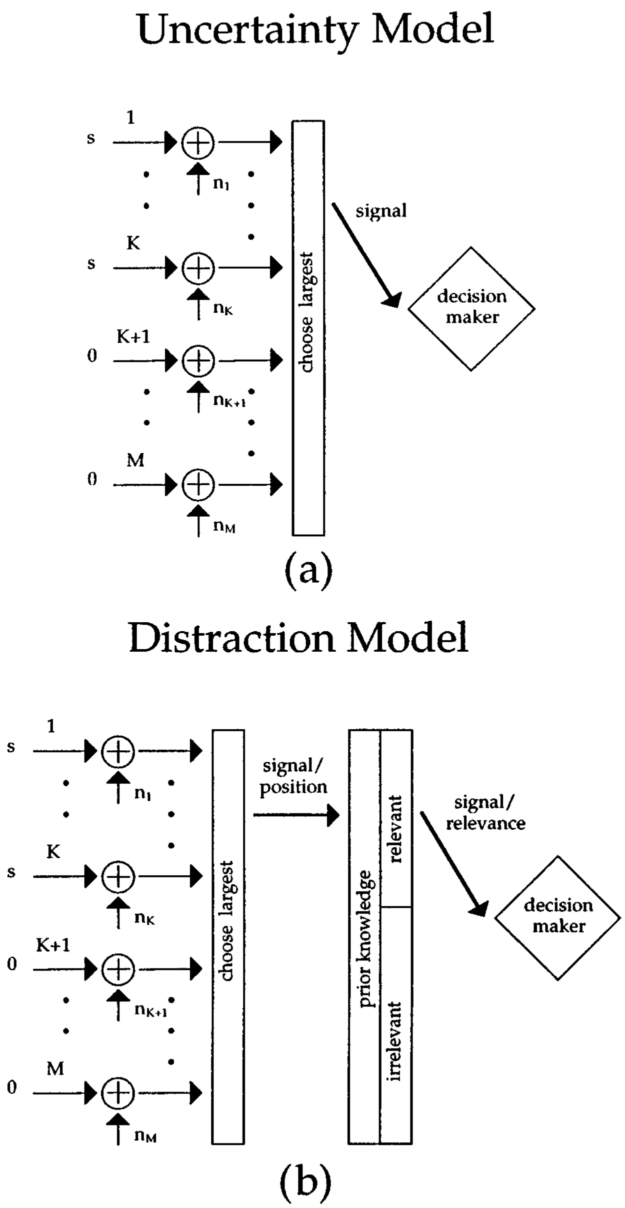

Fig. 1. Uncertainty and distraction models. The two models have similar initial stages: The signals from local analyzers get mixed with additive Gaussian noise, after which the largest signal gets chosen. The uncertainty model makes the decision based directly on this signal; the distraction model first checks the relevance of the maximum signal.

B. Equivalence to an Attention Model

The assumptions that define the distraction model can be restated as the properties of the underlying attentional mechanism.

The maximum rule adopted in the model implies that the system has access to any individual analyzer. The size of the attentional field, consequently, should gather one sample at a time; i.e., the attentional field is assumed to be punctate at the level of the local analyzers. This conclusion does not mean, however, that attention is punctate relative to the input image; at the retinal level the attentional field would correspond to the receptive field of the analyzer that is currently attended. Only attended analyzers can affect the observer's decision.

To pick the maximum among the multiple signals, one should have a mechanism that selects the maximum signal from all available analyzers. We speculate that this task is performed by a low-level preattentive mechanism,[10] which monitors all local analyzers in the visual field. The parameter M in the model, therefore, can be as large as the total number of the analyzers in the visual field.

The properties described situate our attention model at one extreme of the spotlight metaphor[11]: The attentional focus is assumed to move across the visual field and to have the pinpoint focus of a laser pointer at the level of the outputs of the local analyzers. The only way to vary the effective area of such an attention focus is to switch among analyzers with receptive fields of different sizes. The attentional focus is assumed to be sampling only one receptive field at a time.

One might consider such an extreme version of the spotlight metaphor as too restrictive. It could, indeed, be loosened without invalidating the whole approach: Attention itself might be the analog of a receptive field that linearly or nonlinearly integrates weighted outputs of the real receptive fields. We did not elaborate the model along this avenue because that would introduce a number of auxiliary parameters, but the model could be extended in this way.

Moreover, a major simplification of the present model is the assumption that attention reads a single sample during the stimulus presentation, which is certainly not adequate for prolonged presentations. Extension of the distraction model for prolonged presentations is left for the future.

C. The Distraction Model and the Two-Alternative Forced-Choice Task

In the traditional analysis of the two-alternative forced-choice (2AFC) task[5] the observer compares the signals evoked by two stimulus presentations, one of which contains a test stimulus while the other is blank. The larger of the two signals defines which presentation will be named as the test. The same rule is employed in the uncertainty model.[3]

Table 1. Possible Outcomes of 2AFC Experiment Depending on the Relevance of the Samples

Case Number	Blank Interval	Test Interval	% Correct
1	Irrelevant	Irrelevant	50
2	Irrelevant	Relevant	100
3	Relevant	Irrelevant	0
4	Relevant	Relevant	Variable

The distraction model requires a more elaborated decision rule since it deals with two variables: signal and position (see Table 1). In each presentation the observer may attend to the relevant locus or may be distracted. If the observer attends to irrelevant loci in both presentations of the trial (case 1 in the Table 1), it is assumed that sensitivity in the test location drops to zero. Thus the only way to produce an answer would be to guess, resulting in a success rate of 50%. In cases 2 and 3 attention is attracted by a relevant analyzer in only one of the two alternatives. It is assumed that the observer will respond with that alternative. This decision will be correct 100% of the time for case 2 and never for case 3 (0% correct). Finally, relevant loci may be attended in both presentations (case 4). This case should be treated as in the standard 2AFC analysis: The observer chooses the presentation with the larger signal. In a typical case when the number of the relevant channels is a small fraction of the monitored ones ($K \ll M$), cases 1 and 2 are the most common.

The net percent correct for the distraction model is the probabilistic mean of the percent correct values derived for all four cases listed in Table 1, as specified in Appendix C. Appendix D presents a brief description of the uncertainty model to allow comparison between the models. Closed-form solutions for both models were implemented in MATLAB code.

It is possible that in cases 2 and 3 the observer applies a criterion to the relevant location and responds with the other interval if the signal falls below this criterion. Application of such a criterion is not usually considered in analyses of 2AFC tasks, but its effect should be minimal since it merely reassigns the responses between the cases without altering the overall success rate.

3. COMPUTATIONAL EXPERIMENTS WITH THE MODEL

Percent correct in both models is derived as a function of the internal signal-to-noise ratio $c' = s/\sigma$, which is also known as normalized contrast.[3] In computational experiments we set $\sigma = 1$, which made c' identical to the signal variable s. Normalized contrast is the same as Weber contrast up to a multiplier.

The performance of the simulated observer in a 2AFC task is expressed here in terms of detectability $d' = \Phi^{-1}(P_{\mathrm{corr}})$, where Φ^{-1} ( ) is the inverse cumulative normal function. The relationship between d' and c' at small signals can be approximated by a power function $d' \propto (c')^b$. The exponent b is a standard measure for the slope of psychometric function[5]; it is equal to the slope of the detectability curve in double logarithmic coordinates.

A. Simulation 1

To demonstrate that distraction may cause a nonlinear effect, we set the total number of samples at $M = 10^5$ and computed the psychometric functions for a range of relevant samples $K = \{10^5, 10^4, 10^3, 10^2, 10, 1\}$. The results in Fig. 2 show that, when all samples are relevant, i.e., $K = M$, the transducer is almost linear. The nonlinear accelerating effect of distraction increases with reduction of the proportion K/M of relevant samples. (The transducer is not exactly linear when $K = M$, even though the signals in test and blank trials have similar distributions, because the computation of d' assumes Gaussian distributions; the distributions for the maximum signals are not Gaussian.)

B. Simulation 2

The thresholds and the slopes of psychometric functions were evaluated for three cases: (1) for the distraction model with increasing stimulus size K and a large monitored field $M = 10^5$, (2) for the distraction model with a fixed small stimulus $K = 1$ and increasing monitored pool M, and (3) for the uncertainty model with similar assumptions $K = 1$ and a range of M. The first case represents a natural condition for the distraction model: The large number of signals that potentially may attract attention is kept constant; the variable is the number of relevant signals, which may reflect, for example, the size of the test stimulus. The second and third cases represent a typical range under which the uncertainty model was analyzed by Pelli[3]: Only one of a variable number of monitored analyzers is relevant. Setting the same range for two models makes their comparison direct.

The thresholds were defined by $d' = 1$, which corresponds to 76% correct. We estimated the slopes locally

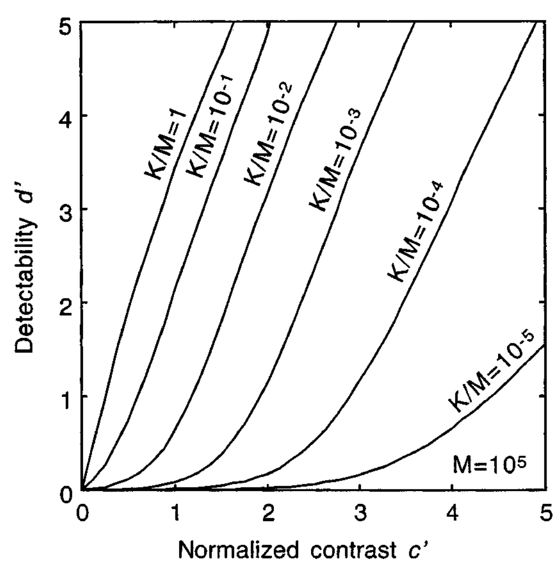

Fig. 2. Effective signal produced by the distraction model in the 2AFC task for a range of ratios K/M. The number of monitored channels M was kept constant at a value of 10^5. The nonlinearity of the transducer increases for small K/M.

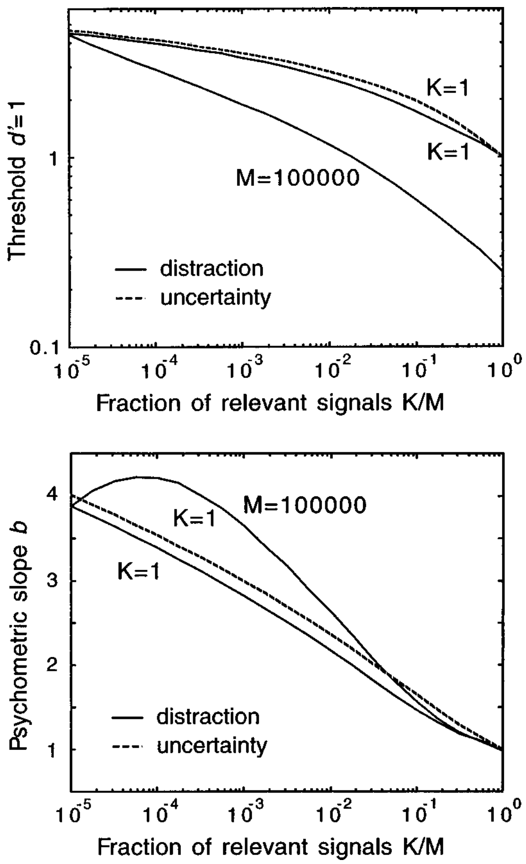

Fig. 3. Top, detection thresholds and bottom, slopes predicted by the distraction and the uncertainty models. The predictions of the uncertainty model are shown for the single relevant channel ($K = 1$) and a variable number of the monitored channels. The predictions of the distraction model are obtained for two conditions: for the single relevant channel ($K = 1$) and for a fixed number of monitored channels ($M = 10^5$). The predictions of the two models for $K = 1$ are similar, indicating that the effects of uncertainty and distraction are comparable.

by taking the derivative at the threshold level. The curves in Fig. 3 represent the thresholds (top) and the slopes (bottom) as functions of the ratio K/M.

The threshold and slope curves computed for the second and third cases of increasing size of the monitored pool ($K = 1$) have similar shapes; the predictions of the uncertainty model are slightly higher than those of the distraction model. Distraction, therefore, has an effect that is similar to, although slightly weaker than, uncertainty.

Our threshold estimates for the uncertainty model are broadly consistent with those provided by Pelli[3] in his Table 1. There is, however, a slight mismatch between the slope estimates: Ours are systematically shallower. A close look at the psychometric functions predicted by either the uncertainty or the distraction model reveals that the slope is not uniform within a single psychometric function and therefore that its estimate depends on where it is measured. In our study we computed a local slope estimate at $d' = 1$, whereas Pelli looked for the maximum-likelihood fit for ten sample contrasts. This methodological difference explains the small mismatch.

The curves produced by the distraction model for the case of a large monitored field with increasing size of the stimulus pool ($M = 10^5$ and a range of K) deviate radically from the curves for the other two cases (Fig. 3, upper panel). The threshold increases dramatically toward the condition of no distraction (no uncertainty in case 3), i.e., for $K/M = 1$, where the lowest value is four times lower than a single-channel threshold of unity value. This improvement in sensitivity is analogous to the uncertainty reduction expected for a fixed number of monitored channels.[3] The psychometric slope (Fig. 3, lower panel) has a maximum at small K/M ratios; for small sets of relevant samples ($K \ll M$) the slope is almost invariant (less than 0.5 change at a value near 4) across a 2-log-unit range of the K/M ratio; for large sets the slope changes faster than in the other two cases.

4. DISCUSSION

The major contribution of the uncertainty model was a clear demonstration of how signals from irrelevant local analyzers may affect an observer's performance. Up to now, however, it has been unclear whether uncertainty is the only manifestation of irrelevant signals. Our analysis and computational experiments illustrate that spatial attention may be another mechanism that is sensitive to irrelevant signals. The results presented here show that such distraction of attention and uncertainty produce similar nonlinear effects. This similarity is a result of the common maximum-based selection rule that is embedded in both models.

Both models require small K/M ratios in order to produce a noticeable nonlinear effect. Since attention can be distracted by analyzers from the whole visual field, the ratio K/M potentially can be extremely small, and the consequent nonlinear effect due to distraction of attention would be significant. Experimentally, the number M of the irrelevant analyzers that may distract attention can be controlled by presenting high-contrast distractor stimuli. The signals in the distractor locations will then always exceed the signals from the blank background field. Top-down control of attention is another factor that may reduce the effective number of the monitored analyzers.

To design a model for the distraction of attention, we had somehow to formalize the very notion of attention. Since there is little consensus in this regard,[10–16] we modeled attention based on the extreme premise that spatial attention operates as a "laser pointer" selecting at each moment the signal from a single local analyzer for high-level analysis. If attention turns out to be fuzzy, one can expect that the reported effects will still be manifest, although likely with less magnitude.

Attentional control is another obscure area that required a radical solution. To make the formulaic derivation possible, we assumed that attention is oriented by a low-level preattentive mechanism that singles out the strongest signal in the visual field and directs attention to that location. This simple model of attentional control, again, does not pretend to be complete; however, we argue that any reasonable model of attentional control must take into account the strength of the signals and have a tendency to orient attention toward the strongest. So, in our computational experiment, we evaluated the effect of the maximum rule in its pure form. In real experiments this effect may be attenuated by top-down attentional control, which makes some parts of the visual field more likely to be attended than the others,[17] by inhibition of return,[18] which does not allow attention to come back for some time to the loci already attended, and by repetitive readings for prolonged stimuli. These factors may alter the gain of the nonlinear effect that is due to distraction but are unlikely to eliminate it completely.

To summarize, from a theoretical viewpoint the distraction of attention may be a key factor in affecting the shape of psychometric functions. This possibility needs to be validated in future experiments.

APPENDIX A: NOTATION

In this paper the following notation is used:

r, etc.	random variables,
r,	a particular value of random variable **r**,
$F_r(x)$,	the cumulative distribution function for random variable **r**, $F_r(x) = P(\mathbf{r} \leqslant x)$,
$f_r(x)$,	the probability-density function for random variable **r**, $f_r(x) = [\mathrm{d}F_r(x)]/\mathrm{d}x$
K,	the number of local analyzers stimulated by the test stimulus,
M,	the number of monitored analyzers,
n,	a Gaussian random variable representing the noise in an analyzer, for which $f_n(x) = (1/\sigma\sqrt{2\pi})\exp(-x^2/2\sigma^2)$,
s,	the signal value that would be produced by the stimulus in a relevant analyzer without the noise infusion,
c',	normalized contrast defined as $c' = s/\sigma$, where σ is the noise spread.

APPENDIX B: RELATIONS FROM PROBABILITY THEORY

The following two relations from probability theory are important for understanding the derivations to be given in Appendixes C and D.

The cumulative distribution function for the maximum of two random variables **r** and **u** is derived from their individual cumulative distributions by

$$F_{\max(\mathbf{r},\,\mathbf{u})}(x) = F_r(x)F_u(x). \quad \text{(B1)}$$

The joint probability density function for a random variable **r** and the binary order relation with a random value **u** is provided by

$$f(\mathbf{r} = x,\ \mathbf{u} \leqslant x) = f_r(x)F_u(x), \quad \text{(B2)}$$

$$f(\mathbf{r} = x,\ \mathbf{u} > x) = f_r(x)[1 - F_u(x)]. \quad \text{(B3)}$$

APPENDIX C: DERIVATION OF THE DISTRACTION MODEL FOR THE TWO-ALTERNATIVE FORCED-CHOICE PARDIGM

Consider random variables **t** and $\bar{\mathbf{t}}$, which represent the attended sample in the test and nontest (blank) trials, respectively. Since an aspect of an attended sample is its relevance, the sample value and relevance form a two-dimensional random variable with a corresponding two-dimensional probability density function. The dimension that corresponds to the sample value is continuous; the relevance-related dimension is discrete with two binary values.

In a blank trial, all M samples are produced by the noise variable **n**. Since attention selects the maximum signal, the part of the joint density function related to relevant attended samples is given by

$$f_{\bar{t}}(x,\ \text{relevant}) = f_{\max_{0<i\leqslant K} n_i}(x)F_{\max_{K<i\leqslant M} n_i}(x) = \frac{\mathrm{d}F_n(x)^K}{\mathrm{d}x}F_n(x)^{M-K}. \quad \text{(C1)}$$

In a test trial the output signals of the irrelevant local analyzers are samples of the noise **n**, whereas the relevant analyzers produce samples of noisy signal $s + \mathbf{n}$. The joint density function for relevant samples in this case is

$$f_{t(s)}(x,\ \text{relevant}) = f_{\max_{0<i\leqslant K} s+n_i}(x)F_{\max_{K<i\leqslant M} n_i}(x) = \frac{\mathrm{d}F_n(x-s)^K}{\mathrm{d}x}F_n(x)^{M-K}. \quad \text{(C2)}$$

The probabilities of relevance and irrelevance of the attended sample in the blank and test presentations are thus

$$P_{\bar{t}}(\text{relevant}) = \int_{-\infty}^{+\infty} f_{\bar{t}}(x,\ \text{relevant})\mathrm{d}x, \quad \text{(C3)}$$

$$P_{\bar{t}}(\text{irrelevant}) = 1 - P_{\bar{t}}(\text{relevant}), \quad \text{(C4)}$$

$$P_{t(s)}(\text{relevant}) = \int_{-\infty}^{+\infty} f_{t(s)}(x,\ \text{relevant})\mathrm{d}x, \quad \text{(C5)}$$

$$P_{t(s)}(\text{irrelevant}) = 1 - P_{t(s)}(\text{relevant}). \quad \text{(C6)}$$

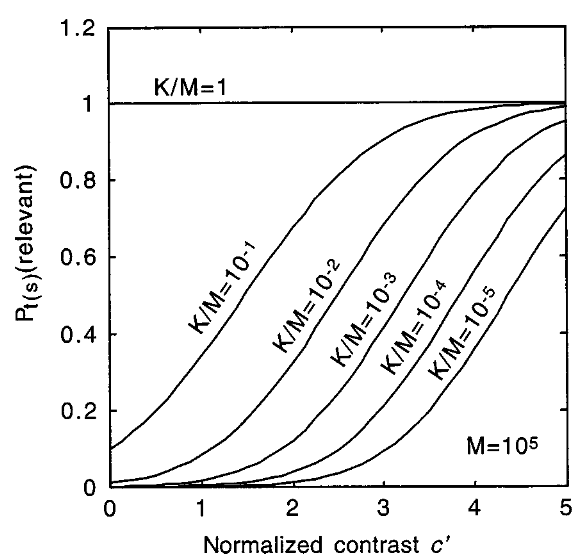

Fig. 4. Each curve illustrates the probability $P_{t(s)}(\text{relevant})$ of the relevant analyzer to be attended in the test trial as a function of signal s (the normalized contrast c') for a particular ratio K/M.

[The reader may notice that $P_{\bar{t}}^{-}(\text{relevant}) = K/M$ and correspondingly $P_{\bar{t}}^{-}(\text{irrelevant}) = 1 - K/M$. Figure 4 illustrates the dependence of the $P_t(\text{relevant})$ profile on the K/M ratio to provide the reader with some intuition.] These probabilities are sufficient to permit us to evaluate the first three rows of Table 1. The last row of the table requires a direct comparison of two relevant decision variables. The parts of the probability density function that are related to the relevant decision variables and the probability of the correct answer are given by

$$f\{[\mathbf{t}(s),\text{relevant}] = x,\ (\bar{\mathbf{t}},\ \text{relevant}) \leqslant x\} = f_{t(s)}(x,\ \text{relevant})F_{\bar{t}}^{-}(x,\ \text{relevant}), \tag{C7}$$

$$P\{[\mathbf{t}(s),\text{relevant}] \geqslant (\bar{\mathbf{t}},\ \text{relevant})\} = \int_{-\infty}^{+\infty} f\{[\mathbf{t}(s),\ \text{relevant}] = x,\ (\bar{\mathbf{t}},\ \text{relevant}) \leqslant x\}\mathrm{d}x. \tag{C8}$$

[The pairs in these formulas are compared by their first components; i.e., $(a,\ \text{relevant}) > (b,\ \text{relevant})$ is equivalent to $a > b$.]

The percent correct in a 2AFC experiment is obtained by summation of the terms corresponding to all four rows of Table 1:

$$P_{\text{corr}}(s) = 0.5P_{t(s)}(\text{irrelevant})P_{\bar{t}}^{-}(\text{irrelevant}) + P_{\bar{t}}^{-}(\text{irrelevant})P_{t(s)}(\text{relevant}) + P\{[\bar{\mathbf{t}}(s),\text{relevant}] \geqslant (\bar{\mathbf{t}},\ \text{relevant})\}. \tag{C9}$$

The third row is zero and therefore is omitted.

APPENDIX D: IMPLEMENTATION OF THE UNCERTAINTY MODEL FOR THE TWO-ALTERNATIVE FORCED-CHOICE PARADIGM

The uncertainty model was implemented in four steps:

$$F_{\bar{t}(x)}^{-} = F_{\max_{0<i\leqslant M} n_i}(x) = F_n(x)^M, \tag{D1}$$

$$f_{t(s)}(x) = f_{\max(\max_{0<i\leqslant K} s + n_i,\ \max_{K<i\leqslant M} n_i)}(x) = \frac{\mathrm{d}F_n(x)^{M-K}F_n(x-s)^K}{\mathrm{d}x}, \tag{D2}$$

$$f[\mathbf{t}(s) = x,\ \bar{\mathbf{t}} \leqslant x] = f_{t(s)}(x)F_{\bar{t}}^{-}(x), \tag{D3}$$

$$P_{\text{corr}}(s) = \int_{-\infty}^{+\infty} f[\mathbf{t}(s) = x,\ \bar{\mathbf{t}} \leqslant x]\mathrm{d}x. \tag{D4}$$

This implementation is substantially simpler than that for the distraction model because it does not require the analysis of separate cases.

ACKNOWLEDGMENTS

The authors thank S. A. Klein, C.-C. Chen, and P. Verghese and anonymous reviewers for helpful comments on the manuscript. This research was supported by National Institutes of Health grant 7890.

L. Kontsevich's e-mail address is lenny@skivs.ski.org.

REFERENCES

1. W. P. Tanner, Jr., "Physiological implications of psychophysical data," Ann. (N.Y.) Acad. Sci. **89**, 752–765 (1961).
2. J. Nachmias and E. C. Kocher, "Visual detection and discrimination of luminous increments," J. Opt. Soc. Am. **60**, 382–389 (1970).
3. D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," J. Opt. Soc. Am. A **2**, 1508–1532 (1985).
4. W. W. Peterson, T. G. Birdsall, and W. C. Fox, "Theory of signal detectability," IRE Trans. Inf. Theory **PGIT-4**, 171–212 (1954).
5. J. M. Foley and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision," Vision Res. **21**, 1041–1053 (1981).
6. G. E. Legge, D. Kersten, and A. E. Burgess, "Contrast discrimination in noise," J. Opt. Soc. Am. A **4**, 391–404 (1987).
7. F. J. J. Blommaert and J. A. J. Roufs, "Prediction threshold and latency on the basis of experimentally determined impulse responses," Biol. Cybern. **56**, 329–344 (1987).
8. M. J. Mayer and C. W. Tyler, "Invariance of the slope of the psychometric function with spatial summation," J. Opt. Soc. Am. A **3**, 1166–1172 (1986).
9. J. Palmer, C. T. Ames, and D. T. Lindsey, "Measuring the effect of attention on simple visual search," J. Exp. Psychol. **19**, 108–130 (1993).
10. U. Neisser, *Cognitive Psychology* (Appleton-Century Crofts, New York, 1967).
11. M. I. Posner, C. R. R. Snyder, and B. J. Davidson, "Attention and the detection of signals," J. Exp. Psychol. **109**, 160–174 (1980).
12. J. M. Wolfe, K. R. Cave, and S. L. Franzel, "Guided search: an alternative to the feature integration model for visual search," J. Exp. Psychol. **15**, 419–433 (1989).
13. A. M. Treisman and G. Gelade, "A feature-integration theory of attention," Cogn. Psychol. **12**, 97–136 (1980).
14. G. Sperling and E. Weichselgartner, "Episodic theory of the dynamics of spatial attention," Psychol. Rev. **102**, 503–532 (1995).
15. C. W. Eriksen and J. D. St. James, "Visual attention within and around the field of focal attention: a zoom lens model," Percept. Psychophys. **40**, 225–240 (1986).
16. D. Kahneman, *Attention and Effort* (Prentice-Hall, Englewood Cliffs, N.J., 1973).
17. K. Nakayama and M. Mackeben, "Sustained and transient components of focal visual attention," Vision Res. **29**, 1631–1647 (1989).
18. M. I. Posner and Y. Cohen, "Components of visual orienting," in *Attention and Performance X*, H. Bouma and D. G. Bouwhuis, eds. (Erlbaum, Hillsdale, N.J., 1984), pp. 531–556.