Toward a Model of Visual Performance: Foundations and Data

Mark S. Rea

I. Introduction

Visual performance has been a central topic of research and discussion in illuminating engineering for many years (see Boyce, 1981; and Rea, 1982a, 1984 for references). This interest is driven largely by practitioners who want to know how lighting affects the performance of workers in industrial and commercial environments. A practitioner might want to know, for example, whether increased illumination levels can lead to faster or more accurate performance in an assembly task. Data and theory that could quantitatively relate lighting and productivity for economic evaluations are, perhaps, the most desirable research goals for lighting practitioners (Clear and Berman, 1981).

Achieving these goals would not, however, be a trivial accomplishment. Many factors associated with the lighting, the task, and the person influence the quantitative relationships between the environment and a person's productivity. Perhaps it is impossible to completely delineate the multiple interactions between lighting, task, and human variables. It should be possible, however, to more narrowly constrict consideration of the number of alternative relationships between these various factors and then to establish plausible and useful calculation procedures of suprathreshold visual performance for lighting practitioners.

The purpose of this paper is to present such a calculation procedure. The plausiblity, and therefore utility, of this proposed model is based upon two propositions. First, *visual* performance must be extracted from *task* performance. The latter is generally an unknown combination of both visual and nonvisual factors that contribute to the behavioral response. The confounding of visual and nonvisual contributions limits the utility of a calculation procedure based on such studies because one does not know how much of the predicted behavior is based upon visual responses under the control of the lighting practitioner and how much is based upon nonvisual (psychological, intellectual, motor, motivational, and emotional) responses evoked by the multitude of other factors. Second, any model of visual performance must be consistent with the literature describing basic visual response. If visual performance can be extracted from task performance, then it should be possible to verify the extraction as an adequate representation of visual processing from data available in the vision literature. Although complete agreement may be difficult, well-established principles in visual sciences limit the number of possible descriptions of suprathreshold visual performance.

This paper describes in some detail, then, the foundations of the proposed model in visual sciences as well as the attempts to extract visual performance from task performance at a simulated realistic task. It is not argued that this visual performance model is complete. A variety of other parameters (*e.g.*, size, luminous uniformity, age of the observer) that are important to the complete specification of visual performance, as well as factors ultimately necessary for making economic evaluations of lighting for task performance (manual dexterity, motivation, aesthetics, interactions with other sense modalities), are not considered. Further, there are some (explicit) assumptions in the model that require further testing. Therefore, this paper and the model serve as a milestone rather than the goal for lighting practitioners. The paper hopefully establishes an appropriate algorithm for a model of suprathreshold visual performance, but it does not provide a complete economic analysis relating lighting and productivity. A subsequent paper (Rea, in press) describes some of the requirements for future studies attempting to describe visual performance more accurately.

II. Foundations

It is clear from a great deal of literature that suprathreshold sensory responses are characterized by compression. Small changes in stimulation about the adaptation level produce corresponding changes in sensation. Although the absolute magnitude of the sensation increases with increasing stimulus intensity, higher stimulation becomes progressively less effective in eliciting incrementally higher sensation. Finally, a level of stimulation will be reached where further increases in stimulus magnitude produce no further increase in sensation magnitude.

Sensory compression is exhibited in a wide variety of invertebrate and vertebrate species, including primates, in various sense modalities, and using a variety of psychophysical and electrophysiological measurement techniques (see Lipetz, 1969; and more recently Hood and Finkelstein, 1979, for references). This general description of sensory compression also charac-

AUTHOR: National Research Council of Canada, Ottawa, Ont.

terizes many other biological "communication" phenomena, including neural reception of internal metabolic changes (Koshland *et al.* 1982). In short, compression is widely accepted as a basic, general description of suprathreshold sensation including, for this paper in particular, suprathreshold visual sensation, or visual response, to luminous modulations.

Suprathreshold sensory compression has the following form, first employed by Naka and Rushton (1966a,b) in describing retinal responses in fish:

$$R/R_{max} = I^n/(I^n + k^n) \quad (1)$$

where: R = response
R_{max} = maximum response
I = stimulus intensity
n = exponent
k = stimulus intensity producing half of maximum response

As already implied, this expression is quite robust in describing visual responses to luminous modulations. The exact values of n, k, and R_{max} in Equation 1 will vary with the experiment and depend upon such factors as the chosen response (*e.g.*, evoked potential or magnitude estimation), the response criterion (*e.g.*, a particular peak or trough from the pattern of evoked potentials), the site of recording (*e.g.*, the retina or the cortex), the chosen stimulus conditions (*e.g*, flashes or gratings), any transformations of stimulus magnitudes (*e.g.*, logarithmic), and the species under investigation (*e.g.*, monkey, cat, or man).

The expression is believed to characterize a physiologically important mechanism in all receptors known as "self-shunting" (Lipetz, 1969). A single photoreceptor (rod or cone) in the retina, for example, is comprised of many photosensitive units connected electrically in parallel. Each small photosensitive unit has what is called an "ionic pump," that produces an electrical potential between the photoreceptor and its environment. Thus, the photoreceptor has a steady-state, resting potential maintained by these ionic pumps. When a photon is captured by one of the photosensitive units in the photoreceptor, the unit depolarizes briefly until the ionic pump can restore the resting potential. This depolarization is like an electrical shunt between the photoreceptor and its environment that produces a voltage drop in the whole photoreceptor. (This, in turn, can produce a chain reaction in subsequent, higher-order neurons that ultimately signal "light" to the brain.) Photon catches by other photosensitive units in the photoreceptor produce similar depolarizations, but each subsequent unit depolarization, or shunting, has relatively less influence on the total response (voltage) of the photoreceptor until, finally, still more photon catches produce no greater response from the photoreceptor. As previously stated, Equation 1 conveniently describes such response compression in photoreceptors.

Compression is exhibited in neurologically highe centers as well as in retinal photoreceptors. One exam ple of visual response compression at the visual co tex, adapted from Albrecht and Hamilton (1982), i shown in **Figure 1.** Presented are the variations in ele trical potentials produced by single neurons in monke cortex in response to the contrast changes of luminous grating. The parameter values for Equatio 1 describing the visual responses are given in the figur caption. Adaptation luminance was 27 cd m^{-2} or 2. log trolands with the 3-mm diameter artificial pupi placed before the monkey's eye.

This particular set of data was used to illustrate visua response compression for a variety of reasons. First, th species investigated, *Macaca fascicularis,* has a visua system almost identical to man (Boynton and Whitten 1970; Valeton and van Norren, 1983), and man is ob viously the species of interest for lighting practice. Se ond, cortical events should be more indicative of th psychophysical responses important to lighting pra tice than neurologically earlier activity. (For exampl retinal responses will not reflect processing from an of the higher visual areas.) Third, the recorded visua responses were to variations in luminous contrast, als employed as stimuli in the present experiment. Fourth the technique for obtaining suprathreshold visua responses was completely different from that employe in the present experiment, so similarities between th results of the two independent experiments would b mutually reinforcing in the development of a supra threshold visual response model. Finally, no other ele trophysiological data with all of the above qualification were found.

From the previous discussion it should not be in ferred that only electrophysiological data can be de scribed by Equation 1. Psychophysical data have als been characterized in this way. Hood and his coworker (Hood *et al.*, 1978; Hood and Finkelstein, 1979; Hoo *et al.*, 1979) have conducted studies specifically address ing the suitability of Equation 1 for describin psychophysical responses. They have found that this ex pression can be used to describe subjective estimate of flash brightnesses, by magnitude estimations, as wel as detection thresholds. Boynton and Whitten (1970 have also used this expression in describing detectio threshold data, and like Hood and his coworkers, hav related psychophysical responses to those recorded ele trophysiologically. Rea (1983) has also used this formula tion to describe the data from a psychophysical stud using measures of speed and accuracy at the numerica verification task employed in the present experiment

Even in studies where this particular formulation ha not been tried, Equation 1 can be used to describe com pressive psychophysical responses. For exampl Blackwell and Blackwell (1980) have presented dat relating the Visibility Level (in this case the contrast of Landolt rings to the probability of correctly identi fying the ring orientation. The dashed line in **Figure 2** is based on Blackwell and Blackwell's "log ogive"; the solid line is based upon Equation 1. Clearly both formulations describe the data well.

Theoretical justification for a particular formulation must rely on criteria other than an adequate curve fit (Valeton, 1983). Both the log ogive and Equation 1 can be fitted to many data exhibiting compression. More work (*e.g.*, Valeton and van Norren, 1983; Valeton, 1983) will have to be conducted to theoretically limit the range of possible "curve fitting strategies." Nevertheless, the robustness of Equation 1 has led to its current preference in the visual sciences for modeling compressive suprathreshold visual responses, and, therefore, forms the basis of the visual performance model presented here.

III. The Numerical Verification Task

In a previous report (Rea, 1981) it was shown that performance, defined in terms of time and errors, at a reading-writing task was strongly affected by the different luminous contrasts between the white paper, the luminance of which was held constant, and the (darker) inks. Further, it was shown that a single function could describe the relationship between performance and contrast, no matter how that contrast was produced. The purpose of the present experiment was to collect more data at the same task using a wider range of background luminances, and, with these data, formulate a visual performance model based upon Equation 1.

A. Methods and procedures

Except for a few minor differences, the stimulus materials, the experimental testing room, and the procedures were like those described in the previous numerical verification experiment (Rea, 1981). Subjects were seated at a desk and asked to compare, from top to bottom, two printed number lists, a reference sheet on the left and a response sheet on the right, for discrepancies. The time to complete the comparisons and the number of errors, both omission and commission, were recorded after every trial.

Data were collected in a black testing room with illumination provided to the work desk from a single fluorescent luminaire having a light emitting aperture of 95.5 x 95.5 cm. A sanded plexiglass diffuser was in the luminaire throughout the experiment. Illumination levels at the center of the task were (approximately) 50, 93, 171 and 700 lx, depending upon the experimental conditions. Illumination levels were adjusted slightly during the experiment to hold task background luminances (*i.e.*, the luminances of the reference sheets) constant at 12, 22, 41, and 169 cd m^{-2} from the position of the subjects' eyes.

Lighting geometry, or direction of illumination, affects task contrast (Rea, 1981). The direction of illumination was changed in this experiment by pivoting the desk (and the subject) about a point at the center of the horizontal task. Eight desk positions were employed in the experiment (**Table 1**).

Subjects were comfortably positioned at a chin rest while seated at the horizontal desk. The viewing angle, from vertical, and distance to the center of the task were 42 degrees and 50 cm, respectively. From the subject's viewing position, the middle digits of the reference and response lists were separated by 7.1 degrees of visual angle. Similarly, the middle digits were 13 and 19 minutes of visual angle wide and high, respectively.

Every reference and response sheet was a column of twenty, five-digit numbers; each five-digit number in the column was separated by a horizontal line. A small calibration square was also printed at the top of each column for photometric measurements. Two sets of 32 reference sheets, printed in black gloss or gray matte ink, were used in this experiment. The photometric qualities of these two sets under gonio and (simulated) hemisphere conditions as well as under similar experimental conditions are described in the earlier report (Rea, 1981). Contrast measurements of the ink calibration squares on one example from each set of reference sheets at the eight experimental desk positions are shown in **Table 1.** Contrast (C_v) is calculated by the formula:

$$C_v = \frac{L_B - L_T}{L_B} \qquad (2)$$

where: L_T is the luminance of the calibration square
and L_B is the luminance of the paper adjacent to the calibration square.

Four groups of 32 high-contrast response lists were employed in this experiment. Three groups were like those used in the previous experiment. The mean frequency of errors in the fourth (new) set was also 3.0. The standard deviation of the list errors in the new set was 1.41; the standard deviations for the other sets were 1.19, 1.32 and 1.48. Contrasts of one example of the response sheets at the eight desk positions are also given in **Table 1.**

Four male and four female subjects between the ages of 19 and 24 (M = 22) years participated in the experiment. These subjects were different from those employed in the earlier experiment, but all had excellent, uncorrected vision as determined by a battery of visual screening tests from a Keystone Ophthalamic Telebinocular. The data from one female subject was not included in the analysis because she dramatically shifted her response strategy partway through the experiment.

The experimental protocol was almost identical to that employed in the earlier experiment. Subjects were asked to compare the reference and response number

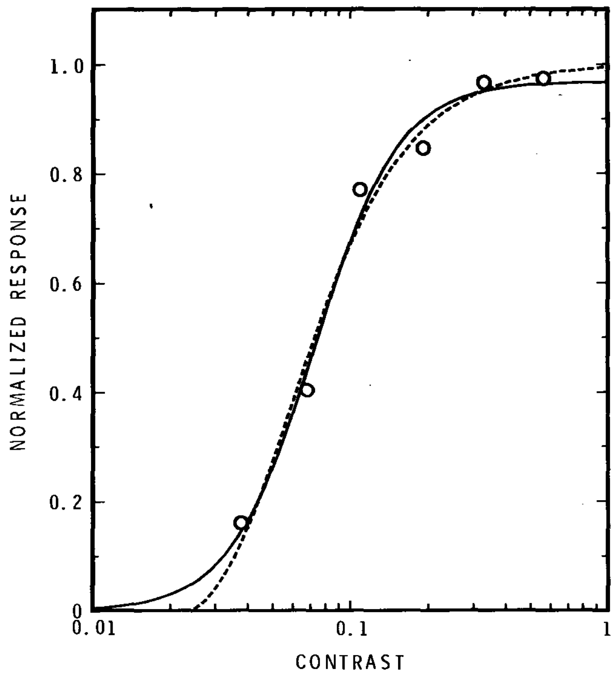

Figure 1—A contrast response function recorded electrophysiologically from the visual cortex (adapted from Albrecht and Hamilton, 1982, Figure 5). Equation 10 (based on Equation 1) describes both the solid and the dashed lines. The dashed line was generated using a threshold term (C_t = 0.023 in Equation 10); the solid line did not include a threshold term (C_t = 0.0). The best fitting, by a least squares criterion, parameter estimates for the data using a threshold term (dashed line) are: n = 1.6, k/L_B = 0.050, and VP_{max} = 1.00. Comparable values without a threshold term (solid line) are: n = 2.7, k/L_B = 0.073, and VP_{max} = 0.97.

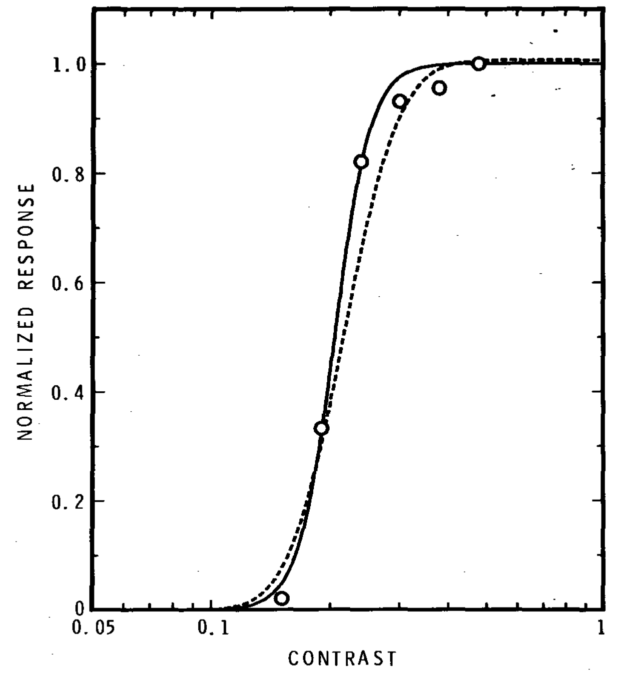

Figure 2—Accuracy in identification of Landolt ring orientation (from Blackwell and Blackwell, 1980, Figure 24). Dashed line is described by the "log ogive" (Equation 3 in Blackwell and Blackwell, 1980) where the free parameters α, γ, and μ (the logarithm of R_{max}) are 0.37, 0.110, and 0.0, respectively. The solid line is based on Equation 1 (or Equation 10, where C_t = 0) where n = 9.61, k/L_B = 0.205, R_{max} (or VP_{max}) = 1.00.

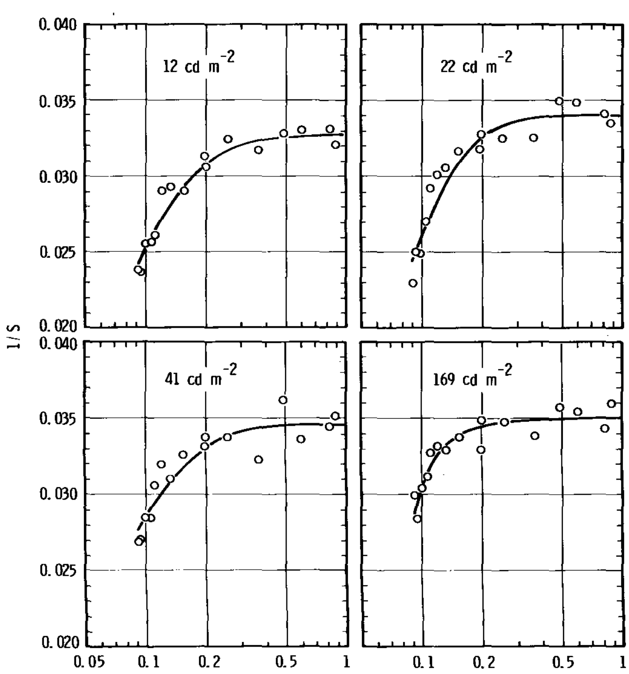

Figure 3—Performance plotted as a function of contrast (scaled logarithmically) at the four background luminances used in the numerical verification task. The reciprocal of the time to compare the reference and response lists (1/S) is used as the performance measure. The solid lines are best fitting curves using Equation 10, which is based on Equation 1, and the four contrast threshold values in Table 3.

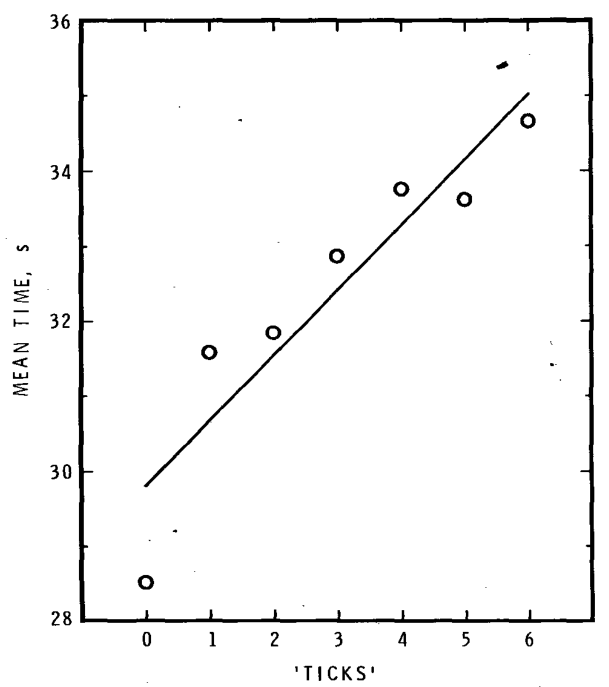

Figure 4—The average time to compare the reference and response lists (S), in seconds (s), plotted as a function of the number of "ticks" (both hits and false positives) that subjects marked on the response lists.

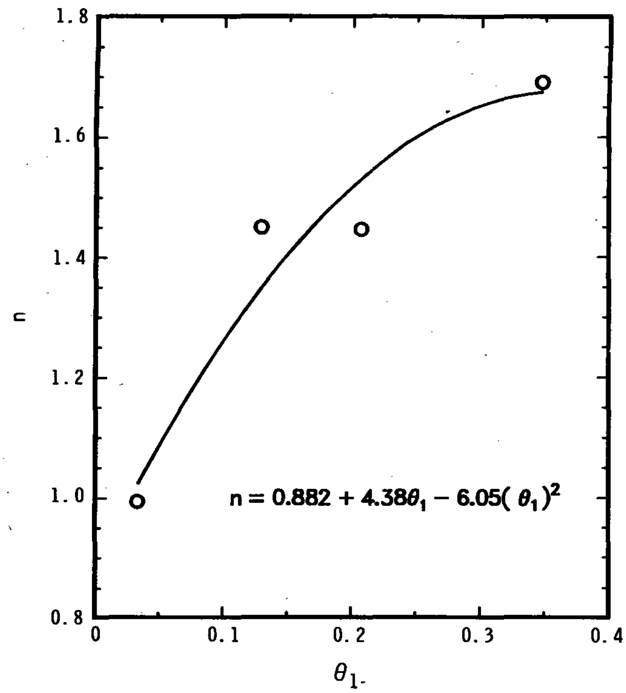

Figure 5—The four estimates of the parameter n, from Table 3, plotted as a function of Θ_1. $\Theta_1 = \log_{10}(\log_{10}(L_B))$, where L_B is the background luminance defined in Equation 2. The solid line, defined by the equation inset into the figure, provides interpolated values of the parameter n between 12 and 169 cd m^{-2}.

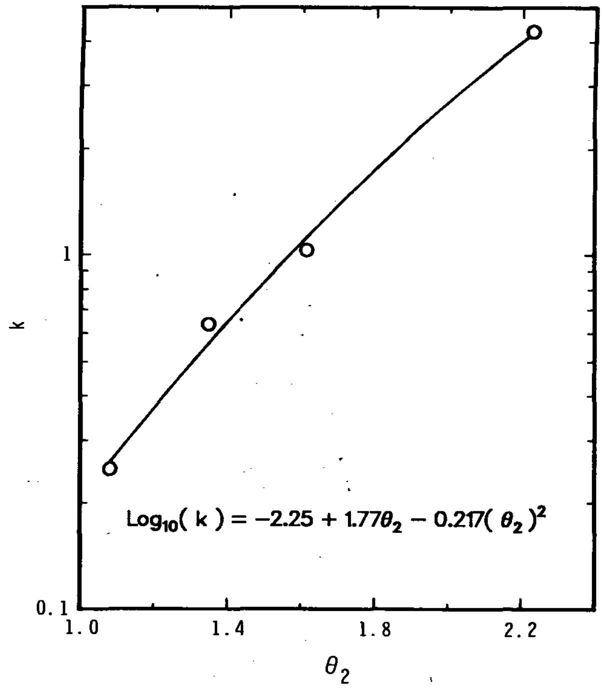

Figure 6—The four estimates of the parameter k, from Table 3, plotted as a function of Θ_2. $\Theta_2 = \log_{10}(L_B)$, where L_B is the background luminance defined in Equation 2. The solid line, defined by the equation inset into the figure, provides interpolated values of the parameter k between 12 and 169 cd m^{-2}.

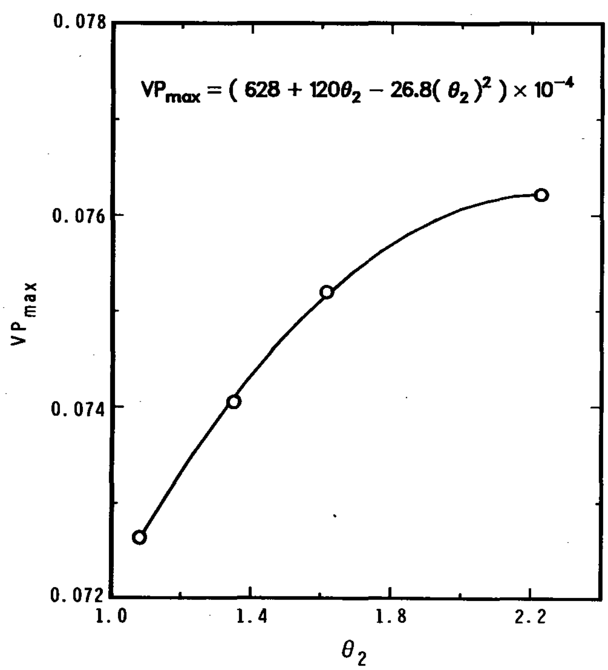

Figure 7—The four estimates of the parameter VP_{max}, from Table 3, plotted as a function of Θ_2. $\Theta_2 = \log_{10}(L_B)$, where L_B is the background luminance defined in Equation 2. The solid line, defined by the equation inset into the figure, provides interpolated values of the parameter VP_{max} between 12 and 169 cd m^{-2}.

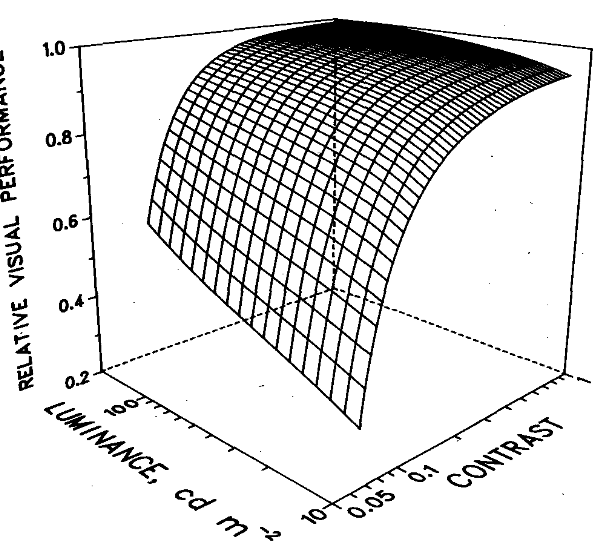

Figure 8—Three-dimensional representation of the visual performance model. Relative visual performance (RVP), scaled linearly, is plotted as a function of contrast (C_v defined in Equations 2 and C2) and luminance (L_B defined in Equation 2), both scaled logarithmically. The levels of RVP for selected values of C_v between 0.08 and 1.0 and values of L_B between 12 and 169 cd m^{-2} are represented by the lines constituting the surface. The levels of RVP for values of C_v between 0.0 and 1.0 and values of L_B between 12 and 169 cd m^{-2} may be calculated from the equations in Appendix C.

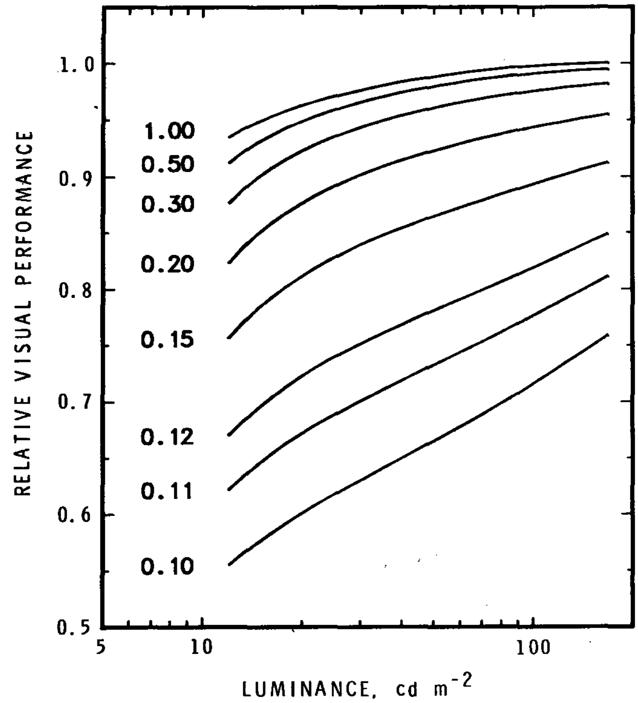

Figure 9—Constant contrast lines from the visual performance model. The curves are labeled in units of contrast (C_v, in Equation 2).

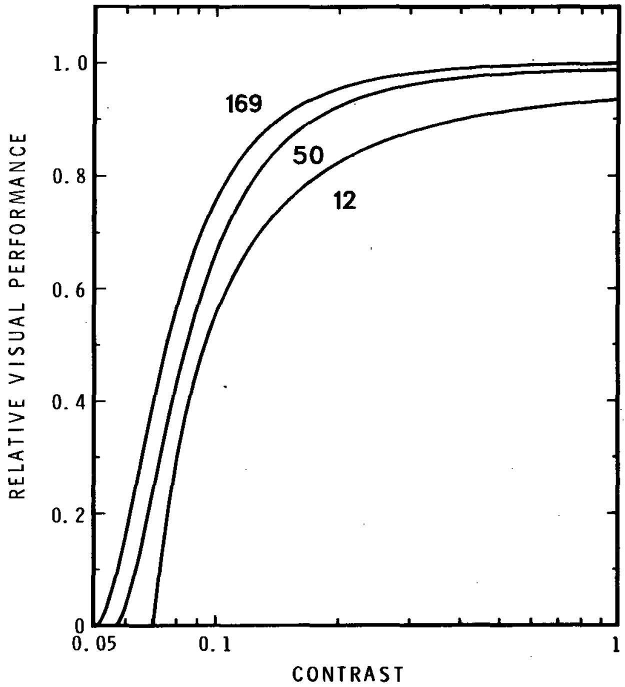

Figure 10—Constant luminance lines from the visual performance model. The curves are labeled in units of background luminance (L_B, in Equation 2).

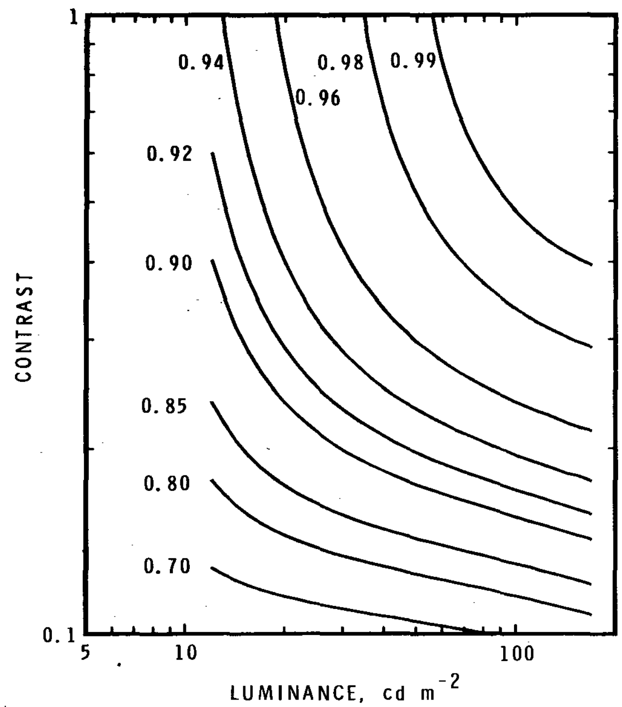

Figure 11—Constant performance lines from the visual performance model. The curves are labeled in units of RVP (Equation C10, in Appendix C). These lines are comparable to threshold functions at various constant criteria.

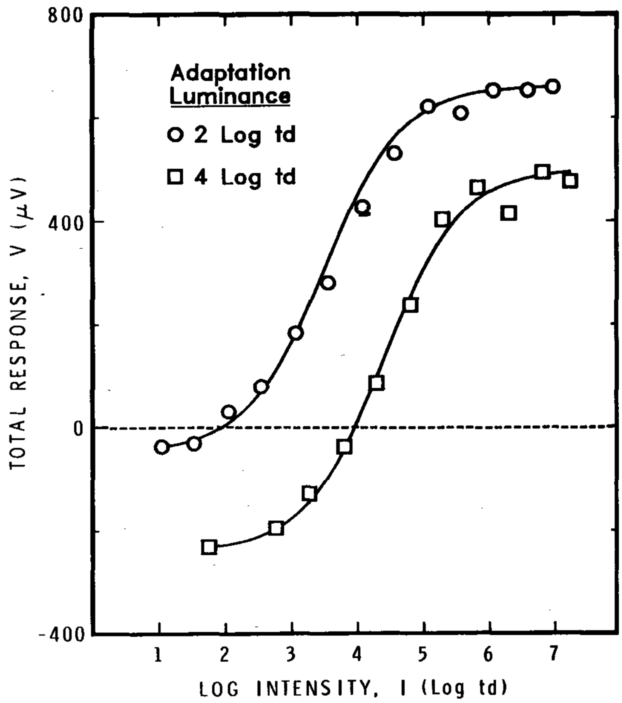

Figure B1—Electrophysiologically recorded cone response functions to luminous increments and decrements at two different adaptation luminances (Adapted from Valeton and van Norren, 1983, Figure 6). A single function, defined by Equation 1 (where n = 0.74, k = 3.2 log trolands, and R_{max} = 1.0), can be shifted in semilogarithmic space (linear response by log luminance) to describe the cone response data at these and other background luminances.

lists as "quickly and accurately" as possible during practice and experimental trials. Discrepancies found by the subjects were marked on the response list with a pen held in the right hand. Total elapsed time (S), misses, and false positives were obtained for each trial. Subjects were not given feedback about their performance during the experiment.

The experimental design was devised to hold constant or minimize, by counterbalancing or randomization, psychological effects like motivation and fatigue. This experimental design can be briefly described as follows. Every subject made the list comparisons under all four illumination levels. To ensure that presentation order was not confounded with illumination level, the illumination levels were counterbalanced across subjects. Eight desk orientations were presented in different orders to the subjects under each of the four illumination levels. Some of the desk orientation presentation orders were counterbalanced across illumination level within subjects. The remainder were counterbalanced across subjects. Four randomly distributed replications of the black gloss and of the gray matte reference lists were presented to the subjects at each of the desk orientations (8 trials/orientation). In all, each subject performed 256 experimental trials during a one-day session lasting approximately six hours. One subject was run per day. It should be noted that the exclusion of the data from one subject prevented complete counterbalancing across the subjects.

B. Results

The data (Appendix A) are quite similar to those published earlier for the numerical verification task (Rea, 1981, 1983; Slater, Perry and Crisp, 1983) and for other tasks (*e.g.*, Poulton, 1969); the relationship between each performance measure and contrast, at a given adaptation level, can be described as suprathreshold sensory compression (Section II). At each background luminance, performance improves rapidly above contrast threshold. As contrast further increases, performance improves but at a decreasing rate until, finally, still higher contrasts produce little, if any, change in the level of performance. The relationships between the reciprocal of the total time to compare the two number lists ($1/S$) at each reference-sheet background luminance (L_B) and contrast (C_v) (**Table A1**) are shown in **Figure 3**.* Best fitting curves, based upon Equation 1, were drawn through the four data sets to help illustrate the different trends and relationships.

The four sets of data in **Figure 3** show the same general form of response compression. Further, each set of data is similar in form to those from cortical neurons in **Figure 1,** at least over the range of contrasts employed in the present experiment (*i.e.*, above C_v = 0.1). While each set of data presented in **Figure 3** shows the same general form of compression, it is also worth noting two important differences between the curves at the different luminance levels. First, for matched contrasts, performance tends to be higher at higher luminances. Second, as contrast increases, performance tends to saturate more quickly at the higher luminances. Essentially, these two trends indicate that performance improves with higher task luminances. Not only does the absolute level of performance increase with higher luminances, but performance is also higher over a larger range of contrasts.

C. Refinements of the 1/S data

Evidence supports the argument that performance in this experiment was largely determined by visual parameters, and the effects of nonvisual factors were minimized. The experimental design, the photometric control of the stimuli, the similarity between time and error response functions, and the similarity between these data and those obtained in other vision experiments (*e.g.*, Albrecht and Hamilton, 1982), all indicate that visual performance, relatively uncontaminated by nonvisual factors, was measured. This evidence might suffice in an argument favoring $1/S$ as the measure of visual performance at the numerical verification task. However, at least two refinements to the $1/S$ data are desirable to improve these data to characterize visual performance.

1. Action time—Weston (1935, 1945) argued that the time taken for subjects to mark targets, the so called "action time," in his Landolt ring search task should be subtracted from the total search time to characterize visual performance more accurately. The same argument can be made for transforming the time data obtained in the present experiment. The average total time taken to compare the two number lists can be plotted as a function of the average number of "ticks" (both hits and false positives) made by subjects in the experiment. As can be seen in **Figure 4,** a simple linear function with a slope of 0.8 can be used to characterize the relationship between these two variables.** Thus, the average time per "tick" is about 0.8 s. It was then a simple matter of subtracting 0.8 s for each hit and false positive in a trial (S_a) from the total time taken by subjects to compare the two number lists (S) in that trial. This transformation of the time data in **Table A1** was used to estimate the so-called reading time (S_r) in the experiment. Thus,

$$S_r = S - S_a \tag{3}$$

2. Time taken to read the response list—The reading time taken to compare the number lists (S_r) is the sum of the time taken to read the reference list (S_{ref}) and the

*Appendix A discusses the reasons for limiting the discussion to $1/S$ as the measure of performance.

**A series of analyses of variance showed that response time ("ticks") was not systematically related to (*i.e.*, did not interact significantly with) any of the experimental variables (luminance, ink pigment density, and desk orientation).

time taken to read the response list (S_{ref}). For a given adaptation luminance (L_B) in this experiment, S_{resp} should be approximately constant because the contrast of the response lists was high and, with the compressive response behavior obtained here and elsewhere (Section II), nearly constant. On the other hand, S_{ref} is expected to vary and to be determined by the contrast of the reference sheet on a particular trial (C_v). When the contrasts of the response and the reference lists are the same, at a given adaptation level, it is reasonable to assume that S_{resp} is equal to S_{ref}. Thus, S_{resp} is a constant (γ) equal to $\frac{1}{2} S_r$ under these conditions. The average response list contrast in **Table 1** was equal to 0.7. From the four curves in **Figure 3** it is possible to estimate the value of 1/S at the four luminance levels for a reference sheet contrast of 0.7 ($C_{0.7}$). Thus, a measure of visual performance (*VP*) can be estimated by taking into account S_a and γ, where:

$$VP = [(S_{L_B,\,C_v} - S_a) - \tfrac{1}{2}(S_{L_B,\,C_{0.7}} - S_a)]^{-1} \qquad (4)$$

$$= (S_{r,\,L_B,\,C_v} - \gamma)^{-1} \qquad (5)$$

$$= (S_{ref,\,L_B,\,C_v})^{-1} \qquad (6)$$

Equation 6 defines the measure of visual performance, VP, for this experiment. Namely, *VP* is the reciprocal of the time taken to read each reference list of a particular contrast (C_v) at a particular adaptation luminance (L_B). **Table 2** presents the values of $\overline{VP}$, the average values of VP.

D. Threshold contrast

Threshold contrasts are estimates of the transition from imperceptible to perceptible contrast at different background luminances, and, therefore, denote the lower limits of contrast perception in a visual performance model. Threshold contrasts have been commonly employed in psychophysical studies (*e.g.*, Graham, 1965; Brindley, 1970) and are even evident in several electrophysiological studies (*e.g.*, Albrecht and Hamilton, 1982; Tolhurst, Movshon and Thompson, 1981). Unfortunately, it is difficult to make estimates of threshold visibility with the procedures employed in the numerical verification task. Subjects tended to quit under difficult conditions so it was impossible to obtain data at the critical, steep transition between threshold contrasts and slightly perceptible contrasts. Threshold contrasts for the number lists had to be estimated at each background luminance by another technique.

Threshold contrasts of the number lists were obtained with a Visual Task Evaluator (VTE) (Blackwell, 1970) over several luminance levels, including the two lowest background luminances used in the numerical verification task. The relationships of these threshold data to background luminance have been described mathematically in an earlier report (Rea, 1982b). However, the VTE elevates the contrast threshold of fine spatial details, like these numbers, relative to "free viewing" (Rea and Ouellette, 1984). To estimate contrast threshold without the intervening VTE optics, a gray number list was placed on the work desk and, by carefully adjusting the lighting geometry, and thus the veiling reflections from the sheet (Section III.A), it was possible to reduce the numbers to a "readability" threshold criterion at a luminance of 67 cd m^{-2}. By normalizing the "readability" contrast threshold function, obtained with VTE, to 0.055 at 67 cd m^{-2}, the contrast value measured following the free viewing method, it was possible to estimate the contrast thresholds (C_t) at the four luminance levels used in the present study (Appendix C, Equation Cl). **Table 3** provides the four values of C_t, derived by these methods.

E. The formulation for visual response compression

Equation 1 may now be rewritten in the following form:

$$VP = [\Delta L^n/(\Delta L^n + k^n)]\ VP_{max} \qquad (7)$$

where:

VP = the level of visual performance, as measured in units of the reciprocal of the time taken to read the reference sheet, from Equation 6;

n and k are as defined in Equation 1;

ΔL is the absolute value of a luminous increment or a luminous decrement from the threshold criterion at a given adaptation luminance.

Thus, for luminous decrements (dark ink on bright background) like those employed in this experiment,

$$\Delta L = L_{T_t} - L_T \qquad (8)$$

where:

L_{T_t} is the luminance of the target at threshold

$= L_B\,(1 - C_t)$

L_T and L_B are as defined in Equation 2;

C_t is the threshold contrast at a given L_B, described in Section III.D, and can be determined from Equation C1 in Appendix C.

Alternatively, Equation 8 can be written as

$$\Delta L = L_B(C_v - C_t) \qquad (9)$$

where:

C_v is the reference sheet contrast.

By using Equation 9 and rearranging terms, Equation 7 can be rewritten, for calculable simplicity, as

$$VP = \{\Delta C^n/[\Delta C^n + (k/L_B)^n]\}\ VP_{max} \qquad (10)$$

where:

$$\Delta C = C_v - C_t$$

There are three unknowns in Equation 10, VP_{max}, n, and k. A nonlinear regression computer routine was used to provide the best estimates by a least squares criterion, of these three parameters at each adaptation luminance. The results of the regressions are presented in **Table 3.**

F. A general algorithm of suprathreshold visual performance

It is desirable to predict visual performance at more background luminances than those actually employed in this study. A series of interpolations was performed on the parameter estimates (n, k, and VP_{max}) in **Table 3. Figures 5** to **7** show the three sets of parameter estimates plotted as a function of Θ_1 or Θ_2, logarithmic transformations of background luminance. Second-degree polynomials were used to estimate, by a least squares criterion, the different parameter values between 12 and 169 cd m^{-2}. It was assumed (see Appendix B) that all of the parameter values in Equation 10 (except C_t) increase monotonically with background luminance. The logarithmic transformations of the abscissae were sometimes employed to limit nonmonotonic behavior by the interpolation routines. The best fitting polynomials and expressions are presented with each figure.

These expressions enable one to calculate relative visual performance (RVP) for alphanumeric reading material of any contrast, where the paper is brighter than the ink and the luminance of the paper is between 12 and 169 cd m^{-2} (or, for white matte paper, between illuminances of 50 and 700 lx). Maximum RVP corresponds to a contrast of 1.0 at 169 cd m^{-2}. Appendix C describes the procedures for calculating RVP.

Figure 8 is a graphical representation of RVP as a function of contrast and luminance, both scaled logarithmically. Examination of this surface from several perspectives illustrate more clearly the influence of contrast and luminance on visual performance according to this model.

RVP is plotted as a function of luminance for several contrasts in **Figure 9.** Two features are readily apparent from this figure. First, visual performance changes very little from medium to high contrasts. For example, there is virtually no difference in RVP at contrasts between 0.5 and 1.0; changes in contrast are much more dramatic between lower contrasts (*e.g.*, between 0.10 and 0.15). Second, the changes in RVP across luminance at high contrasts (*e.g.*, 0.5) are slight, whereas, luminance is much more important to RVP at low contrasts (*e.g.*, 0.15). These two points lead to the conclusion that visual performance will be relatively high and stable for medium to high contrast tasks, but veiling reflections and reduced illumination levels can be very important for low contrast tasks.

Figure 10 shows RVP plotted as a function of contrast at three luminance levels. Several points are worth noting in this figure. First, visual response compression is readily apparent in this figure; RVP is characterized by a steep initial slope and saturation as contrast increases. Second, the point of saturation is higher at higher luminances; RVP is always slightly better at higher luminances, even for medium and high contrasts. Third, the "knees" of the RVP response functions are higher and more prominent at higher luminances, indicating that a high level of visual performance can be maintained at lower contrasts the greater the level of illumination. Finally, this figure shows the shift in absolute contrast threshold, the lower limit of RVP, with higher luminance.

Figure 11, augments this last point concerning contrast threshold. Contrast, scaled logarithmically, is plotted as a function of luminance, also scaled logarithmically; the curves trace constant levels of RVP. These constant criterion curves are, in principle, *identical* to directly obtained threshold functions relating contrast and luminance. It is important to note in **Figure 11**, that the highest constant criterion levels of visual performance are unattainable at lower luminances, even at the highest contrasts.

IV. Discussion

Figures 8 through **11** help illustrate the predictions of the model as well as the similarity of these predictions to findings from a variety of electrophysiological and psychophysical studies.

A. Threshold contrast

Many electrophysiological and psychophysical studies have, for example, used threshold behavior to describe visual response (*e.g.*, Boynton and Whitten, 1970; Albrecht and Hamilton, 1982; Tolhurst *et al.*, 1981). Threshold contrast has been shown repeatedly to decrease with adaptation luminance (*e.g.*, Blackwell, 1946), and the model makes predictions of threshold consistent with these observations.

The exact shapes of the various threshold functions in **Figure 11** may be different from other published data for a variety of reasons (Rea, 1982b). The independent (alphanumeric stimuli) and dependent (time) variables on which the model is based are somewhat unorthodox for the vision community, although they were specifically chosen to be "realistic" and therefore relevant to illuminating engineering. Further, these "threshold" curves were derived from the model and reflect the suprathreshold data obtained in the numerical verification task. They were not obtained from a lengthy threshold experiment in which subjects were asked to adopt various constant criteria and judge the visibility of stimuli at different contrasts and luminances. In principle, there should be no difference in the data obtained by either procedure, but it seems impractical, if not impossible, to expect subjects to adopt and then maintain all of the various criteria necessary to generate the data required for functions comparable to those in **Figure 11.** Therefore, it is unlikely that a successful model of suprathreshold visual performance can be developed from threshold data alone.

The CIE model of visual performance (CIE, 1981a,b) is based on just one threshold function. Higher constant criterion functions, called Visibility Levels (VLs), are obtained by multiplying the single threshold curve by fixed multiples (CIE, 1981a,b). This assumption, that the various VL curves are parallel in log contrast versus log luminance space, is not justified by the data obtained in this experiment and is therefore inconsistent with the predictions from the RVP model (Appendix C).* In this respect, at least, the RVP model presented here is probably better at making *a priori* predictions, and, therefore, in attempting to describe suprathreshold visual performance at a realistic, reading-writing task.

B. Suprathreshold response

Studies directly measuring suprathreshold visual performance have shown that it will increase with both contrast and luminance (*e.g.*, Weston, 1935, 1945; Boyce, 1973, Smith and Rea, 1980; McNelis, 1973). The RVP model makes predictions similar to the data generated in these studies by predicting an interaction between luminance and contrast. The model predicts that suprathreshold visual performance will increase only slightly with luminance when the contrast is high and more strongly when contrast is low. As with the threshold data, however, the model predictions and the data from these various studies do not always agree quantitatively. It will be argued in a later paper (Rea, in press), however, that the results of properly controlled visual performance studies (*e.g.*, McNelis, 1973) should yield results very similar to the predictions generated by the model. The contrast response functions obtained from cortical cells in monkeys (**Figure 1**), for example, are both qualitatively and quantitatively similar to predictions from the model.** This marked similarity, known as "psychophysical isomorphism" (Kaufman, 1974; Dawis, 1981), is likely the result of the experimental control over the many "nonvisual" factors that can contribute to the measured responses. Still, it should be realized that the RVP model is based largely upon the latency of response and most electrophysiological data, *e.g.*, Albrecht and Hamilton (1982), are based upon the amplitude of response. The relationships between the amplitude and latency response functions can be complex (Ermolaev and Kleinman, 1983; van der Tweel *et al.*, 1979), but similarities have been shown (Vaughn *et al.*, 1966). More studies in electrophysiology may more clearly delineate the relationship between amplitude and latency.

C. Subjective impressions

It is also important to note, that the model seems to be consistent with our everyday experiences. One has little problem reading high contrast newsprint under virtually any domestic or commercial lighting system. As the model predicts, high contrast reading materials, like a newspaper, are relatively insensitive to variations in illuminance or contrast reducing veiling reflections. On the other hand, low contrast reading materials, like some blueprints or photoduplications, are quite sensitive to such variations. It is much more important for people viewing such material to have lighting systems that give high illumination and reduce disability glare and veiling reflections. The model quantitatively underscores the importance of these everyday occurrences.

V. Conclusions

These broad confirmations of the model aside, however, it must be stressed that the model still requires extension and testing. Not all of the important stimulus parameters have been studied, and there are a variety of inadequately tested assumptions built into the model. For example, the size of the printed text was not varied in the experiment. Also, targets brighter than the background (as with visual display terminals) have not been studied. Importantly too, it was assumed (Appendix B) that n, k, (or k_0) and VP_{max} should increase monotonically with luminance (**Figures 5, 6** and **7**). Further, it was assumed that contrast threshold data from different subjects could be linked with the suprathreshold data reported in this paper to make a more complete model.

Although these and other issues need to be investigated, it seems unlikely that the model is grossly inaccurate in representing suprathreshold visual performance. Visual response compression, which forms the basis of the model, is a widely observed phenomenon. The model makes predictions consistent with documented trends in threshold and suprathreshold behavior and, importantly, is consistent with our everyday experiences. Finally, a sensitivity analysis showed that reasonable variations in estimates of the parameter, n, did not significantly affect the model predictions over the range of contrasts and luminances encountered in commercial environments.

There is no question that the model must be extended and revised to some degree, but, based upon the

*It is possible to change the shapes of the threshold curves in the CIE model, depending upon a variety of estimates of free parameters established after results are obtained (Rea, 1982b). For a given set of parameters, however, the constant criterion (VL) function would be parallel in a log contrast versus log luminance coordinate system.

Albrecht and Hamilton (1982) did not use a threshold term when fitting their data with Equation 1. Adding an appropriate threshold term makes the parameter estimates comparable to those for the model (Figure 1**). Conversely, the deletion of a threshold term when fitting the present data, makes the parameter estimates like those published by Albrecht and Hamilton.

arguments made above, the model would appear to have utility for lighting practitioners as an interim algorithm for calculating relative visual performance at alphanumeric reading tasks. Therefore, the RVP model should be considered as a milestone toward the ultimate goal for an economic model of task performance.

VI. Acknowledgments

This paper is a contribution from the Division of Building Research, National Research Council of Canada.

VII. References

Albrecht, D.G., Hamilton, D.B., "Striate cortex of monkey and cat: Contrast response function," *J. Neurophysiology*, 48 (1), pp. 217–237, July 1982.

Blackwell, H.R., "Contrast thresholds of the human eye," *JOSA*, 36 (11), pp. 624–643, November 1946.

Blackwell, H.R., "Development of procedures and instruments for visual task evaluation," *Illuminating Engineering*, 65, pp. 267–291, April 1970.

Blackwell, H.R., Blackwell, O.M., "Population data for 140 normal 20–30 year olds for use in assessing some effects of lighting upon visual performance (Compendium Paper No. 1)" 1980 Complete Compendium of IERI Technical Data in Support of International Commission on Illumination, Publication CIE No. 19/2, 1980.

Boyce, P.R., "Age, illuminance, visual performance and preference" *Lighting Research and Technology*, 5 (3), pp. 125–144, 1973.

Boyce, P.R., *Human Factors in Lighting*, Macmillan Publishing Co., New York, 1981.

Boynton, R.M., Whitten, D.M., "Visual adaptation in monkey cones: Recordings of late receptor potentials," *Science*, 170, pp. 1423–1426, December 1970.

Brindley, G.S., *Physiology of the retina and visual pathway*, Edward Arnold Ltd., London, Second Edition, 1970.

Clear, R., Berman, S., "Cost-effective visibility-based design procedures for general office lighting," *J. Illuminating Engineering Society*, 10 (4), pp. 228–236, July 1981.

Commission internationale de l'éclairage (CIE), Technical Committee 3.1, "An analytic model for describing the influence of lighting parameters upon visual performance: Vol. I: Technical Foundations," Publication CIE No. 19/2.1, 1981.

*Commission internationale de l'*éclairage (CIE), Technical Committee 3.1, "An analytic model for describing the influence of lighting parameters upon visual performance; Vol. II: Summary and Application Guidelines," Publication CIE No. 19/2.2, 1981.

Dawis, S.M., "The compression model: A re-examination," *Vision Research*, 21, pp. 1511–1515, 1981.

Ermolaev, R.Y., Kleinman, D., "The effect of background illumination on pattern onset visual evoked potentials," *Electroencephalography and Clinical Neurophysiology*, 55, pp. 546-556, 1983.

Graham, C.H., ed., *Vision and visual perception*, John Wiley and Sons Inc., New York, 1965.

Hood, D.C., Ilves, T., Maurer, E., Wandell, B., Buckingham, E., "Human cone saturation as a function of ambient intensity: A test of models of shifts in the dynamic range," *Vision Research*, 18, pp. 983–993, 1978.

Hood, D.C., Finkelstein, M.A., Buckingham, E., "Psychophysical tests of models of the response function," *Vision Research*, 19, pp. 401–406, 1979.

Hood, D.C., Finkelstein, M.A., "Comparison of changes in sensitivity and sensation: Implications for the response-intensity function of the human photopic system," *J. of Experimental Psychology: Human Perception and Performance*, 5 (3), pp. 391–405, August 1979.

Kaufman, L., *Sight and mind: An introduction to visual perception*, Oxford University Press, 1974.

Koshland, D.E., Goldbeter, A., Stock, J.B., "Amplification and adaptation in regulatory and sensory systems," *Science*, 217, pp. 220– 225, July 1982.

Lipetz, L.E., "The transfer functions of sensory intensity in the nervous system," *Vision Research*, 9, pp. 1205–1234, 1969.

McNelis, J.F., "Human performance: A pilot study." *J. Illuminating Engineering Society*, 2 (3), pp. 190–196, April 1973.

Naka, K.I., Rushton, W., "S-potentials from colour units in the retina of fish (*Cyprinidae*)," *J. Physiology*, 185, pp. 536–555, 1966a.

Naka, K.I., Rushton, W., "An attempt to analyse colour reception by electrophysiology," *J. Physiology*, 185, pp. 556–586, 1966b.

Poulton, E.C., "Skimming lists of food ingredients printed in different brightness contrasts," *J. of Applied Psychology*," 53 (6), pp. 498–500, 1969.

Rea, M.S., "Visual performance with realistic methods of changing contrast," *J. Illuminating Engineering Society*," *10 (3), pp. 164–177, April 1981.*

Rea, M.S., "An overview of visual performance," Lighting Design and Application, 12 (11), pp. 35–41, November 1982a.

Rea, M.S., "The validity of the relative contrast sensitivity function for modelling threshold and suprathreshold responses," *The integration of visual performance criteria into the illumination design process*, Public Works Canada, Ottawa, Canada, January 1982b.

Rea, M.S., "Effects of Haidinger's brushes on visual performance," *J. Illuminating Engineering Society*," 12 (3), pp. 197–203, April 1983.

Rea, M.S., "Overview of task performance in the illuminated environment," *Illumination Engineering Learning Aid*, Vol. II, June 1984.

Rea, M.S., Ouellette, M.J., *An assessment of the Blackwell visual task evaluator, model 3X*, Division of Building Research, National Research Council Canada, DBR Paper No. 1160, 1984.

Slater, A.I., Perry, M.J., Crisp, V.H.C., "The applicability of the CIE visual performance model to lighting design," *Light and Lighting '83, Proceedings of the 20th Session*, Amsterdam, Publication CIE No. 56, pp. D114/1–4, 1983.

Smith, S.W., Rea, M.S., "Relationships between office task performance and ratings of feelings and task evaluations under different light sources and levels," *Proceedings of the 19th Session*, Kyoto, CIE Publication No. 50, pp. 207–211, 1980.

Tolhurst, D.J., Movshon, J.A., Thompson, I.D., "The dependence of response amplitude and variance of cat visual cortical neurons on stimulus contrast," *Experimental Brain Research*, 41, pp. 414– 419, 1981.

van der Tweel, L.H., Estevez, O., Cavonius, C.R., "Invariance of the contrast evoked potential with changes in retinal illuminance," *Vision Research*, 19, pp. 1283–1287, 1979.

Valeton, J.M., "Photoreceptor light adaptation models: An evaluation," *Vision Research*, 23 (12), pp. 1549–1554, 1983.

Valeton, J.M., van Norren, D., "Light adaptation of primate cones: An analysis based on extracellular data," *Vision Research*, 23 (12), pp. 1539–1547, 1983.

Vaughan, H.G., Costa, L.D., Gilden, L., "The functional relation of visual evoked response and reaction time to stimulus intensity," *Vision Research*, 6, pp. 645–656, 1966.

Weston, H.C., "The relation between illumination and industrial efficiency: The effect of size of work," Her Majesty's Stationery Office, London, pp. 1–14, 1935.

Weston, H.C., "The relation between illumination and visual efficiency: The effect of brightness contrast," Industrial Health Research Board, Great Britain Medical Research Council, London, Report No. 87, pp. 1– 35, 1945.

Acknowledgments

The author wishes to express his gratitude to M.J. Ouellette for input at every level in the preparation of this manuscript. Thanks also go to B. Guzzo for running the subjects and the following individuals who made many helpful comments on earlier drafts: S. Berman, R. Biggs, P. Boyce, R. Clear, R. Corth, V. Crisp, D. DiLaura, A. Levy, I. Lewin, M. Perry, A. Robertson, P. Russell, A. Slater, S. Squillace, G. Webber.

This paper is a contribution from the Division of Building Research, National Research Council of Canada.

Table 1—Stimulus sheet contrasts

Desk Angle*	Reference Sheets: Black gloss	Gray matte	Response Sheet
0	0.199	0.095	0.670
7.5	0.202	0.092	0.672
15	0.260	0.100	0.675
20	0.371	0.107	0.684
25	0.491	0.112	0.698
30	0.606	0.121	0.714
45	0.835	0.134	0.759
90	0.894	0.156	0.826

*These values are the clockwise azimuth angles, in degrees, of the vertical plane defining the subject's line of sight to the center of the task with respect to the center of the luminaire. The 0-degree azimuth angle, for example, corresponds to the vertical plane bisecting the subject's nose, the center of the task and the center of the luminaire.

Table 2—Visual Performance ($\overline{VP}$)

C_v	L_B: 12	22	41	169
0.092	0.0389	0.0362	0.0463	0.0554
0.095	0.0386	0.0415	0.0467	0.0504
0.100	0.0441	0.0412	0.0513	0.0570
0.107	0.0443	0.0472	0.0511	0.0597
0.112	0.0457	0.0543	0.0585	0.0660
0.121	0.0551	0.0574	0.0638	0.0676
0.134	0.0560	0.0590	0.0602	0.0665
0.156	0.0551	0.0632	0.0666	0.0699
0.199	0.0640	0.0638	0.0687	0.0667
0.202	0.0610	0.0680	0.0712	0.0749
0.260	0.0688	0.0667	0.0711	0.0745
0.371	0.0658	0.0670	0.0650	0.0705
0.491	0.0703	0.0778	0.0826	0.0789
0.606	0.0717	0.0775	0.0706	0.0774
0.835	0.0720	0.0738	0.0743	0.0726
0.894	0.0671	0.0712	0.0775	0.0801

Entries are values of visual performance ($\overline{VP}$), the average of the reciprocal of the time required to read the reference sheet [$(S_{ref})^{-1}$]. L_B is the luminance of the reference sheet in units of cd m^{-2}. C_v is the contrast of the numbers as defined in Equation 2.

Table 3—VP parameters

C_t	L_B	n	k	VP_{max}
0.0702	12.1	0.993	0.249	0.0726
0.0632	22.4	1.45	0.634	0.0740
0.0581	41.2	1.45	1.03	0.0752
0.0506	169	1.69	4.27	0.0762

Appendix A

Table A1 gives the response measure (total time, misses, false positives and score) averages for the combinations of sixteen contrasts and four background luminances produced in the experiment. Analyses of variance (ANOVAs) for the experimental design variables are presented in **Table A2** for each response measure. The response measure called score is based upon the other three and calculated by the formula:

$$\text{Score} = (T - E)\ 100\ /\ (S + 5) \tag{A1}$$

where:

T = total number of comparisons per trial (always 20)

E = number of errors, both misses and false positives, committed per trial

S = total time to complete the comparisons per trial (in seconds).

This measure was devised by Smith and Rea (1980) to monotonically increase with visibility and to incorporate both speed and accuracy, traditional measures of visual performance (Weston, 1935, 1945). The score formulation is a somewhat arbitrary measure of performance, most strongly determined by total time (S). (There was a correlation of −0.95 between the trial scores and the trial times in this experiment.) In fact, errors play only a minor role in determining the score values. Nevertheless, and as noted in the earlier report (Rea, 1981), the contrast response functions for total time, misses, and false positives are nearly identical in these experiments.* Ideally then, the relative weightings of the primary measures would make little difference in a strictly linear characterization of performance combining time and errors. Score as defined in Equation A1 is not linear, however. The constant (5 s) in the denominator of Equation A1 introduces a nonlinearity into the score formulation. Although the constant is relatively unimportant at most response times observed in this experiment (usually longer than 25 s), it would obviously be more important at shorter response times.

It is still advantageous to minimize the importance of errors in a score formulation for this experiment, despite the similarity of the contrast response functions for all of the primary measures. The absolute levels of both error measures are low in these experiments. Even small, random fluctuations in the subjects' behaviors could strongly influence trial scores if these integer, dependent variables were more strongly weighted. To limit the influence of random perturbations in the false

* The marked similarity between the primary measures total time, misses and false positives, indicates that the classic, psychological "speed-accuracy trade-off" was not reflected in the response averages. This identity of the total time and the error responses is strong, but indirect, support for the hypothesis that these averaged results are largely based upon visual rather than nonvisual factors.

positive or miss responses, total time *should be* weighted more heavily than errors to characterize visual performance, assuming time and errors vary the same way with contrast. As stated in the previous paragraph, this was the case.

In summary, only measures based upon trial times are discussed as indices of visual performance in this report because a) the score formulation includes a small nonlinearity, b) misses and false positives are less reliable measures of performance than total time, c) manuscript brevity, and, importantly, d) the contrast response function for the reciprocal of time is not substantively different from that for score, the reciprocal of misses or the reciprocal of false positives. The reciprocal of trial times was used because it is more convenient to think of the measure of performance increasing with increasing visibility.

Table A1—Means and standard errors (SE) for the four dependent variables used in the experiment.

	Score		Time (S)		Misses		False Positives	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
L_B = 12 cd m^{-2}								
C_v								
0.092	42.2	1.55	41.9	1.69	0.786	0.173	0.071	0.050
0.095	42.0	1.37	42.2	1.57	0.643	0.156	0.107	0.079
0.100	45.7	1.94	38.9	1.86	0.786	0.149	0.071	0.050
0.107	45.0	1.58	38.8	1.47	0.786	0.166	0.071	0.050
0.112	46.0	1.83	38.2	1.65	0.893	0.165	0.036	0.036
0.121	50.6	2.03	34.4	1.45	0.786	0.173	0.036	0.036
0.134	50.8	1.66	34.1	1.44	0.714	0.169	0.036	0.036
0.156	51.2	1.55	34.4	1.16	0.321	0.104	0.000	0.000
0.199	54.9	1.99	31.9	1.37	0.464	0.131	0.000	0.000
0.202	54.7	2.12	32.6	1.74	0.357	0.092	0.000	0.000
0.260	56.6	1.98	30.8	1.23	0.393	0.119	0.000	0.000
0.371	55.7	2.04	31.5	1.53	0.429	0.120	0.036	0.036
0.491	57.5	2.15	30.5	1.52	0.429	0.174	0.000	0.000
0.606	58.5	2.16	30.2	1.46	0.250	0.098	0.000	0.000
0.835	56.9	1.88	30.1	1.27	0.607	0.181	0.000	0.000
0.894	56.2	1.93	31.2	1.24	0.321	0.127	0.000	0.000
L_B = 22 cd m^{-2}								
C_v								
0.092	41.8	1.87	43.4	2.41	0.536	0.131	0.393	0.130
0.095	44.9	1.77	39.9	2.14	0.679	0.163	0.107	0.060
0.100	44.0	1.46	40.0	1.66	0.750	0.160	0.071	0.050
0.107	48.6	2.09	36.9	1.90	0.536	0.120	0.071	0.050
0.112	51.6	1.60	34.2	1.55	0.393	0.165	0.036	0.036
0.121	53.0	1.99	33.2	1.51	0.571	0.166	0.000	0.000
0.134	54.4	2.05	32.7	1.49	0.250	0.098	0.036	0.036
0.156	55.1	1.73	31.6	1.34	0.464	0.120	0.000	0.000
0.199	55.6	1.55	31.4	1.10	0.214	0.079	0.000	0.000
0.202	56.5	1.66	30.4	1.12	0.464	0.120	0.000	0.000
0.260	56.0	1.68	30.7	1.09	0.393	0.119	0.071	0.050
0.371	57.4	2.11	30.7	1.45	0.321	0.104	0.000	0.000
0.491	60.9	2.08	28.6	1.38	0.286	0.087	0.000	0.000
0.606	59.8	1.74	28.7	1.04	0.357	0.117	0.000	0.000
0.835	58.4	1.79	29.3	1.10	0.500	0.109	0.000	0.000
0.894	57.3	1.50	29.8	1.02	0.464	0.131	0.000	0.000

Table A1 (continued)

	Score		Time (S)		Misses		False Positives	
	Mean	**SE**	**Mean**	**SE**	**Mean**	**SE**	**Mean**	**SE**
L_B = 41 cd m^{-2}								
C_v								
0.092	46.6	1.64	37.1	1.36	0.750	0.175	0.179	0.074
0.095	47.5	1.64	36.9	1.46	0.571	0.188	0.107	0.079
0.100	48.8	1.59	35.0	1.19	0.857	0.152	0.107	0.079
0.107	50.0	1.60	35.1	1.41	0.500	0.121	0.036	0.036
0.112	53.0	1.49	32.6	1.08	0.429	0.120	0.036	0.036
0.121	55.0	1.63	31.2	1.13	0.571	0.130	0.000	0.000
0.134	55.1	1.80	32.1	1.26	0.143	0.067	0.000	0.000
0.156	56.4	1.73	30.6	1.20	0.500	0.109	0.000	0.000
0.199	57.7	1.67	30.1	1.09	0.214	0.079	0.000	0.000
0.202	57.5	1.68	29.6	0.97	0.536	0.150	0.000	0.000
0.260	58.9	1.92	29.6	1.18	0.214	0.079	0.000	0.000
0.371	56.7	1.95	30.9	1.35	0.321	0.104	0.000	0.000
0.491	60.7	1.66	27.6	0.89	0.571	0.120	0.000	0.000
0.606	58.4	1.88	29.7	1.18	0.321	0.127	0.000	0.000
0.835	59.5	2.02	29.0	1.16	0.357	0.128	0.000	0.000
0.894	59.9	1.84	28.4	1.02	0.464	0.109	0.000	0.000
L_B = 169 cd m^{-2}								
C_v								
0.092	52.0	1.84	33.3	1.42	0.679	0.137	0.071	0.050
0.095	49.8	1.80	35.1	1.57	0.643	0.180	0.107	0.060
0.100	52.1	1.58	32.8	1.04	0.607	0.107	0.107	0.060
0.107	54.3	2.01	32.0	1.23	0.536	0.150	0.000	0.000
0.112	56.3	1.73	30.4	1.08	0.536	0.109	0.000	0.000
0.121	57.6	1.81	30.1	1.06	0.321	0.116	0.000	0.000
0.134	57.0	1.92	30.3	1.18	0.464	0.120	0.000	0.000
0.156	57.7	1.51	29.6	0.94	0.429	0.130	0.000	0.000
0.199	56.6	1.63	30.3	1.02	0.464	0.167	0.000	0.000
0.202	59.7	1.86	28.6	1.05	0.393	0.139	0.036	0.036
0.260	59.2	1.68	28.7	0.97	0.464	0.141	0.000	0.000
0.371	58.6	1.99	29.5	1.17	0.393	0.119	0.036	0.036
0.491	60.9	1.80	28.0	0.99	0.393	0.119	0.000	0.000
0.606	60.0	1.69	28.2	0.89	0.464	0.120	0.000	0.000
0.835	60.1	2.36	29.0	1.31	0.357	0.106	0.000	0.000
0.894	61.5	1.75	27.8	0.97	0.286	0.101	0.000	0.000

Table A2—Analyses of variance (ANOVAs) for the four dependent variables

1. Dependent Variable = SCORE

EFFECT	SS	DF	MS	F	P
L	7,514.1	3	2,504.7	78.628	.001
O	9,335.3	7	1,333.6	41.865	.001
I	25,829	1	25,829.0	810.84	.001
A	93,437	6	15,573.0	488.87	.001
LO	1,043.5	21	49.690	1.5599	–
LI	1,390.8	3	463.62	14.554	.001
LA	6,278.2	18	348.79	10.949	.001
OI	3,333.9	7	476.27	14.951	.001
OA	2,197.2	42	52.315	1.6423	.005
IA	1,497.3	6	249.55	7.8338	.001
LOI	525.82	21	25.039	0.78602	–
LOA	7,826.9	126	62.118	1.9500	.001
OIA	667.66	42	15.897	0.49903	–
LIA	960.93	18	53.385	1.6759	.05
LOIA	2,855.7	126	22.664	0.71147	–

2. Dependent Variable = TIME (S)

EFFECT	SS	DF	MS	F	P
L	4,608.8	3	1,536.3	89.798	.001
O	5,029.0	7	718.43	41.994	.001
I	12,034	1	12,034.0	703.40	.001
A	49,333	6	8,222.2	480.60	.001
LO	751.82	21	35.801	2.0926	.005
LI	1,093.5	3	364.51	21.306	.001
LA	4,220.5	18	234.47	13.705	.001
OI	2,161.0	7	308.72	18.045	.001
OA	1,357.0	42	32.310	1.8886	.001
IA	353.83	6	58.971	3.447	.005
LOI	429.94	21	20.473	1.1967	–
LOA	5,452.8	126	43.276	2.5296	.001
OIA	308.63	42	7.3483	0.42952	–
LIA	781.44	18	43.413	2.5376	.001
LOIA	2,427.7	126	19.268	1.1262	–

L = light level
O = desk orientation
I = ink pigment density
A = subject

Table A2 (Continued)

3. Dependent Variable = Number of MISSES

EFFECT	SS	DF	MS	F	P
L	3.6490	3	1.2163	2.5296	–
O	4.9682	7	0.70974	1.476	–
I	15.563	1	15.563	32.366	.001
A	49.671	6	8.2785	17.217	.001
LO	8.9358	21	0.42552	0.88494	–
LI	2.3008	3	0.76693	1.5950	–
LA	13.910	18	0.77276	1.6071	–
OI	7.8878	7	1.1268	2.3435	.025
OA	14.481	42	0.34479	0.71705	–
IA	7.6596	6	1.2766	2.6549	.025
LOI	9.2662	21	0.44125	0.91766	–
LOA	53.224	126	0.42242	0.87849	–
OIA	14.046	42	0.33442	0.69550	–
LIA	7.6172	18	0.42318	0.88008	–
LOIA	48.035	126	0.38123	0.79283	–

4. Dependent Variable = Number of FALSE POSITIVES

EFFECT	SS	DF	MS	F	P
L	0.1808	3	0.060268	1.78902	–
O	1.6585	7	0.23693	6.9984	.001
I	1.2857	1	1.2857	37.978	.001
A	1.5056	6	0.25093	7.4121	.001
LO	0.81920	21	0.039009	1.1523	–
LI	0.16518	3	0.05506	1.6264	–
LA	0.43638	18	0.024244	0.71612	–
OI	1.3571	7	0.19388	5.7268	.001
OA	3.0212	42	0.071933	2.1248	.001
IA	0.8471	6	0.14118	4.1703	.001
LOI	1.1562	21	0.05506	1.6264	.05
LOA	4.0011	126	0.031755	0.93799	–
OIA	1.6975	42	0.040418	1.1939	–
LIA	0.63951	18	0.035528	1.0495	–
LOIA	3.8516	126	0.030568	0.90293	–

L = light level
O = desk orientation
I = ink pigment density
A = subject

Appendix B

A. VP_{max}

The assumption that VP_{max} increases with background (adaptation) luminance has support from the data by Valeton and van Norren (1983). In their study, a microelectrode was placed adjacent to a single photoreceptor in the retina of a monkey. Changes in electrical potential were recorded for changes in light level about a given adaptation luminance. Adaptation luminances ranged from darkness to 7 log photopic trolands. A single function, based upon Equation 1, described the relative cone response data for both increments and decrements about any given adaptation luminance.

Figure B1 shows the single-cone response function that was shifted vertically (arithmetic ordinate) and horizontally (logarithmic abscissa) to fit their data at different adaptation luminances. As adaptation luminance increases the curve shifts systematically toward the right along the abscissa and downward along the ordinate. Thus, the absolute levels of the maximum and minimum responses change relative to the zero response crossings (there is zero response at the adaptation luminances). With increases in the adaptation luminance, the absolute value of the maximum response, for increments, decreases while the absolute value of the minimum response, for decrements, increases.

In the present experiment only luminous decrements were used. Thus, one expects the absolute value of the decrement response to increase with background luminance, based upon the data from Valeton and van Norren. This is consistent with data in **Figure 3** and with the interpolation routine employed to estimate the values of VP_{max} (**Figure 5**).

B. *n and k*

Table 3 shows a general tendancy for n and k to increase with background luminance for the data in this experiment. For the model it was assumed that estimates of n and k at background luminances between 12 and 169 cd m^{-2} could be estimated from monotonically increasing interpolations (**Figures 6** and **7**). The assumption that k increases with adaptation luminance is clearly correct since its magnitude is heavily dominated by the background luminance (L_B) over the range of the model. The parameter k_0, however, may be defined as the abscissa value, relative to the origin, producing half of maximum response, and it can be estimated from k in **Table 3**. Dividing each estimate of k by the corresponding background luminance gives values of k_0. It is also true that k_0 tends to increase with background luminance, further supporting the assumption that k increases with background luminance.

The assumption that n increases with adaptation luminance is somewhat more tenuous, however. Although there was a tendency for n to increase monotonically with background luminance in this experiment, the true values of n might not increase with light level. Given this uncertainty in the true values of n, a sensitivity analysis was conducted to see how predictions of relative visual performance (RVP in Appendix C, Equation C10) would change if n were constant a) at the lowest estimate of n in **Table 3** (n = 1.0), b) at the highest estimate (n = 1.7), and c) at the average estimate (n = 1.4). Variations in the estimates RVP never exceeded 10 percent and rarely exceeded 4 percent over the entire range of the model. Thus, reasonable variations in the parameter n do not seriously affect the model predictions.

Appendix C

The procedure for calculating relative visual performance (RVP) in **Figure 11** is given below. Strictly speaking, the model is applicable to young adults reading positive contrast (dark ink on bright background) letters or numbers, each character subtending a visual angle of 0.2 degrees in width. The range of luminances in the model is between 12 and 169 cd m^{-2}. The reciprocal of time necessary to read the alphanumeric stimuli "as quickly and accurately as possible" is the measure of visual performance.

1. Select task background luminance, L_B, within the range of

$$12 \leq L_B \leq 169 \text{ cd m}^{-2}.$$

2. Calculate threshold contrast, C_t, at the selected L_B, where

$$C_t = 0.0418\,[\,(0.308/L_B)^{0.4} + 1.0)^{2.5} \tag{C1}$$

3. Select task luminance contrast, C_v, where L_B is greater than the target luminance, L_T, and

$$C_v = \frac{L_B - L_T}{L_B} \tag{C2}$$

4. Calculate the parameters n, k, and VP_{max}, where

$$n = 0.882 + 4.38 \times \Theta_1 - 6.05 \times \Theta_1^2 \tag{C3}$$

where $\Theta_1 = \log_{10}[\log_{10}(L_B)]$

$$\log_{10}k = -2.25 + 1.77 \times \Theta_2 - 0.217 \times \Theta_2^2 \tag{C4}$$

where $\Theta_2 = \log_{10}(L_B)$

$$VP_{max} = 0.0628 + 0.0120 \times \Theta_2 - 0.00268 \times \Theta_2^2 \tag{C5}$$

5. Calculate the predicted level of visual performance, VP, where,

$$VP = \{\Delta C^n/[\Delta C^n + (k/L_B)^n]\}\,VP_{max} \tag{C6}$$

where:

$$\Delta C = C_v - C_t$$

For values of $\Delta C < 0$, $VP = 0$

6. Visual performance relative to that at L_B = 169 cd m^{-2} and C_v = 1.0 (RVP) may be determined by the following equation:

$$RVP = VP/f \tag{C7}$$

where:

$$f = 0.0760 = VP_{max} \text{ at } L_B = 169 \text{ cd m}^{-2} \text{ and } C_v = 1.0$$

Toward a Model of Visual Performance: A Review of Methodologies

M. S. Rea

Introduction

The lighting practitioner directly influences the speed and accuracy of visual response (*i.e.*, visual performance) by affecting the illumination level and lighting geometry in a space. Yet visual performance is clearly but one aspect of performance at a realistic task; training, intelligence, motivation, and many other factors also play an important role in defining a person's level of task performance.[1,2] Special procedures and measurement techniques are required to minimize the impact of these "non-visual factors" on measured behaviour when trying to isolate visual performance from task performance. As argued in this paper, many studies that describe performance of realistic tasks have not incorporated such procedures. What is more, two or more visual factors may be confounded in some of these studies. Thus, the measured behaviour has been influenced by some unknown combination of visual parameters. Although these studies give a general flavor of how lighting and other stimulus parameters affect visual performance, it is usually impossible to delineate unambiguous relations between human visual responses and stimulus conditions. If the lighting practitioner is to follow recent, more refined lighting recommendations,[3] it will be necessary to have more refined information from research to implement the recommendations.

The present paper is a critical evaluation of literature from the area of illuminating engineering on the performance of human subjects at a variety of visual tasks. It is concluded that most of these experiments have procedural problems or have dealt with more than visual performance alone. Because visual performance cannot be isolated in the experiments, most of them cannot be used to set precise guidelines for relating lighting and visual performance, nor can they be used to validate or extend the visual performance model (**Figure 1**)* presented in the companion paper.

Experiments to investigate performance at several kinds of visual task are examined in the present paper. Each of the subsequent sections is devoted to one task. Common problems with procedures and measurements are identified and discussed in terms of the difficulties they create in specifying the relationship between lighting and visual performance. Finally, recommendations for future experimentation and measurement are proposed that can lead to precise lighting recommendations and facilitate assessment and advancement of the visual performance model proposed in the companion paper.[4]

Numerical verification task

Performance of the numerical verification task forms the basis of the model. Data using the same task have already been reported.[5-9] Briefly, subjects are required to find and then indicate discrepancies between juxtaposed printed number lists "as quickly and accurately as possible" on a given trial. The lists are examined once from top to bottom, and the time to complete the comparisons and note the number of errors of both omission and commission are recorded for each trial. In one set of numbers lists (response lists) all the sheets are printed in high contrast ink on white paper; in the other set of numbers lists (reference lists) the sheets are printed with different types of ink on white paper. The lists are usually examined under different levels of illumination and lighting geometry. Experimental combinations of all these factors produce different task background (adaptation) luminances and contrasts.

Dependent variable

A performance measure, score, based on total time and errors identified during each trial[5] was used as the dependent variable in all the numerical verification studies prior to this.[4]

$$\text{Score} = (20 - E)/(S + 5) \qquad (1)$$

where

E = number of errors per trial (both misses and false positives)
S = total time (s) to compare the lists.

This measure can affect the characterization of visual performance for a variety of reasons described in *Appendix A* of the earlier report by Rea.[4] There it was noted that score has a slightly nonlinear relation with the primary measures, misses, false positives, and

**Except where noted, all reported contrast values describe stimuli in which the target is darker than the background. Contrast (C) is defined as: $C = (L_B - L_T)/L_B$, where L_B is the luminance of the background and L_T is the luminance of the target.*

total time. Further, total time does not completely isolate the changes in visual response due to experimental manipulation because it is a composite of the times to 1) mark discrepancies, 2) read the lists of fixed contrast (*i.e.*, response lists), and 3) read lists of different contrasts (*i.e.*, reference lists).

The visual performance model was based upon time to read the reference lists while performing the numerical verification task. It is argued that this refined measure best characterizes visual performance for reading alphanumeric stimuli. Direct comparison of visual performance model predictions and score values from the numerical verification task can be made in **Figure 2**, which presents score data from *Appendix A* of the previous report and model predictions of relative visual performance (RVP) from equation (C7) of *Appendix C*[4]. The dashed lines through the score data are described by equation (A1), *Appendix A*. All RVP predictions were normalized to units of score using one common factor; in this case the factor was equal to the average of the four values of R_{max}, one for each background luminance, determined from equation (A1).

Although the model predictions are similar to the performance score values, it is clear from **Figure 2(a-d)** that the score data are flatter at medium and high contrasts than the model predictions. Further, the model predictions become less like the performance score values as illumination level (background luminance) is reduced. Thus, the characterization of visual performance depends not only on the experimental design and protocol but also on any transformations of the response measure. The dependent measure used in any study must be carefully scrutinized to determine whether it is an appropriate representation of visual performance. Ideally, the performance measure should be based simply on visual response times or errors.

Confounding subject groups with experimental variables

Despite the use of a common task, other factors specific to each of the studies using the numerical verification task (*e.g.*, subjects, viewing angles, size of numbers) can affect the observed levels of performance. Further, many undefined factors such as motivation, manual dexterity, or intelligence can also influence the results. The experimental conditions employed by Rea[6] are (not surprisingly) almost identical to those reported later.[4] In the earlier study a different group of young adult subjects with excellent uncorrected vision and positioned in a chin rest performed the numerical verification task at a fixed background luminance of 67 cd m^{-2}. **Figure 3** shows the mean scores from that study, plotted with the highest (169 cd m^{-2}) and lowest (12 cd m^{-2}) curves through the score data from **Figure 2**. The curve through the score data from Rea's earlier study[6] was obtained using equation (A1) in *Appendix A*.

Although the two sets of data are similar, the absolute levels of performance score were different in the two experiments. Subjects in the earlier study had higher scores, on average, than did subjects in the more recent report. Thus, even at a task run under essentially identical conditions and using subjects nominally equivalent, the absolute levels of performance for two populations of subjects differed for unknown reasons.

At present, model predictions of visual performance will have to be normalized to experimental data. Each set of data can differ in the absolute level of performance owing to differences in the experimental variables or, as shown, to unspecified differences amoung sampled subject populations. A different normalization factor will probably be required, therefore, whenever new subject populations are employed.

Applying the model predictions to different experiments having changed subject populations presents no difficulties, but utilizing various subject populations within one experiment can be very problematic. It is possible to compare model predictions with experimental data only roughly if variations in the stimulus parameters are confused with sub-populations of subjects within an experiment. For example, if one sub-population of subjects is given a task under one illumination level and another sub-population is given the same task under another illumination level,[10] it will be impossible to determine whether the observed differences in performance are due to differences in illumination, to differences in the subject groups, or to both factors.

In principle, very large sub-populations should have very small differences in their average performance potential. It is not always practical, however, to employ very large populations in one experiment. Even fairly large sub-populations (*e.g.*, n = 40; see section entitled **Tinker's reading experiments**) may be different in visual capabilities, motivation, manual dexterity, etc. As these factors influence measured performance, it is always better to reduce the likelihood of potential problems by ensuring that subject groups (sub-populations) are not confounded with experimental manipulations. Although confounding subject groups with stimulus variables in an experimental design has not been a problem in studies of the numerical verification task, it was a problem in several of the studies to be discussed later in this paper.

Photometric measurements

Clearly, errors in the photometric specification of the stimulus conditions can adversely affect comparisons between experimental data and model predictions. In studies by Rea,[4,6] Slater *et al.*,[8] and Slater and Perry,[9] so-called "calibration squares" were printed with the numbers on the reference and response sheets so that target luminances as well as background luminances could be measured to determine task contrasts. For a known lighting-task-subject geometry, then, it was possible to specify task contrast. Only in lighting conditions approaching those in a hemisphere do variations in geometry become inconsequential to task contrast. Directional lighting can greatly influence the perceived task contrast and, therefore, the expected level of visual performance.

In a recent study by Rea *et al.*[11] it became clear that subjects deliberately try to avoid contrast-reducing veiling reflections in order to maintain a high level of performance. In previous studies by Rea[4,6] and by Slater and Perry[9] subjects' head positions were fixed in a chin rest so that the lighting-task-subject geometry and therefore the task contrast were known. Smith and Rea[5] seated subjects in front of a white viewing booth roughly simulating hemispherical lighting conditions. Head and body positions were not restrained in any way, so that variations in head position did not seriously affect task contrast because the lighting was non-directional.** But although head position could not seriously affect task contrast, head and body movements could affect apparent size of the stimuli. As Rea *et al.*[11] also point out, subjects tend to move closer to the task as illumination levels are reduced, increasing the apparent size of the stimuli and thereby permitting better performance. It is therefore impossible to use the Smith-Rea data to isolate completely the effects of illumination level from apparent task size in assessing visual performance. This type of confounding is also present in several of the studies discussed below.

Conclusions

The numerical verification task can be useful in characterizing visual performance, but special precautions with regard to experimental design and execution must be implemented. In particular, task contrast, luminance, and size must be carefully controlled. An appropriate measure of performance, free of distortion, must also be defined before the data can be used to validate and advance a visual performance model.

Tinker's reading experiments

Early reading performance tests were conducted by Tinker using different task contrasts and various levels of illumination.[12-15] In his 1959 experiment, for example, several groups of subjects performed speed-reading tests.[15] Of special interest were two groups of university students of unknown age or ophthalmic capability; one group of forty saw printed passages with dark ink on white paper under six illumination levels, and another group of forty saw passages printed with dark ink on grey paper under the same six illumination levels.

Confounding with task contrast

As pointed out by Poulton,[10] changing the reflectance of the task background affects not only task contrast but adaptation luminance as well. Performance is known to vary with both parameters, and when they are confounded, as in this experiment, differences in performance cannot be attributed to one factor alone. Nevertheless, the contrasts of the two printed targets employed by Tinker were both quite high (according to the reported reflectances). The compressive visual response to contrast should produce nearly equivalent levels of performance for high contrasts at equal background luminances. Assuming that all other factors are constant, performance differences should therefore be mainly attributable to changes in background reflectance.

A serious confounding in Tinker's experiment (identified in the previous section) was between subject groups and background reflectances. Differences in the visual capabilities of the sub-populations or in their psychological profiles (*e.g.*, intelligence or motivation) could have contributed to the observed differences in reading performance at the different task background reflectances. In fact, there were reliable differences between the two sub-populations at matched task background luminances, indicating that even for relatively large sub- populations (n = 40), absolute differences in group performance potential can affect the results.

Photometric measurements

For reasons outlined in the previous section, there is also uncertainty about the accuracy of the reported photometric values that define the visual conditions experienced by the two groups of subjects. First, the reported values were reflectance measurements of paper and ink under a lighting geometry different from the lighting conditions actually experienced by subjects during the experiment. Further, documentation was not provided as to how the reflectance measurements were obtained from the very small letters in the printed text. Unless luminance measurements are obtained from some "calibration square" printed with the text it is unlikely that the reported values accurately represent ink density. Second, although both groups of subjects saw the reading passages under six

It should be noted that a calibration square was not printed on the task sheets used by Smith and Rea.[5] The contrast of the task materials had to be estimated by a comparison method. The luminance of the printed digits was estimated by measurements of a larger piece of paper matched in brightness to the printed digits. Although the task contrasts were estimated under hemispherical lighting and did not vary when subjects adjusted their postures, there may have been some discrepancies between the contrasts actually seen by the subjects and the estimated contrasts that were reported in the study.

levels of illumination ranging from 54 to 4300 lx, it is not clear that lighting geometry and thus task contrast did not change with changes in illumination level. Finally, it is possible that subjects systematically adjusted their posture under the different lighting geometries to improve task contrast and thus maintain a high level of performance. In truth, the reported photometric values may only be crudely representative of those actually seen by the subjects.

Although there is a consistent performance difference across illumination level for the two subject groups (confounded, of course, by the two background reflectances), the levels of performance across luminance are approximately constant. Even the relative performance of the two subject groups is questionable, however. As discussed in the preceding section, reading performance can probably be maintained if a subject systematically moves closer to the task as illumination is reduced. It seems very likely that the subjects in Tinker's experiment were unrestrained and behaved in this manner. As in the Smith-Rea study,[5] it is difficult to use these data to advance a visual performance model since illumination level and apparent size of text were probably confounded.

Dependent variable

The measure of performance for this experiment was the time to read a passage. At least two problems with the dependent measure make it of questionable utility for characterizing visual performance. First, the levels of comprehension were assumed to be high but were never verified, although Tinker reported that previous tests gave high levels of comprehension near 100 percent. In principle, the single dependent measure of reading speed will not reflect any possible differences in "accuracy" of reading (*i.e.*, comprehension). Without some direct verification of a high correlation between time and errors (*e.g.*, Rea[6]), reading speed may be a distorted representation of visual performance. Second, the very high levels of comprehension imply a very easy test for university students. Depending upon the kind of questions, it is conceivable that university students could answer many of the questions correctly without even reading the passages. In essence, Tinker's comprehension measure may be completely insensitive to visual performance. A performance measure should be realistic, but it must also be sensitive to the parameters manipulated in the experiment.

Conclusions

Although Tinker's experiment may give a general indication of visual performance at a reading task, a variety of serious procedural problems make it impossible to use these data, or those from his earlier studies, to evaluate the present, or any other, model of visual performance quantitatively. Other, more carefully controlled studies of reading performance would have to be used to validate and advance the visual performance model. Special care must be taken in this kind of experiment, however. Reading for comprehension may actually require more thinking than seeing,[16] and precautions must therefore be taken to ensure that the recorded responses reflect visual performance alone.

Experiments using visual search

Many investigators have employed search tasks.[10,17-20] The most widely known was first employed by Weston.[21,22] In this kind of experiment subjects are asked to search an array of Landolt rings for those having a specific gap orientation. The arrays can be printed in different sizes and inks and on different background reflectances. In an attempt to explore parametrically some of the factors important to visual performance, subjects are presented with the arrays under different levels of illumination. As argued above, the results obtained in such experiments depend not only on the visual factors employed in the experiment but also on factors such as the accuracy of the photometric measurements, experimental design and protocol, subjects' behavior, and scoring procedures.

Weston's early Landolt ring experiments

Documentation Although Weston's Landolt ring search tasks[21,22] are commonly-cited studies of visual performance, they were not adequately documented. Lighting geometry and task specularity (both target and background) interacted to affect luminance and contrast. Nowhere does Weston document how ink and paper reflectances were measured or under what conditions. Consequently, very little is known about the stimulus conditions actually presented to his subjects. Further, in his 1945 studies[22] both paper and ink reflectances were varied to change task contrast, so that it is imposible to attribute the measured effects to changes in contrast, background luminance, or both. Weston did not provide details about the positions assumed by subjects during testing. Presumably they were unrestrained and could adjust their posture and position in performing the task to suit themselves. Again, one must assume that illuminance and apparent size are confounded in the study. The visual capabilities of the subjects are unknown. Neither is it clear what experimental design was employed. One does not know, for example, whether all subjects saw all conditions. Weston's early landmark efforts in investigating visual performance must be lauded, but it would be impossible to have confidence in more than the general trends indicated by these experiments because of inadequate documentation.

Dependent measure The performance scoring procedure is one of the most serious problems with

Weston's experiments. He was the first to attempt to combine speed and accuracy in a performance metric, and many others have followed his example; several are discussed in this paper. Accuracy in Weston's experiment was defined as the number of hits (*i.e.*, targets found and marked) relative to the total number of targets in the array.

A proper measure of accuracy should also consider the number of correct rejections (*i.e.*, non-targets found identified and not marked). Further, without knowing both hits and correct rejections it is impossible to tell how many stimuli were actually inspected during the recorded time (one minute or shorter) and therefore impossible to make a proper estimate of speed (stimuli inspected per unit time). Consequently, his performance score is completely inappropriate as a measure of visual speed and accuracy. More will be said about the importance of determining the number of stimuli actually inspected per unit time in the **Boynton and Boss** section of this paper.

Consistency and reasonableness of results The two sets of performance curves for 3.0-min gap sizes reported by Weston in 194522 are shown in **Figure 4**. Not only are the two sets inconsistent in their prediction of performance but some of the smoothed functions are counter to known trends of visibility versus luminance. For example, the slopes of some of the performance versus luminance functions for low contrast stimuli are flatter than those for high contrast stimuli. In effect, this indicates that light level is less important to low-contrast material than to high-contrast material, a prediction that can be easily dispelled by a simple demonstration. If one looks through rotating, crossed-polaroid filters at high and low contrast stimuli on a luminous background, the low-contrast stimulus disappears much more quickly than the high-contrast stimulus as background luminance is reduced.

Conclusions One is left to conclude that Weston's studies cannot be used as a basis of validation or extension of a visual performance model owing to inadequate procedural details, inappropriate scoring procedures, and inconsistent results.

Modern experiments

Boyce Boyce[20] employed Weston's Landolt ring search task with some procedural differences. Boyce used total, self- paced scanning time as a measure of visual performance. Unlike Weston, who identified the importance of eliminating response times from estimates of visual performance, Boyce ignored response (action) time. Errors (of both omission and commission) were also used to index performance.

Although total time and errors changed with variations in illumination, these dependent variables were not compared directly to discover whether they gave similar or dissimilar representations of visual performance. Actually, the data were probably too variable to permit proper comparison. As it is not known whether errors and total time are alike, it is uncertain whether either response measure actually characterizes visual performance.

Four sets of Landolt rings were presented to all subjects, one set for each combination of two gap sizes (1.5 and 2.4 min) and two contrasts (C = 0.4 and C = 0.7). Actually, it is not clear from the published report[20] whether the reported task contrasts were obtained under the actual experimental conditions. In a personal communication, Boyce has confirmed that the photometric measurments were obtained under the conditions actually experienced by the subjects. The different sets of Landolt rings were presented under five different levels of illumination. Subjects' viewing angles and distances were controlled, avoiding systematic confounding of illumination level and apparent size, but a different set of subjects was used for each of the five light levels. Because subject groups are confounded with illumination levels, comparisons between performance values and the predictions from a visual performance model would not necessarily coincide.

Waters and Loe In the experiment by Waters and Loe,[19] arrays of high-contrast (C = 0.9) Landolt rings (gap size = 1.5 min) were presented to subjects under different light levels and different light sources. Every subject saw all combinations of light level and light source in a Latin square experimental design.

It is not clear how the contrast value reported for the Landolt rings was obtained, but measurements were made "in the position they were placed for the tests (p. 7)."[19] For very high contrast targets like those used in the experiment, however, it is not important to determine the contrast values accurately owing to the predicted saturation in visual response at medium to high contrasts. Despite uncertainties in the measurement technique, therefore, predictions of visual performance at a high contrast task should not be grossly inaccurate. Unfortunately, subjects were not adequately restrained while performing this task. As illumination levels were reduced, they could move closer to the task to maintain a high level of performance.

Performance was defined as the sum of the hits and correct rejections divided by the "net inspection time." Net inspection time was taken to be the total time to scan the array, row by row, minus the "action time" to mark targets in the array. There was no justification for combining net inspection time, hits, and correct rejections in this way; the measure of performance must therefore be suspect. Nevertheless, this combination seems to be appropriate since both hits and correct rejections as well as action time are taken into account.

Conclusions Modern experiments using the classic Landolt ring search task have used methodologies that are improvements on those used earlier by Weston.[21,22] Largely typified by the two discussed above, these experiments incorporate techniques that do not completely overcome possible confoundings. Boyce[20] controlled viewing angles and distances but used different subject groups for different illumination levels. Waters and Loe[19] used the same subjects throughout, but did not constrain head position or posture. Beyond these problems, and for reasons to be outlined, even the best controlled search tasks may be inappropriate for estimating visual performance.

Boynton and Boss

Boynton and Boss[18] conducted an unusual and well documented study that did not incorporate the Landolt ring matrix. An array of sixteen, randomly located stimuli was presented to young adults with normal visual acuity under different conditions of stimulus contrast (luminous decrements) and background luminance. Viewing distance was held constant while subjects searched for a small square (9 min on a side) among 15 circles (10 min in diameter). For half the trials 16 circles were presented (*i.e.*, no square was presented). The time between onset of the stimulus array and a button press by the subject, indicating that a square was or was not included in the stimulus array, was measured for every trial. A maximum of 20 s was allowed for a trial. **Figure 5** shows the results of the study by Boynton and Boss[18] characterized by the two measures of performance. The cumulative percentage of trials in which subjects correctly reported no square (percentage of correct responses) and the cumulative percentage of trials in which subjects correctly identified squares (percentage of targets reported) are plotted as a function of search time, τ, in seconds. For a stimulus array of 100 percent contrast, the search times required to reach the maximum level of performance regularly decrease as background luminance increases (**Figure 5a, c**). Similarly, for a background luminance of 1370 cd m^{-2}, the search times to reach maximum levels of performance regularly decrease as the stimulus array contrast increases (**Figure 5b** and **d**). Qualitatively, these findings are like those predicted in the visual performance model; visual performance should increase regularly with both background luminance and target contrast. Quantitatively, however, it is impossible to compare the visual performance model and the search data because the experimental paradigm used by Boynton and Boss, and possibly in all search experiments, probably confounds the number of targets searched with the difficulty of the visual conditions.

Response time and number of stimuli searched Response times can be affected by the visual processing time for each target (*i.e.*, visual performance) or, obviously, by the number of targets actually processed; the fewer to be searched the shorter the response times. The main problem with the Boynton and Boss procedure is that one does not know how many stimuli were actually processed. Their recorded times are based upon an unknown mixture of visual processing time and searching time. To isolate visual performance in these experiments one would have to estimate, and then eliminate, search time, which is comprised of time for each eye movement and total number of eye movements. It is probably not possible to use a simple constant to estimate search time; it is logical that subjects will systematically look at fewer stimuli under better visual conditions (high luminance and high contrast). By definition, targets become more visible as the stimulus conditions improve. Under these more conspicuous stimulus conditions, then, subjects would make fewer eye movements to find the target. The manner of visual search changes, therefore, with the level of visibility, so that the number of targets actually searched tends to be systematically fewer as visual conditions improve. As a consequence, searching time is probably confounded with the difficulty of the stimulus conditions. Without accurate eye movement data, yielding both time of each eye movement and how many were made, it would be impossible to estimate searching times or to evaluate the data in terms of a visual performance model.

Conclusions In the Landolt ring search task the stimuli are usually arranged in a matrix. One might suppose that subjects would scan the stimuli in a regular manner (*e.g.*, top left to bottom right), looking at every target once. This may not always be true. Some rows or rings within rows may be missed inadvertently. Similarly, some rows or rings may be scanned repeatedly. If these are random occurences, it is only necessary to collect more data to determine the relative changes in visual performance with variations in contrast, luminance, size, etc. It seems possible, particularly with the configuration of stimuli in the Boynton and Boss experiment, that there may be systematic changes in the number of stimuli scanned, depending upon the quality of the stimulus conditions. Because the number of targets scanned may be confounded with the difficulty of the visual conditions (even in the Landolt ring search studies), precautions must be taken to determine the number of targets actually searched as well as the time to perform the task. Otherwise visual performance cannot be evaluated.

Contrast threshold task

Several investigators[23-25] have used contrast threshold tasks to identify the lower limits of visual performance. Usually, luminous increments are presented to subjects on various luminous backgrounds. The probabilities of detecting luminous increments (contrasts) of different magnitude on a given luminous background are obtained only after many trials. A sigmoid function generally describes the relation be-

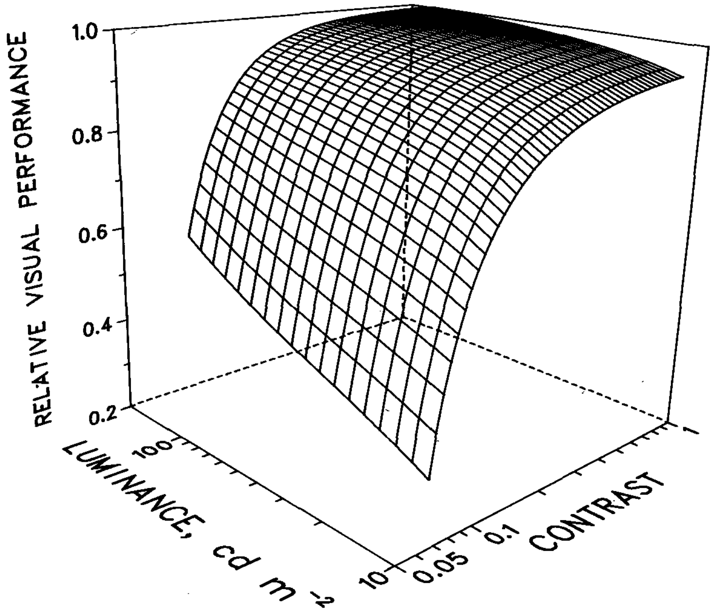

Figure 1—Three-dimensional representation of visual performance model for alphanumeric reading material.[4] Relative visual performance, scaled linearly, is plotted as a function of task contrast and background luminance, both scaled logarithmically

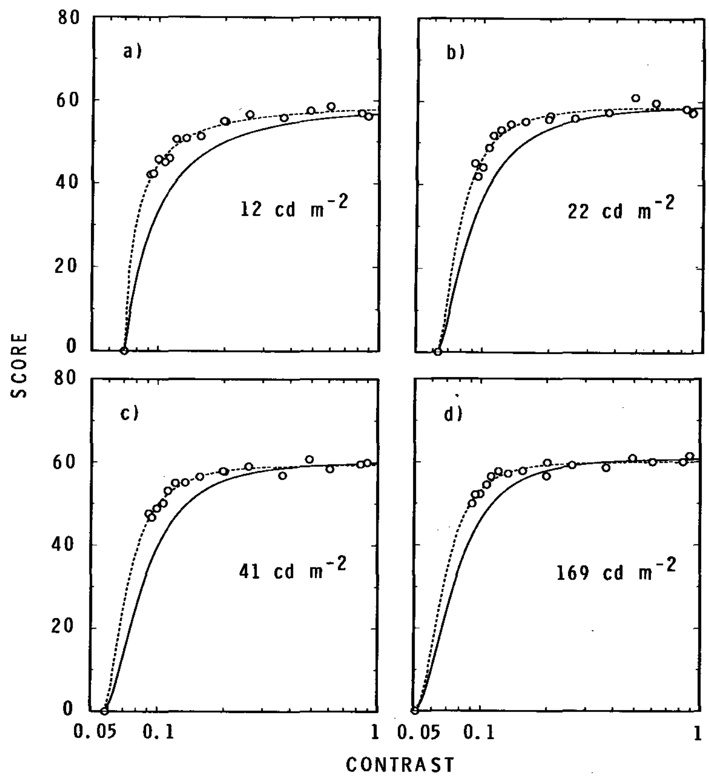

Figure 2—Comparison of score values for the numerical verification task and predictions of visual performance model.[4] The four curves through the score data were obtained from equation (A1). Model predictions (solid lines) are normalized using one factor equal to 60.66

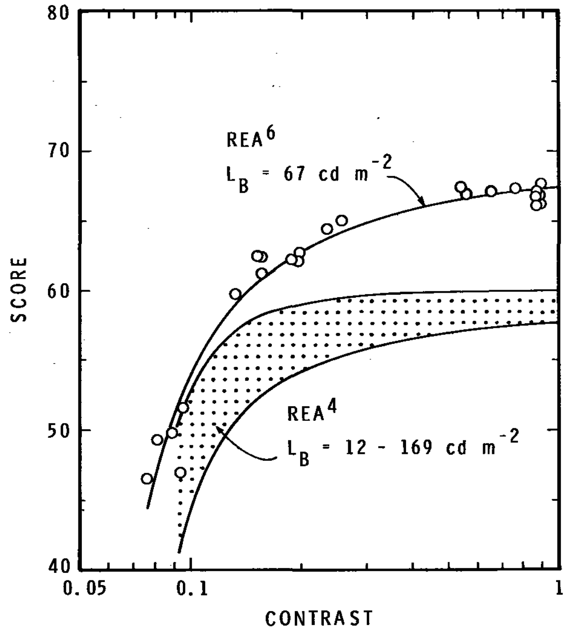

Figure 3—Comparison of score values at the numerical verification task.[4,6] The shaded area is bounded by the 12 and 169 cd m^{-2} score curves from Figure 2. The solid curve through the score values defined by Rea[6] is obtained from equation (A1)

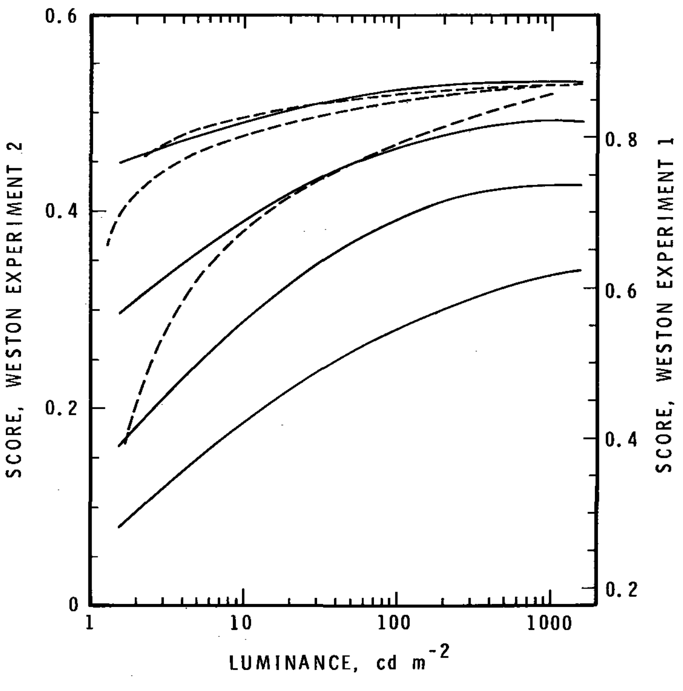

Figure 4—Comparison of visual search performance in two experiments conducted by Weston.[22] Both sets of curves represent performance scores for subjects searching for Landolt rings with a specific (3-min) gap orientation. The two sets of data were scaled differently in an attempt to match the trends in performance score at high and medium contrasts. Dashed lines are the performance curves at three task contrasts (top to bottom, 0.92, 0.68 and 0.36) smoothed by Weston for his first experiment; score values may be obtained from the right ordinate. Solid lines are the performance curves smoothed by Weston at four task contrasts (top to bottom, 0.97, 0.56, 0.39 and 0.28) for his second experiment; score values may be obtained from the left ordinate. (Reprinted with permission HMSO, London)

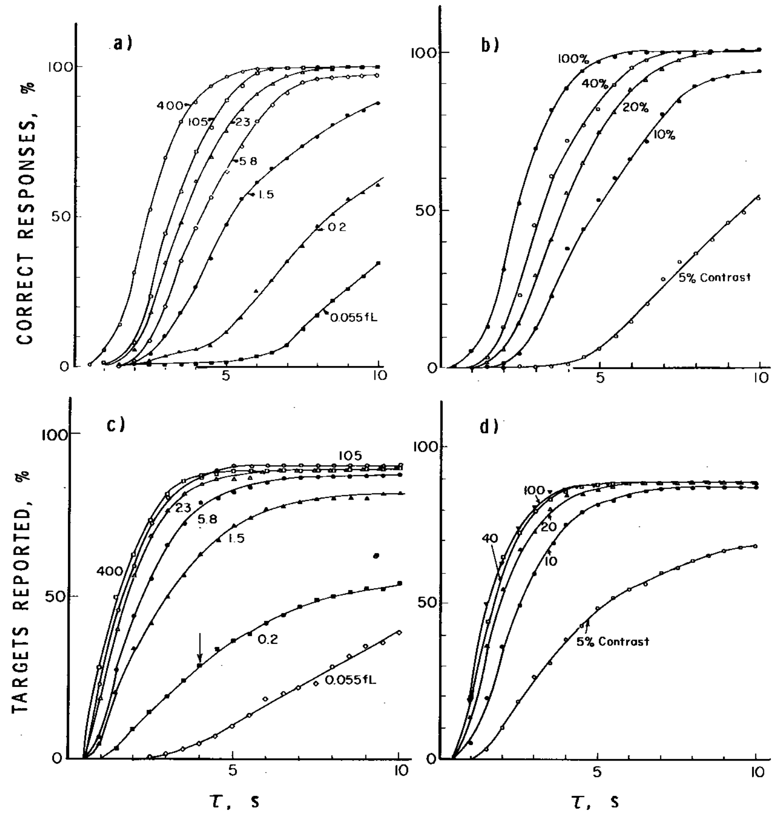

Figure 5—Visual search data, from Boynton and Boss.[18]

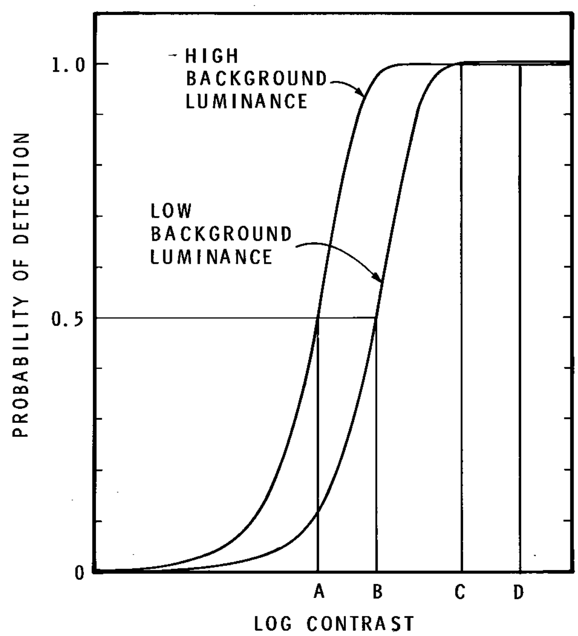

Figure 6—Probability of detection curves for luminous increments on high and low background luminances. The probability of detection equal to 0.5 is taken as the constant criterion giving visual threshold for any background luminance. Abscissa values for points A and B are contrast thresholds at the two background luminances, defined as equally visible in the CIE model.[31] Abscissa values for points C and D are contrast values four times those at A and B, respectively, and at their respective background luminances are defined as equally visible in the CIE model. Contrast is defined as the difference between the luminance of the target and that of the background divided by the luminance of the background

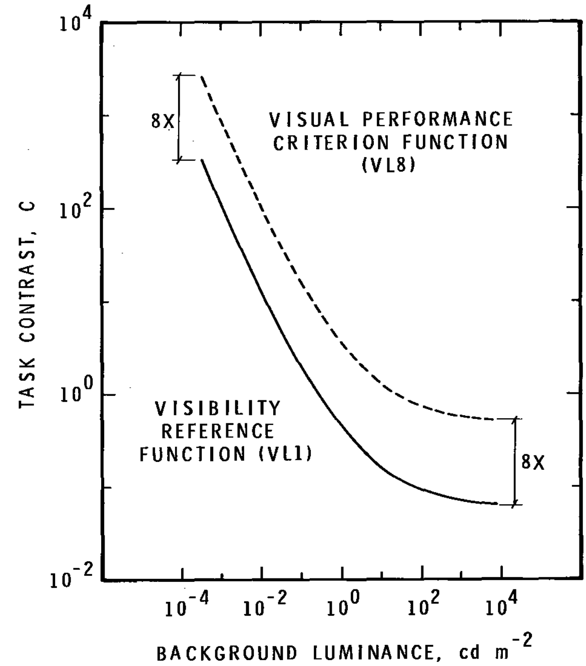

Figure 7—Visibility Reference Function (VL1) and Visual Performance Criterion Function (VL8) for the CIE visual performance model (see Figure 3-24, IES[32]). VL1 represents the threshold contrast for a circular, 4-min diameter, luminous increment on various background luminances. Contrast in this figure is defined as the difference between the luminance of the target and that of the background divided by the luminance of the background. VL8 is an assumed equal-visibility curve for contrasts eight times those for VL1

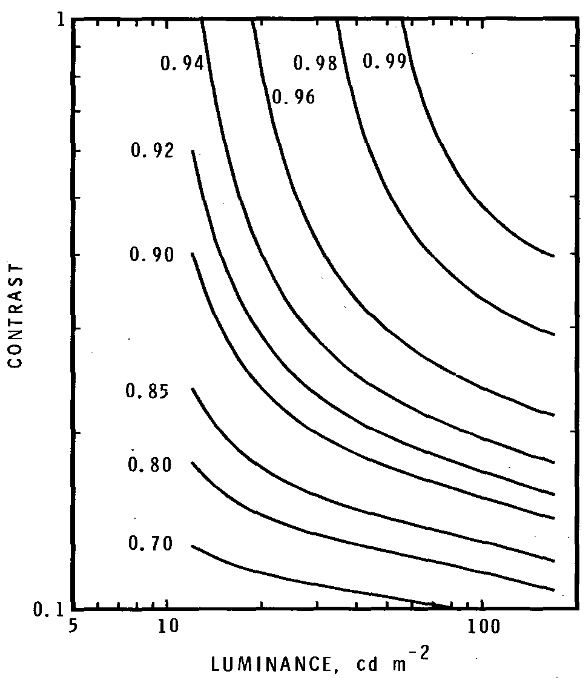

Figure 8—Lines of constant criterion, visual performance (see Figure 11, Rea[4]). Curves are labelled in units of relative visual performance (RVP)

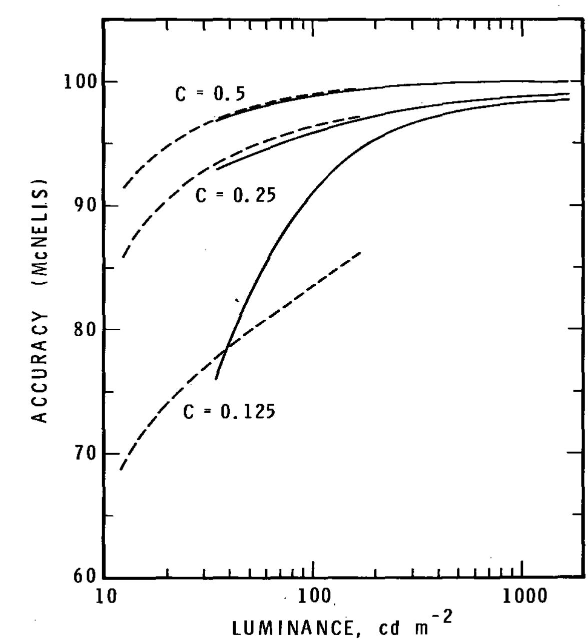

Figure 9—Comparison of performance at a numerical recitation task (from McNelis[1]), solid lines, and predictions from the visual performance model (from Rea[4]), dashed lines. Model predictions not normalized.

tween the logarithm of the luminous increment on a fixed background and the probability of detecting the increment, ranging from 0 to 1 (**Figure 6**).

From the probability of detection curves a visual threshold can be determined. This is defined as a constant, psychophysical criterion producing a fixed level of sensation, usually taken as the break-point between perceptible and imperceptible. For a detection task at one background luminance, contrast threshold is represented by a fixed probability of detection, p; usually p = 0.5. Once the criterion has been established it is possible to derive a contrast threshold function that defines contrast threshold at various background luminances. That is, the contrast required to reach threshold, obtained from each probability of detection curve, is plotted against the background luminance for which the probability of detection data were obtained (*e.g.*, VL1 in **Figure 7**).

Visibility level

A model based largely upon detection threshold studies by Blackwell and his co-workers[24-28] has been developed to characterize suprathreshold as well as threshold visual performance.[30,31] In this CIE model suprathreshold levels of visual performance are represented by so-called Visibility Levels (VL) derived from the contrast threshold values. The contrast threshold curve for detecting a small disc, called the Visibility Reference Function, VL1, has been taken to represent the basic relation between background luminance and threshold contrast.[32] Higher criterion levels representing constant suprathreshold levels of visual performance are obtained by simply multiplying the Visibility Reference Function by a fixed factor (see caption, **Figure 6**). VL8 in **Figure 7**, for example, is the equal visibility curve relating background luminance to contrasts eight times those at VL1.

Probabilities of detection equal to unity In the CIE model,[30,31] two parameters, M and s, of a cumulative normal distribution describe the sigmoids (or so-called normal ogives) fitting the probability of detection data at each background luminance (*e.g.*,**Figure 6**). M is the mean of the normal distribution of contrast values (corresponding to the 0.5 probability of detection) and s is its standard deviation. The latter parameter characterizes the steepness of the sigmoid. Blackwell reports (see Technical Note 16[31]) that the ratio s/M is a constant equal to 0.37 for all background luminances.

The probability of detection curves for all background luminances eventually become equal to unity at high contrast. It is quite unlikely, however, that the high contrast targets will, in fact, produce the same level of visual performance at different background luminances. Reaction times to equally high contrast stimuli are different at different background luminances.[33] Similarly, the levels of performance at the numerical verification task are systematically different for the same high-contrast targets at various background luminances.[4] These results are consistent with the wealth of literature on the Pulfrich pendulum phenomenon (*e.g.*, Lit[34]).*** It is therefore impossible to use this methodology directly to discriminate between levels of suprathreshold visual performance.

By the same token, the probability of detection methodology cannot be used to validate a central assumption in the CIE model, namely, that contrasts on the same VL function (*e.g.*, VL8 in **Figure 7**) will produce the same level of suprathreshold visual performance. When all of the probabilities of detection associated with high contrast targets are equal to unity, it is not only impossible to use those data to discriminate among levels of suprathreshold visual performance (discussed in the previous paragraph) but also impossible to validate the assumption that all the points on a single Visibility Level function give the same level of visual performance. Another technique, perhaps using reaction time, would have to be employed to test the assumption directly.

It seems unlikely, however, that this assumption is valid.[4] It has been shown that constant criterion curves from the visual performance model for luminous decrements are not simple multiples of the lowest (threshold) criterion function. Graphically, the constant criterion curves are not parallel in a space defined by log contrast (ordinate) versus log background luminance (abscissa) axes; the separation in contrast between constant criterion curves is greater at low luminances than at high (**Figure 8**). This increased separation at low luminances happens because the level of maximum visual performance associated with maximum contrast at a given background luminance increases systematically as background luminance increases. Since it was impossible to attain as high a level of visual performance at low background luminances as at high, even for maximum contrast (unity for luminous decrements), it is impossible for the constant criterion curves (*i.e.*, Visibility Levels) to be parallel throughout the log contrast versus log background luminance space, as supposed in the CIE model and illustrated for luminous increments in **Figure 7**.

Effects of guesswork or response bias on probability of detection In these experiments the probability of chance detection (guessing) is determined in an attempt to isolate the visual threshold from the non-sensory

******A high contrast pendulum swinging in one plane will appear to swing in an elliptical path when viewed binocularly if the two eyes are not adapted to the same light level. The latency of the signal from the eye adapted to the lower light level is lower in reaching binocular cells in the brain than is the signal from the eye adapted to the higher level. This difference in latency, dependent upon the difference in the adaptation luminances, presumably produces the illusion.***[35]

contributions to the subjects' responses. One of two options is generally employed to characterize the visual threshold. With the first option a criterion equal to chance detection is chosen as the probability of detection to represent visual threshold. If, for example, one is forced to decide whether a target is presented to the left or to the right of a fixation point, the subject has a 50 percent chance of guessing the correct location. Probabilities greater than chance (in this case $p = 0.5$) are taken as visual events. Thus, a 0.5 probability of detection criterion is taken as visual threshold, the break-point between perceptible and imperceptible. Response bias presents a difficulty with this approach. For example, if subjects are biased to say "right" all the time, the best score would be $p = 0.5$, no matter how high the contrast. This, of course, would never happen, but if subjects have some degree of response bias and it is unknown, all of the probability of detection data would reflect both visual response and psychological bias.

With the second option, the raw probability of detection (p) data are mathematically transformed, according to the following formula, into a new set of probability of detection (p′) data in an attempt to correct for guesswork.[29]

$$p' = (p - q)/r \tag{2}$$

where q is the probability of guessing the correct answer r is the probability of guessing the wrong answer and $r + q = 1$.

As with the first method, chance behavior is typically taken as an estimate of q. For the example given, both r and q are equal to 0.5. Thus, for $p = 0.5$, $p' = 0.0$. Values of p′ greater than 0.0 can be taken as greater than threshold. As with the first method, response bias can play a part in the transformed probabilities of detection. Blackwell and Scott,[29] for example, have shown that the values of q and r cannot always be characterized simply by chance detection; their transformed probability of detection data are still combinations of seeing and non-randon guessing. Steps had to be taken to transform these data further in order to eliminate response bias. Blackwell and Scott used an empirical, graphical correction procedure to adjust values of p′. It applies only to their data, but such efforts illustrate the difficulty of isolating visual response in probability of detection data.

In many experiments (not just detection tasks) response biases have been neither assessed accurately nor documented. In these cases the probability of detection data should be suspect. The choice of criterion probabilities to represent threshold may be arbitrary and unrepresentative of the visual response. Occasionally, special experimental techniques or methods of analysis (as in the Blackwell and Scott experiments) must be employed to eliminate response bias.

Conclusions The probability of detection methodology has been used to model suprathreshold visual performance, but there can be fundamental difficulties in the response measure that limit its utility. First, high contrast targets associated with probabilities of detection equal to unity (or the same Visibility Level value) at different background luminances do not always produce the same level of visual performance. Second, special care must be taken to ensure that the probability of detection curves are not contaminated by subject response biases.

A timed recitation task

McNelis[1] conducted an unusual study of visual performance when on each trial an observer was required to name two briefly-presented (0.4 s) small letters (8 min in size) separated by a 10-degree visual angle. The letters were presented at different contrasts (luminous decrements) and at different background luminances. Accuracy (correct identifications) was used as the dependent variable. Performance was better for the left letter than for the right letter. As confirmed by eye movement data, there was not enough time for subjects to identify the second (right) letter in 0.4 s except at high luminances and contrasts. McNelis' accuracy data for the left letter have been redrawn from data supplied by him and presented in **Figure 9.**

Methodological issues

The McNelis study was well documented, but it did not make entirely clear whether subjects' heads were restrained. He has confirmed in a personal communication that head positions were indeed restrained by a chin rest. Postural changes in response to reductions in luminance were prevented, so that apparent size and luminance were not confounded in the study. McNelis also notes in a personal communication that contrast measurements were difficult to make and that there may be some uncertainty in the values; such uncertainty makes predictions of performance, particularly for low contrast stimuli, more tenuous.

Professor S.W. Smith points out in discussion of the McNelis study that subjects did not operate at a chance level of accuracy; correct identifications approached zero under difficult conditions. Subjects were apparently biased in their responses against errors of commission (*i.e.*, false positives). Because the number of correct responses due to guesswork were apparently very small, the simple measure of accuracy as the dependent variable was probably (but not certainly) representative of visual performance.

Comparison with predictions from the visual performance model

Other than the potential problem of estimating absolute contrast, the methodologies employed in the McNelis study appear to avoid most of the problems outlined in this paper. For this reason it seems appropriate to compare predictions of the visual performance model developed by Rea in an earlier paper[4] with McNelis' accuracy data. Both the model predictions and McNelis' accuracy data are shown in **Figure 9**. The model predictions at the three contrasts were all normalized by use of a single factor of 100.

Agreement between McNelis' data and the model predictions are excellent at the two higher contrasts. Even at the lowest contrast they are reasonably good. Difficulties in both model predictions and the McNelis data are evident, however, in **Figure 9**. First, the model predictions at the lowest contrast are probably inaccurate to some degree; the slope of this lowest function does not continue to decrease with increases in background luminance, as one would expect. More refined estimates of visual performance at the lowest contrast values are required. Second, the contrast values reported by McNelis are, as he indicated, probably inaccurate. Examining his data for the highest luminance, it is clear that the difference in performance is greater for contrasts of 0.5 and 0.25 than for 0.25 and 0.125. It would be more reasonable to expect the difference in performance to be larger for the latter two contrasts than for the former pair. Inaccuracy in evaluating the contrast of task materials is, of course, problematic in evaluating the visual performance model. Again, care must be taken to specify contrasts of task materials accurately, particularly at the lowest contrasts, because visual performance changes most rapidly just above contrast threshold.

Conclusions

The study by McNelis[1] appears to avoid most of the procedural problems outlined in this paper. One would expect, then, good agreement between predictions from the visual performance model and the reported accuracy data. Indeed, there was marked similarity between model predictions and the data, despite obvious procedural differences between the numerical verification task and those used by McNelis. It appears that the McNelis procedures could easily be used to extend and validate the visual performance model.

Discussion and conclusions

In the past it may have been less important to obtain precise estimates of stimulus conditions or visual performance as distinct from overall task performance. Traditional recommendations of illumination levels for various tasks were made without specifying factors important to vision, such as task reflectance and contrast.[32] Only very general information about task performance was necessary to make lighting recommendations of this kind and, perhaps rightly, investigators did not strive for experimental precision greater than that required by the various national and international sanctioning bodies. Perhaps, too, practitioners have been concerned with lighting real tasks and many investigators deliberately studied overall performance at different kinds of simulated, realistic tasks. These experiments gave a general indication of how performance is affected by lighting, but for many of them it was not possible to isolate the contribution of each visual factor to performance nor to disentangle the visual factors from non-visual ones that contribute to performance. It was not possible, therefore, to determine how important lighting is to visual performance at a task.

There has been recent emphasis on more exacting specification of the conditions important for visual performance, as reflected in more refined recommendations of illuminance.[3] Size, contrast and the age of the worker are important factors that must be specified when determining levels of illumination. If it is necessary to follow the more precise recommendations of lighting quantity and quality for visual performance, then it is necessary to have more explicit knowledge of the relationships between luminous parameters and visual performance.

The model proposed by Rea[4] is an attempt to provide quantitative predictions of visual performance at an alphanumeric reading task from the photometric specification of the luminous stimuli. Based upon a variety of arguments, the proposed model seems to be an adequate, first-order characterization of visual performance at that kind of task. It is not necessarily far-reaching in its predictive capabilities (*e.g.*,, different text sizes and age groups are not included), nor has it been completely verified. For reasons outlined in the present paper, problems with experimental procedures, photometric measurements, and indices of performance limit the utility of many earlier studies for validating and extending the visual performance model. It is therefore necessary to design and control future experiments more carefully in order to assess and extend the model predictions. The recommendations for future experiments on visual performance identified in this paper are summarized in **Recommendations** below.

Although more experiments must be executed to validate and extend the visual performance model, the close correspondence of the model predictions with well controlled electrophysiological[36] and psychophysical[1] studies implies that it will be an important step in developing a general model of suprathreshold visual performance. When the necessary experiments have been completed and a final visual performance model is established, it will be possible to make more precise lighting recommendations, and, importantly, to justify them.

Recommendations

It goes without saying, perhaps, that experiments should be well designed, executed, and analyzed. Barber,[37] for example, identified important issues in conducting experiments and describing the results. Care should be taken to adhere to this credo in future experiments on visual performance. Common, recurring problems specific to past studies of visual performance have been identified in this paper and deserve special mention.

Independent variables

Accurate measurements of apparent luminance, contrast, and size must be obtained, and must represent the conditions actually experienced by the subjects.

Dependent variables

Indices of performance should not distort visual response. Simple measures of time and errors rather than arbitrary scoring indices should be used. The measures should be sensitive to changes in visual response across experimental conditions. Equal probabilities of detection at different background luminances, for example, do not necessarily result in equivalent levels of suprathreshold visual performance. The response measures should also limit the influence of non-visual factors.

Experimental design

Proper experimental designs must be used to eliminate various sources of confounding. Subject sub-populations should not be confounded with experimental variables. Absolute differences in subjects' response levels can mask differences in performance due to experimental variables. Systematic behavioral changes in response to experimental variables should be eliminated. For example, leaning forward as illuminance levels are reduced confounds the effects on vision of apparent size and background luminance. Similarly, in a search experiment the number of stimuli inspected may be confounded with background luminance.

Acknowledgements

The author gratefully acknowledges the continuing assistance of M. Ouellette during the preparation of this manuscript. Thanks are also extended to the following for their comments on an earlier manuscript: S. Berman, R. Biggs, P. Boyce, R. Clear, R. Corth, V. Crisp, D. DiLaura, A. Levy, I. Lewin, M. Perry, A. Robertson, P. Russell, A. Slater, S. Squillace, G. Webber. This paper is a contribution from the Institute for Research in Construction, National Research Council of Canada.

References

1. McNelis, J.F. 1973. Human performance—a pilot study, *J of the IES* 2(no. 3):190-196.

2. Rea, M.S. 1984. Overview of task performance in the illuminated environment. *Illuminating engineering learning aid.* Nat'l. Research Council of Canada, Division of Bldg. Research.

3. IES of N. America. 1981. *IES lighting Handbook* Ref. Vol. New York.

4. Rea, M.S. Toward a model of visual performance: foundations and data. *J of the IES* 15(no. 2).

5. Smith, S.W. and Rea, M.S. 1980. Relationships between office task performance and ratings of feelings and task evaluations under different light sources and levels. Proceedings 19th Session. CIE Publication No. 50:207-211. Kyoto.

6. Rea, M.S. 1981. Visual performance with realistic methods of changing contrast. *J of the IES* 10(no. 3):164-177.

7. Rea, M.S. 1983. Effects of Haidinger's brushes on visual performance. *J of the IES* 12(no. 3):197-203.

8. Slater, A.I., Perry, M.J., and Crisp, V.H.C. 1983. The applicability of the CIE visual performance model to lighting design. *Light and lighting.* Proceedings of the 20th session. CIE Publication No. 56. pp. D114/1-4. Amsterdam.

9. Slater, A.I., and Perry, M.J. 1984. Visual performance under reference and realistic conditions. *Proceedings of the CIBS Nat'l. Lighting Conference* Cambridge, UK.

10. Poulton, E.C. 1969. Skimming lists of food ingredients printed in different brightness contrasts. *J of Applied Psychology* 53(no. 6):498-500.

11. Rea, M.S., Ouellette, M.J., and Kennedy, M.E. 1986. Lighting and task parameters affecting posture, performance and subjective ratings. Nat'l. Research Council of Canada, Institute for Research in Construction.

12. Tinker, M.A. 1939. The effect of illumination intensities upon speed of perception and upon fatigue in reading. *J of Ed Psych* 30(no. 8):561-571.

13. Tinker, M.A. 1943. Illumination intensities for reading newspaper type. *J of Ed Psych* 34:247-250.

14. Tinker, M.A. 1952. The effect of intensity of illumination upon speed of reading six-point italic print. *Amer J of Psych* 65(no. 4):600-602.

15. Tinker, M.A. 1959. Brightness contrast, illumination intensity and visual efficiency. *Amer J of Optometry and Archives of American Academy of Optometry* 36(no. 5):221-236.

16. Rea, M.S. 1982. An overview of visual performance. *Lighting Design + Application* 12(No. 11)13.

17. Stone, P.T. and Groves, S.D. 1968. Discomfort glare and visual performance. *Transaction of the IES (London)* 33(no. 1):9-15.

18. Boynton, R.M., and Boss, D.E. 1971. The effect of background luminance and contrast upon visual search performance. *Illuminating Engineering* April 1971, pp. 173-186.

19. Waters, I. and Loe, D.L. 1973. Visual performance in illumination of differing spectral quality. Visual

Performance Research Proj. Univ. College Environ. Research Group.

20. Boyce, P.R. 1973. Age, illuminance, visual performance and preference. *Lighting Research and Technology* 5(no. 3):125-144.

21. Weston, H.C. 1935. The relation between illumination and industrial efficiency, 1. The effect of size of work. HMSO. pp. 1-14. London.

22. Weston, H.C. 1945. The relation between illumination and visual efficiency—The effect of brightness contrast. Industrial Health Research Bd. Great Britain Medical Research Council, Report No. 87:1-35 London.

23. Hecht, S., Peskin, J.C., and Patt, M. 1938. Intensity discrimination in the human eye. II. Relationship between ΔI/I and intensity for different parts of the spectrum. *J of Gen Physiol* 22:7-19.

24. Graham, C.H., and Kemp, E.H. 1938. Brightness discrimination as a function of the duration of the increment in intensity. *J of Gen Physiol* 21:635-650.

25. Blackwell, H.R. 1946. Contrast thresholds of the human eye. *JOSA* 36(no. 11):624-643.

26. Blackwell, H.R. 1953. Psychophysical thresholds: Experimental studies of methods of measurement. Engineering Research Inst. Univ. of Mich.

27. Blackwell, H.R. 1959. Development and use of a quantitative method for specification of interior illuminating levels on the basis of performance data. *Illuminating Engineering* 54:317-353.

28. Blackwell, O.M. and Blackwell, H.R. 1971. Visual performance data for 156 normal observers of varying age. *J of the IES* 1:3-13.

29. Blackwell, H.R., and Scott, D.E. 1973. Analysis of visual performance data obtained in a Landolt ring task without response limitation. *J of the IES* 2(no. 4):445-460.

30. CIE, Technical Committee. 1972. A unified framework of methods for evaluating visual performance aspects of lighting. CIE Publication No. 19.

31. CIE Technical Committee. 1981. An analytical model for describing the influence of lighting parameters upon visual performance. Vol. 1: Technical Foundations. CIE Publication No. 19/2.1.

32. IES of N. America. 1972. *IES Lighting Handbook* New York.

33. Vicars, W.M., and Lit, A. 1975. Reaction time to incremental and decremental target luminance changes at various photopic background levels. *Vision Research* 15:261-265.

34. Lit, A. 1949. The magnitude of the Pulfrich stereophenomenon as a function of binocular differences of intensity at various levels of illumination. *Amer J of Psych* 62(no. 2):159-181.

35. Levick, W.R., Cleland, B.G., and Coombs, J.S. 1972. On the apparent orbit of the Pulfrich phenomenon. *Vision Research* 12(no. 8):1381-1388.

36. Albrecht, D.G., and Hamilton, D.B. 1982. Striate cortex of monkey and cat:contrast response function. *J of Neurophysiol* 48(no. 1):217-237.

37. Barber, T.X. 1976. *Pitfalls in human research* New York: Pergamon Press Inc.

Appendix A

All the score data in **Figures 2** and **3** were fitted with the same general form of equation (1) developed by Rea.[4] This equation is repeated in equation (A1) below in a slightly different form. Parameter estimates for each set of data were obtained by a least squares criterion using a non-linear regression computer routine.

$$R/R_{max} = \Delta C^n/(\Delta C^n + k_0{}^n) \qquad \text{(A1)}$$

where

R = score in equation (1)
R_{max} = a free parameter equal to the maximum score for a given background luminance, L_B
n = a free parameter affecting the shape of the curve
k_0 = a free parameter, also affecting the shape of the curve, representing the increment in contrast above threshold contrast (C_t) producing half maximum response.

$$\Delta C = C_v - C_t \qquad \text{(A2)}$$

$$C_v = \text{task contrast} = (L_B - L_T/L_B) \qquad \text{(A3)}$$

where

L_B = task background (*e.g.*, paper) luminance
L_T = task target (*e.g.*, printed digit) luminance
$L_B \geq L_T$
C_t = C_v at contrast threshold