

# Merging Psychophysical and Psychometric Theory to Estimate Global Visual State Measures from Forced-Choices

Robert W. Massof,<sup>1</sup> Karen M. Schmidt,<sup>2</sup> Daniel M. Laby,<sup>3</sup> David Kirschen,<sup>4</sup> and David Meadows<sup>5</sup>

<sup>1</sup>Johns Hopkins University, <sup>2</sup>University of Virginia, <sup>3</sup>Harvard University, <sup>4</sup>University of California Los Angeles, <sup>5</sup>University of Florida

Email: bmassof@jhmi.edu

**Abstract.** Visual acuity, a forced-choice psychophysical measure of visual spatial resolution, is the *sine qua non* of clinical visual impairment testing in ophthalmology and optometry patients with visual system disorders ranging from refractive error to retinal, optic nerve, or central visual system pathology. Visual acuity measures are standardized against a norm, but it is well known that visual acuity depends on a variety of stimulus parameters, including contrast and exposure duration. This paper asks if it is possible to estimate a single global visual state measure from visual acuity measures as a function of stimulus parameters that can represent the patient's overall visual health state with a single variable. Psychophysical theory (at the sensory level) and psychometric theory (at the decision level) are merged to identify the conditions that must be satisfied to derive a global visual state measure from parameterised visual acuity measures. A global visual state measurement model is developed and tested with forced-choice visual acuity measures from 116 subjects with no visual impairments and 560 subjects with uncorrected refractive error. The results are in agreement with the expectations of the model.

## 1. Background

Visual resolution (a.k.a. visual acuity) often is measured by determining the smallest size Landolt ring for which the orientation of the gap in the ring can be identified reliably [1]. As shown in Figure 1, on each trial the Landolt ring is presented at one of four randomly chosen orientations and the observer is forced to respond with one of four categories: “up”, “down”, “left”, or “right”. The response is scored dichotomously as “correct” (1) or “incorrect” (0). Different size stimuli are presented multiple times and the frequency of correctly identifying the stimulus orientation is estimated as a function of stimulus size.

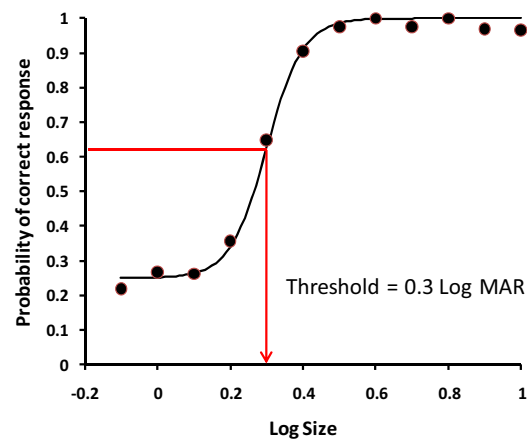


**Figure 1.** Examples of Landolt rings at different orientations. The observer must identify the orientation of the gap.

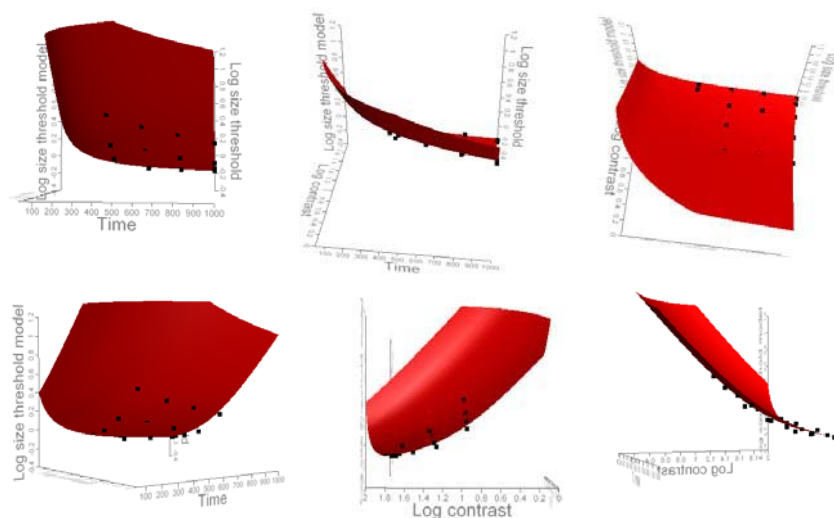
As illustrated with simulated data in Figure 2, by convention the visual resolution threshold is defined as the stimulus size that corresponds to the probability of correctly identifying the stimulus orientation that is halfway between 100% and chance performance (25% in this example). [2] The curve fit to the simulated data in Figure 2 is a cumulative logistic (approximation to a cumulative log normal distribution) with the lower bound constrained to chance, i.e.,  $P_c(S|\hat{S}) = 0.25 + \frac{.75 \exp(a[S-\hat{S}])}{1 + \exp(a[S-\hat{S}])}$ , where  $S$  is the log stimulus size and  $\hat{S}$  is the log stimulus size at threshold.

The visual resolution threshold depends on how far off axis the stimulus is from the line of sight, the average luminance of the stimulus, stimulus contrast, stimulus exposure duration, and other physical parameters of the stimulus. The visual resolution threshold also depends on the observer's light exposure history prior to the stimulus presentation and on the health state of the observer's visual system. Besides size thresholds, one can use the Landolt ring resolution task to measure contrast thresholds, exposure duration thresholds, etc. as a function of the other stimulus parameters.

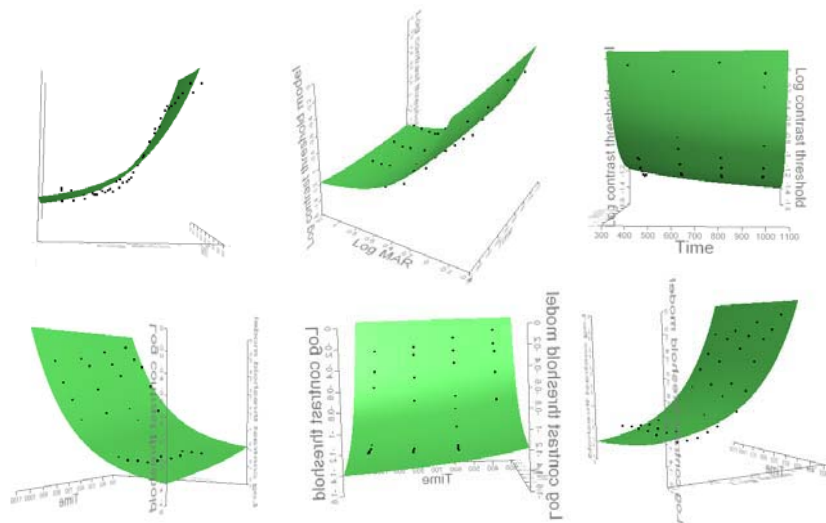
Figure 3 illustrates log size thresholds as a function of log contrast and exposure duration displayed at different orientations of the axes and Figure 4 similarly shows log contrast thresholds as a function of log size and exposure duration. These



**Figure 2.** Simulation of typical psychophysical visual acuity data plotted as probability of correctly identifying the orientation of the gap in the Landolt ring as a function of stimulus size, expressed as log minimum angle of resolution (MAR). The solid curve drawn through the data is a cumulative logistic with a lower bound at chance performance. Visual acuity, or the size threshold, is defined as the stimulus size that results in a probability correct half way between chance and 100%.



**Figure 3.** Log size threshold (ordinate), measured as a function of contrast (Log contrast) and stimulus exposure duration (Time). The black points are average thresholds for 116 normally-sighted subjects. The red surface is the fit of a vision psychophysical model based on stimulus reciprocities. The different panels illustrate six different orientations of the threshold surface.



**Figure 4.** Log contrast threshold (ordinate) for detecting the orientation of a Landolt ring was measured as a function of stimulus size (Log MAR) and exposure duration (Time). The black points are average thresholds for 116 normally-sighted subjects. The green surface is the fit of a vision psychophysical model based on stimulus reciprocities.

thresholds (points) are averages for 116 observers who have normal vision (points).

The threshold surfaces in Figures 3 and 4 are generated by a vision psychophysical model that is based on well known stimulus reciprocity laws (Ricco's law, Bloch's law, Piper's law, and Pieron's law). The log size threshold surface as a function of log contrast ( $C$ ) and exposure duration ( $T$ ) is defined as  $\hat{S}(C, T) = \ln\left(\frac{k_1}{T e^C} + k_2\right)$  [3-4] and the log contrast threshold surface as a function of log size ( $S$ ) and exposure duration is defined as  $\hat{C}(S, T) = k_3 \ln\left(\frac{k_4}{T}\right) + \frac{k_5}{e^S}$  [4-5].

The probability of correctly identifying the orientation of a stimulus with log size, log contrast, and exposure duration  $S, C, T$  is  $P_c(S, C, T) = P_c\left(S, C, T | \hat{S}(C, T)\right) = P_c\left(S, C, T | \hat{C}(S, T)\right)$ . This truism gives rise to the possibility that the parameterized size and contrast thresholds represent a single latent trait of the person,  $\theta$ , relative to a function of  $S, C$ , and  $T$ , i.e.,  $P_c(S, C, T | \theta, I(S, C, T))$ , where  $I(S, C, T)$  is a constructed latent variable that represents processing of the multi-parameter stimulus by the average visual system for a defined population. This concept could be extended to an arbitrary number of stimulus parameters, i.e.,  $P_c(S, C, T, \dots | \theta, I(S, C, T, \dots))$ . Thinking of resolution thresholds in this way opens the door to the possibility of measuring the global state of the visual system, relative to the average visual system, with a single parameter-independent variable, the person measure  $\theta$ .

## 2. Theory

Landolt ring  $j$  is an object on a display screen that has the spatial-temporal dependent spectral radiant energy distribution  $Q_j(\mathbf{x}, \mathbf{y}, \mathbf{t}, \lambda)$ . The environmental optics and the optics of the eye of observer  $n$  produce an image of the Landolt ring on the observer's retina with spectral radiant energy distribution  $Q_{nj}(\mathbf{x}, \mathbf{y}, \mathbf{t}, \lambda)$ . The retinal image is sampled by a matrix of cone and rod photoreceptors with space-time dependent spectral absorption coefficients specific to person  $n$ :  $L_n(\mathbf{x}, \mathbf{y}, \mathbf{t}, \lambda)$ ,  $M_n(\mathbf{x}, \mathbf{y}, \mathbf{t}, \lambda)$ ,  $S_n(\mathbf{x}, \mathbf{y}, \mathbf{t}, \lambda)$ ,  $R_n(\mathbf{x}, \mathbf{y}, \mathbf{t}, \lambda)$  for the long, middle, and short wavelength-sensitive cones and rods, respectively. Thus, the retinal image of Landolt ring  $j$  for person  $n$  is transformed to the retinal image light absorption vector field

$Q_{nj}(x, y, t) = [\int_0^\infty a_n(x, y, t, \lambda) Q_{nj}(x, y, t, \lambda) d\lambda]$  for  $a = L, M, S, R$ . A vector of neurophysiologic space-time operators transform the light absorption vector field to a visual sensation vector field,  $V_{nj}(x', y', t') = \Phi_n\{Q_{nj}(x, y, t)\}$ , which corresponds to a visual sensory representation of Landolt ring  $j$  for observer  $n$ . Visual sensation is a vector field because the visual sensory representation of the retinal image corresponds to spatial and temporal variations in color and brightness, a personal experience of the observer. [6]

Observer  $n$  knows what he is looking for and has expectations for the appearance of the Landolt ring at each orientation  $\omega$ ,  $\hat{V}_{n\omega}(x', y', t')$ . Observer  $n$  performs a comparator operation on the visual sensation vector field and expected vector field to produce a scalar decision variable for stimulus  $j$  that is a function of orientation,

$D_{nj}(\omega) = \Psi_n\{V_{nj}(x', y', t'), \hat{V}_{n\omega}(x', y', t')\}$  (e.g., vector cross correlation at each orientation across  $x', y', t'$ ). Observer  $n$ 's response for stimulus  $j$  would be the orientation that produces the maximum (or minimum, depending on the comparator operation) value of  $D_{nj}(\omega)$ .

For a given population, the decision variable for stimulus  $j$  can differ between people because of fixed and stochastic between and within person differences in expectations and decision rules, i.e., differences in  $\hat{V}_{n\omega}(x', y', t')$  and  $\Psi_n$ , and because of fixed and stochastic differences between and within people in the global state of the visual system, i.e. differences in  $Q_{nj}(x, y, t, \lambda)$  due to differences in retinal image formation,  $a_n(x, y, t)$  due to space-time differences in the spectral absorbance of photopigments, and  $\Phi_n$  due to spatial and temporal variations in amplitude, gain, space and time constants, and other neurophysiological processing characteristics. If  $\bar{D}_j(\omega)$  is the expected value for the population of the decision variable for stimulus  $j$  as a function of orientation, then  $D_{nj}(\omega) = \bar{D}_j(\omega) + \theta_n + e_n$ , where  $\theta_n$  is the fixed difference and  $e_n$  is the stochastic difference for person  $n$  from the expected value.

The task of the observer is to identify the correct orientation of the Landolt ring,  $\omega_c$ . In terms of the theory, the observer chooses the orientation that corresponds to the maximum value of  $D_{nj}(\omega)$ . The probability that observer  $n$  makes the correct choice for stimulus  $j$  is the probability that  $D_{nj}(\omega_c) > D_{nj}(\omega)$  for all  $\omega \neq \omega_c$ . If the density function for  $e_n$  is a logistic that approximates a standard normal distribution, then the probability that  $D_{nj}(\omega_c) < D$  is  $Pr(D_{nj}(\omega) < D) = 1 - \frac{\exp(D - \bar{D}_j(\omega) - \theta_n)}{1 + \exp(D - \bar{D}_j(\omega) - \theta_n)}$  and the probability that  $D_{nj}(\omega_c) = D$  is  $Pr(D_{nj}(\omega_c) = D) = \frac{\exp(D - \bar{D}_j(\omega_c) - \theta_n)}{(1 + \exp(D - \bar{D}_j(\omega_c) - \theta_n))^2}$ . If  $e_n$  is stochastically independent with respect to orientation, given all possible values of the decision variable  $D$ , then the probability of person  $n$  making the correct choice for stimulus  $j$  is

$$Pr(D_{nj}(\omega_c) > D_{nj}(\omega)) = \int_{-\infty}^{\infty} Pr(D_{nj}(\omega_c) = D) \prod_{\omega \neq \omega_c} Pr(D_{nj}(\omega) < D) dD,$$

and the probability of making the incorrect choice of orientation  $i$  for stimulus  $j$ ,  $\omega_i$ , is

$$Pr(D_{nj}(\omega_i) > D_{nj}(\omega)) = \int_{-\infty}^{\infty} Pr(D_{nj}(\omega_i) = D) \prod_{\omega \neq \omega_i} Pr(D_{nj}(\omega) < D) dD.$$

The probability of choosing the correct orientation out of  $N_\omega$  equally likely and mutually

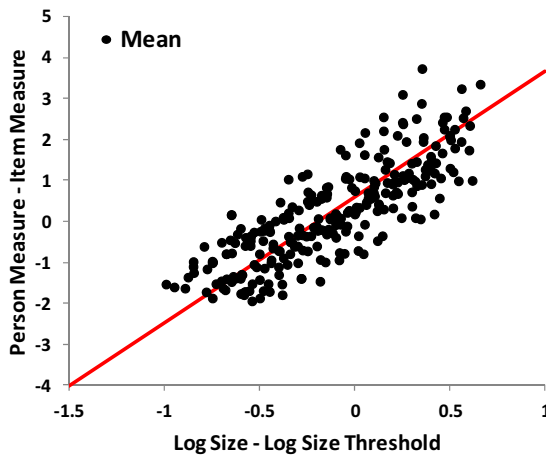
$$\text{exclusive orientations is } P_c = \frac{\Pr(D_{nj}(\omega_c) > D_{nj}(\omega))}{\Pr(D_{nj}(\omega_c) > D_{nj}(\omega)) + \sum_{i=1}^{N_\omega-1} \Pr(D_{nj}(\omega_i) > D_{nj}(\omega))}.$$

If  $\Pr(D_{nj}(\omega_c) > D_{nj}(\omega)) = \Pr(D_{nj}(\omega_i) > D_{nj}(\omega))$  for all  $i$ , then  $P_c = \frac{1}{N_\omega}$ , which is chance performance. This condition for chance performance is met when

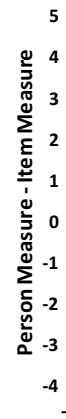
$$\Psi_n\{V_{nj}(x', y', t'), \hat{V}_{n\omega_i}(x', y', t')\} = \Psi_n\{V_{nj}(x', y', t'), \hat{V}_{n\omega_c}(x', y', t')\} \text{ for all incorrect orientations } i.$$

Subjects' choices of Landolt ring orientations for different combinations of stimulus parameters are scored as correct (1) or incorrect (0). From the assumption of local independence of observations, a maximum likelihood estimation routine is used to estimate  $\theta_n$  for every person and  $\bar{D}_j(\omega_c)$  for every stimulus from the matrix of response scores. [7] For Landolt ring size, contrast, and exposure duration, we define  $I(S, C, T)_j \equiv \bar{D}_j(\omega_c)$ .

The hypothesis of a global visual system state variable predicts that  $\theta_n - I(S, C, T) = a_{nS}(S - \hat{S}(C, T)_n) + b_{nS}$  and  $\theta_n - I(S, C, T) = a_{nC}(C - \hat{C}(S, T)_n) + b_{nC}$  for all combinations of S, C, T. Figure 5 is a scatter plot of the average  $\theta_n - I(S, C, T)$  vs the average  $S - \hat{S}(C, T)_n$  of 560 subjects with uncorrected refractive error (points) and Figure 6 is the similar scatter plot of the average  $\theta_n - I(S, C, T)$  vs  $C - \hat{C}(S, T)_n$  for the same subjects (each point corresponds to a pair of mean values for a unique combination of S, C, T).



**Figure 5.** Scatter plot of the difference between the estimated global visual state variable (i.e., person measure,  $\theta_n$  for person  $n$ ) and the stimulus-dependent expected value of the decision-variable for the population (i.e., item measure,  $I(S, C, T)_j$  for stimulus  $j$ ) vs. the difference between log stimulus size ( $S_j$  for stimulus  $j$ ) and the log size threshold ( $\hat{S}(C, T)_n$  for person  $n$ ). Each point is the average pair of values across 560 subjects for each combination of stimulus contrast ( $C$ ) and exposure duration ( $T$ ). The red line is the predicted relationship.



**Figure 6.** Scatter plot of the difference between the estimated global visual state variable (i.e., person measure,  $\theta_n$  for person  $n$ ) and the stimulus-dependent expected value of the decision-variable for the population (i.e., item measure,  $I(S, C, T)_j$  for stimulus  $j$ ) vs. the difference between log stimulus contrast ( $C_j$  for stimulus  $j$ ) and the log contrast threshold ( $\hat{C}(S, T)_n$  for person  $n$ ). Each point is the average pair of values across 560 subjects for each combination of stimulus size ( $S$ ) and exposure duration ( $T$ ). The red line is the predicted relationship.

The observed linear relationships in Figures 5 and 6 are consistent with our hypothesis that a single parameter-free variable,  $\theta_n$ , can be used to measure the global state of the visual system (at least for uncorrected refractive error) and predict multivariable visual resolution thresholds.

### 3. References

1. International Organization for Standardization. Optics and Optical Instruments. Visual Acuity Testing: Method of Correlating Optotypes, ISO 8597:1994. Geneva, Switzerland: International Organization for Standardizations; 1994.
2. Alexander KR, Wei X, Derlacki DJ. Visual acuity and contrast sensitivity for individual Sloan letters. *Vision Res.* 1997;37:813-819.
3. Nachmias J. Effect of exposure duration on visual contrast sensitivity of square-wave gratings. *J Opt Soc Am.* 1967;57:421-427.
4. Kelly DH. Motion and vision. II. Stabilized spatio-temporal threshold surface. *J Opt Soc Am.* 1979;69:1340-1349.
5. Alexander KR, McAnany JJ. Determinants of contrast sensitivity for the tumbling E and Landolt C. *Optom Vis Sci.* 2010;87:28-36.
6. Massof RW, Bird JF. A general zone theory of color and brightness vision. I. Basic formulation. *J Opt Soc Am.* 1978;68:1465-1471.
7. Royal KD, O'Neill TR. Using the CUTLO procedure to investigate guessing. *Rasch Measur Trans* 2011;251:1319-1320.