

THE SHAPE OF THE UNDERLYING DISTRIBUTIONS IN ABSOLUTE IDENTIFICATION EXPERIMENTS

Bruce A. Schneider
Centre for Research on Biological Communication Systems
University of Toronto at Mississauga
bschneid@utm.utoronto.ca

Abstract

In signal-detection analyses of one-dimensional, n -alternative, absolute-identification (AI) experiments it is usually assumed that the n stimuli give rise to n equal-variance, normal-distributions (EVNDs) along a uni-dimensional decision axis. However, Parker et al. (2002) have argued that equal-variance Laplace distributions (EVLDS) provide a better fit to AI data. This result is somewhat counter-intuitive, especially if the distribution of effects along the decision axis are thought to arise from noise (or an accumulation of small errors) in the decision process, which, according to the central limit theorem, should give rise to normal distributions. Here, we show that even when the data from AI experiments are generated from EVNDs, EVLDS will characterize the results, whenever the data are averaged across sessions (either within- or between-subjects) in which the underlying acuity (separation between distributions) is changing, a situation that is likely to occur whenever there are changes in gain-control.

In an absolute identification experiment, one of n stimuli is presented on a trial, and the observer is asked to identify it, usually by pressing one of n buttons. When the stimuli vary along only one physical dimension, it is typically assumed that the presentation of stimulus k ($1 \leq k \leq n$) gives rise to a response, r , along a unidimensional decision axis. Because of variability in the stimulus, and/or in its perceptual processing, the value of this response along the decision axis is assumed to vary from trial to trial. In a simple model of this process, each of the responses, r , evoked by repeated presentations of stimulus k , are assumed to be normally distributed, with mean $= \mu_k$, and standard deviation $= \sigma$. This characterization of the decision process is illustrated in Figure 1, for $n = 8$. In this model the observer is assumed to place $n-1$ criteria along this uni-dimensional decision axis, thereby dividing the decision axis into n response regions. The locations of these criteria are assumed to be influenced by stimulus probabilities, and payoffs for correct and incorrect responses. Hence, if the response elicited by stimulus k falls in response region j , the observer will identify it as stimulus j . The data from such an experiment therefore consists of the proportion of times stimulus k was identified as stimulus j for ($1 \leq k \leq n$) and ($1 \leq j \leq n$). Note that if the decision axis is normalized by dividing it by σ , the n distributions along the decision axis are all assumed to have standard deviations $= 1$. If the location of μ_1 is arbitrarily set to 0, there are $2n - 2$ free parameters ($n-1$ means, and $n-1$ criterion locations) in the model. Because there are $n(n-1)$ degrees of freedom in the data, the number of degrees of freedom remaining when the data are used to determine the $2n - 2$ parameters of the model is $n^2 - 3n + 2$. Parker et al. (2002) have described a procedure in which the parameters of the model are selected so as to minimize Pearson's Chi Square, i.e., minimize

$$\chi^2 = \frac{\sum_{k=1}^n \sum_{j=1}^n (R_{k,j} - E_{k,j})^2}{E_{k,j}} \quad (1)$$

where $R_{k,j}$ is the number of times stimulus k was labeled stimulus j , and $E_{k,j}$ is the expected number of times that stimulus k should be labeled as stimulus j , given the model shown in Figure 1.

However, when Parker et al. (2002) attempted to fit this equal-variance, normal-distribution (EVND) model to the data from a 4-stimulus absolute identification (AI) experiment (the four stimuli were 25, 30, 35, and 40 dB SPL 1-kHz tones), they found systematic deviations of the data from the model that were consistent with the notion that the kurtosis of the underlying distributions along the decision axis was much larger than that characteristic of normal distributions. This led them to test a model in which the distributions along the decision axis were assumed to be equal-variance Laplace (EVLD) distributions where the event, r , elicited by stimulus k has the probability density

$$\frac{1}{\sqrt{2} \sigma} e^{-\frac{\sqrt{2} |r - \mu_k|}{\sigma}}. \quad (2)$$

Parker et al. (2002) found that the EVLD model provided an excellent fit to the data. Subsequent experiments (yet to be published) have confirmed that the EVLD model provides a much better fit to the data in both the auditory and visual realms than the EVND model.

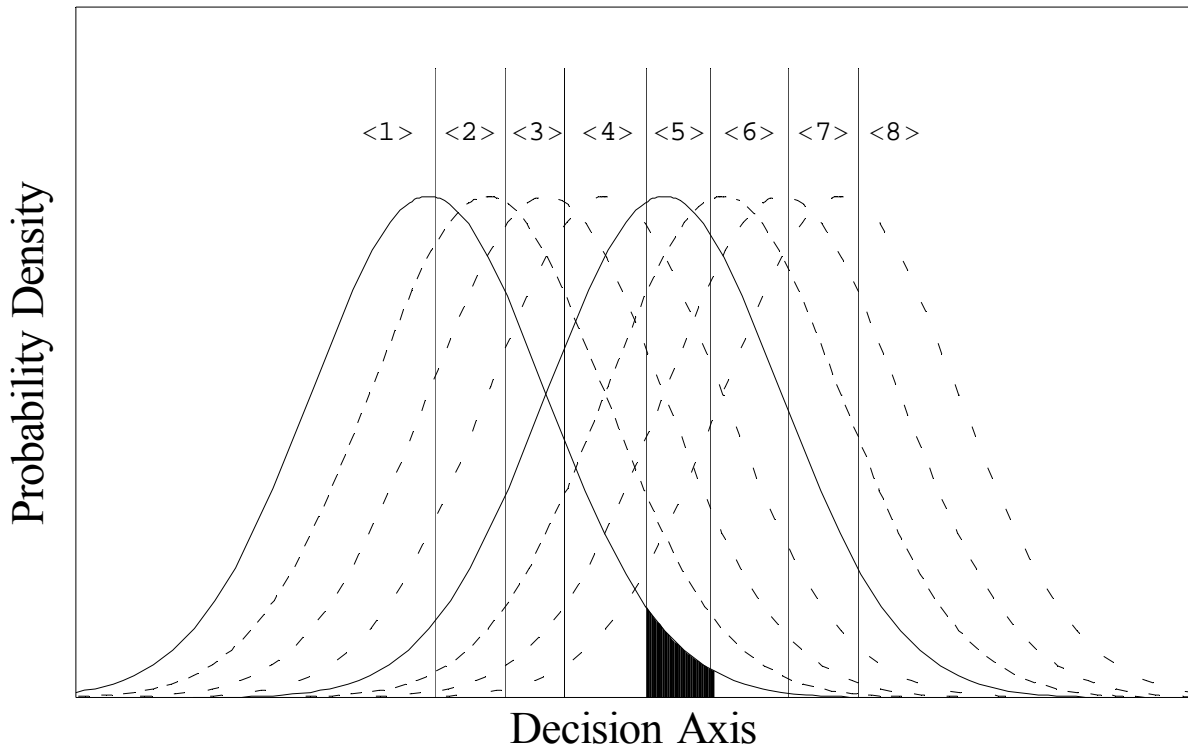


Figure 1. The EVND Model for an 8-alternative AI Experiment. Each stimulus, k , $\{1 \leq k \leq 8\}$ gives rise to a normal distribution of events along the decision axes. The vertical lines represent the criteria which divide the decision axis into 8 response regions (denoted by $\langle k \rangle$). All 8 distributions have equal variances. The shaded portion is the proportion of times stimulus 1 is labeled stimulus 5.

That the EVLD model provides a much better fit to the data from absolute identification experiments than the EVND model is somewhat surprising given the likely nature of sources of variance in perceptual tasks. Normally we assume that variability along the decision axis reflects noise in the stimulus or noise in the nervous system. Noise in the nervous system is often assumed to arise from a cumulation of errors at sensory-, perceptual-, and cognitive-level processing of the stimuli. If so, according to the central limit theorem, we would expect internal noise to be normally distributed. Why then does the EVLD model provide a better fit to the data than does a EVND model? One possibility is that averaging of data within or across individuals leads to a data set that is better fit by an EVLD model than by a EVND model. To evaluate this possibility, we generated AI data sets assuming an EVND model. These data sets differed in several ways. We then averaged the response matrices from these different data sets, and attempted to fit the average response-probability matrix with EVND and EVLD models. The specific ways in which the individual data sets differed are described below.

S1. Data sets generated by randomly perturbing the μ_k values across sets.

In this simulation $n = 8$ and $\sigma = 1$ in all data sets. In the first data set, the location of μ_k was set to $k-1 + x_k$, where the x_k are independent samples from a normal distribution whose mean is 0 and whose standard deviation is 0.25, subject to the constraint that the μ_k values must increase with k . Then the location of the seven criteria were assumed to fall midway between adjacent pairs of μ_k values, i.e., $c_j = (\mu_j + \mu_{j+1})/2$, for $1 \leq j < 8$. The model for the second data set was generated in a similar fashion. Figures 2A and 2B present two of the 10,000 models generated in this fashion. Response-probability matrices for each of these 10,000 models were then generated where, $p_{k,j}$, the probability that stimulus k is labeled stimulus j is

$$p_{k,j} = \frac{1}{\sqrt{2\pi}} \int_{r=c_{j-1}}^{c_j} e^{-\frac{1}{2}(r-\mu_k)^2} dr, \quad (3)$$

where $c_0 = -\infty$, and $c_8 = \infty$. These 10,000 response-probability matrices were then averaged to produce an average data set, which is labeled “ μ perturbed”.

S2. Data sets generated by randomly perturbing the c_j values across data sets.

In this simulation $n = 8$ and $\sigma = 1$ in all data sets. In the first data set, the location of μ_k was set to $k-1$. Then a criterion, c_j , was randomly selected from a rectangular distribution whose boundaries were $3\mu_j/4 + \mu_{j+1}/4$, and $\mu_j/4 + 3\mu_{j+1}/4$, for $1 \leq j < 8$. The model for the second data set was generated in a similar fashion. Figure 2C and 2D present two of the 10,000 models generated in this fashion. The 10,000 response-probability matrices so generated were then averaged to produce an average data set labeled “ c perturbed”.

S3. Data sets generated by randomly changing the discriminability of the stimuli.

In this simulation $n = 8$ and $\sigma = 1$ in all data sets. The discriminability of the stimuli was randomly varied by multiplying all μ_k by a random variable, α , drawn from a Rayleigh distribution whose variance parameter was set equal to 1. In the first data set, the location of μ_k was set to $\alpha(k-1)$. Then the location of the seven criteria were assumed to fall midway between

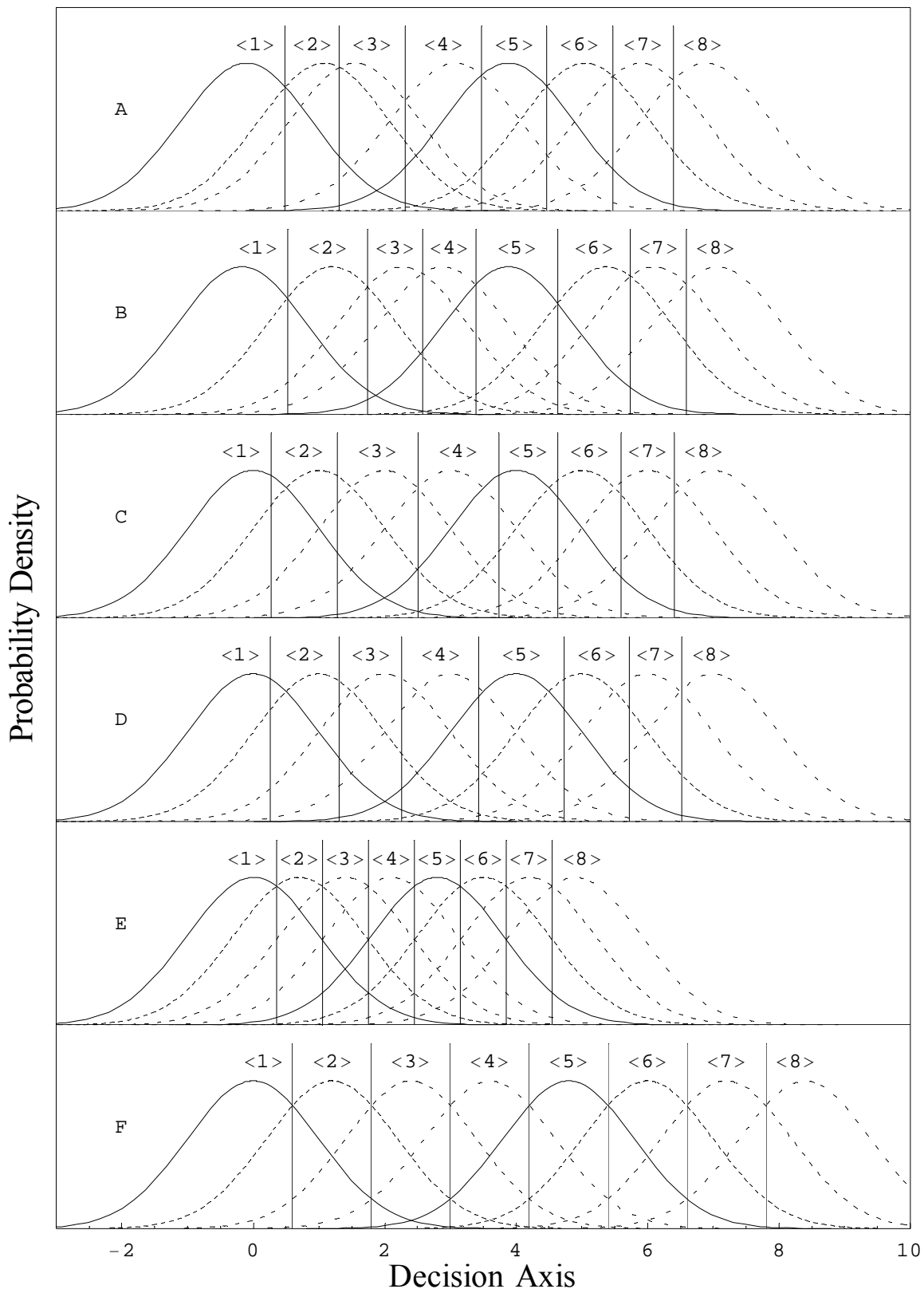


Figure 2. 2A and 2B. Instances in which the positions of the means of the distributions are randomly varied with criteria located at the intersections of adjacent distributions. 2C and 2D. Instances in which the separation between the means of the distributions remains constant but the locations of the criteria are varied. 2E and 2F. Instances in which the range of the distributions varies with criteria occurring at the intersections of adjacent distributions.

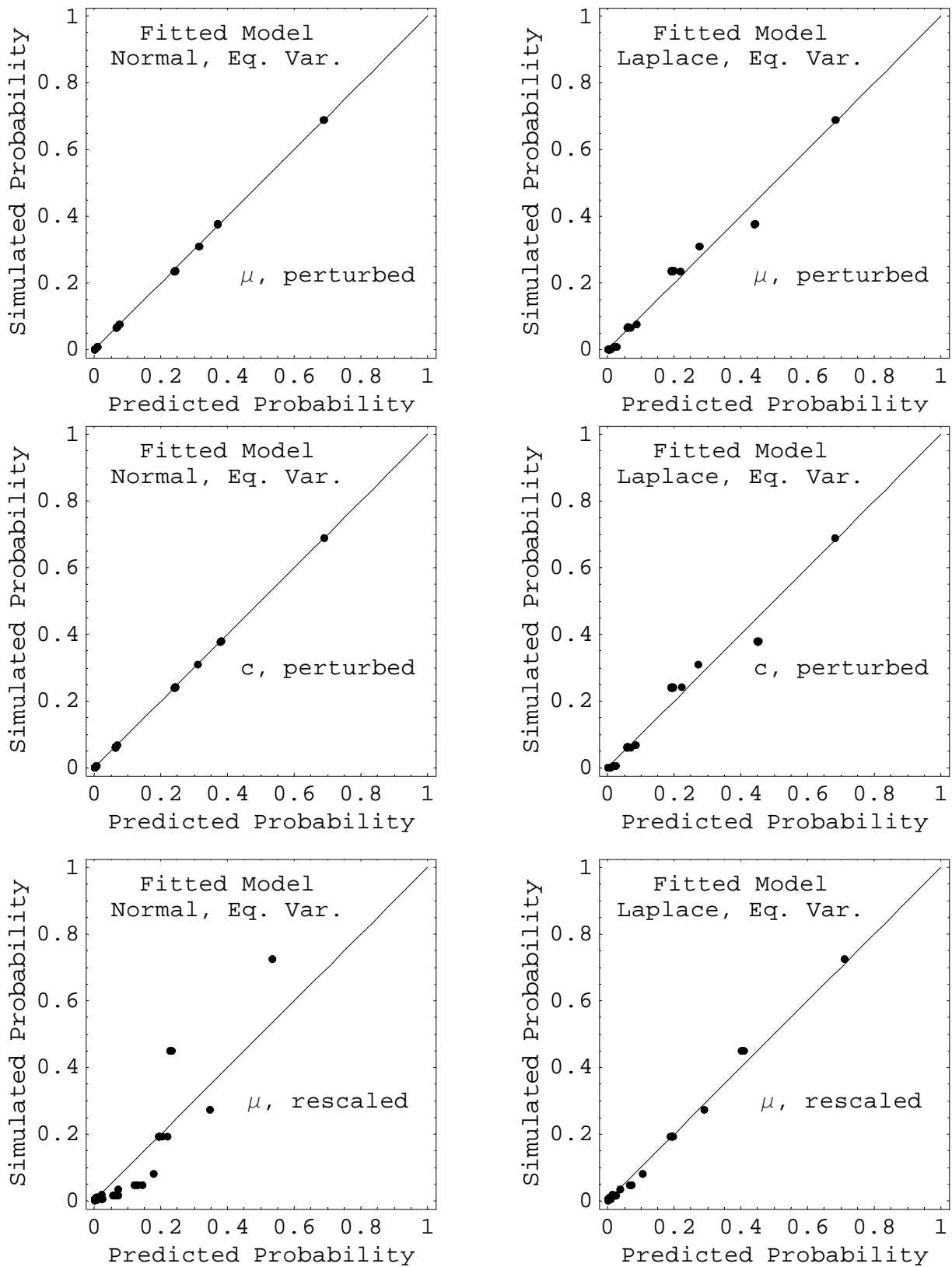


Figure 3. Simulated probabilities of a response given a stimulus were generated, assuming an EVND Model, when μ was perturbed, c was perturbed, and μ was rescaled (see text). The EVND (left column) and EVLD (right column) models were then fit to these 3 cases. This figure plots the simulated probabilities as a function of the probabilities predicted by the fitted models.

adjacent pairs of μ_k values, i.e., $c_j = (\mu_j + \mu_{j+1})/2$, for $1 \leq j < 8$. The model for the second data set was generated in a similar fashion. Figure 2E and 2F show two of the 10,000 models generated in this fashion. Note that multiplying the mean values by a scalar is exactly the same as leaving the locations fixed but altering σ by dividing by α . The 10,000 response-probability matrices so generated were then averaged to produce an average data set labeled “ μ rescaled”.

Results of the simulations

EVND and EVLD models were fit to the average response-probability matrices in all three conditions. Figure 3 plots the simulated probabilities as a function of the probabilities predicted by the two models for the average response probability matrices generated in S1, S2, and S3. Figure 3 shows that when the data are generated from a normal model, randomly perturbing the μ_k away from their mean values, or randomly changing the locations of the criterion, leads to an average response-probability matrix that is virtually perfectly described by the EVND model, and reasonably-well described by a EVLD model. Note, however, when the discriminability of the stimuli is varied randomly, data generated using equal-variance normal distributions leads to a response-probability matrix that is poorly described by the EVND model but well-described by the EVLD model. Additional simulations showed that jointly perturbing the means and criterion locations produced average response-probability matrices that were nearly perfectly described by an EVND model. However, any combination of perturbations which included changing the discriminability of the stimuli produced average response-probability matrices that could not be described by an EVND model but were well described by an EVLD model.

Implications for signal-detection models of AI experiments

When the response probability matrices are generated from an EVND model in which there is random variation in the locations of the means of the n distributions along the decision axis, but no change in the average discriminability of the stimuli, these response probabilities will be well-described by an EVND model. The same will hold true if there are random variations in the locations of the criteria. However, if the overall discriminability of the stimuli changes (variation in the standard deviation of their distributions along the decision axis), response-probability matrices generated from an EVND model will not be well-described by such a model, but rather will appear to fit an EVLD model. Within a subject, overall discriminability could be changing due to changes in gain control (changes in a nonlinear gain-control mechanism will change overall discriminability). Hence, if the gain applied to the input during an AI session is not uniform within a session or is not uniform across sessions, response-probability matrices obtained by averaging over changes in discriminability will be very well fit by an EVLD model, but not by an EVND model. In addition, if response-probability matrices are averaged across individuals who have different acuities, the average response-probability matrix will be well fit by an EVLD model even when the data from each individual are generated from an EVND model, and are well fit by an EVND model. Hence, averaging over individuals who differ with respect to sensory acuity, will lead to the conclusion that an EVLD model provides the better description of AI experiments, when, in fact, the distributions along the decision axes are normal in shape.

Reference

Parker, S., Murphy, D.R., & Schneider, B.A. Top-down gain control in the auditory system: Evidence from identification and discrimination experiments. *Perception & Psychophysics*, 64, 598-615.