

A random-ray model for Visual Search and Object Recognition

papers\rnRndw2003_NS.doc aug 7, 2003 word count: 3981 (text+ references)

Adam Reeves, Nayantara Santhi*, and Stefano DeCaro**

Department of Psychology, Northeastern University, Boston MA 02115, USA.

* Now at: Laboratory for Sleep Medicine, Brigham and Women's Hospital and
Harvard Medical School, Boston MA 02115, USA.

** Now at: Foo-Yin University, 151 Chin-Shueh Rd.,
Kaohsiung, Taiwan ROC

ABSTRACT

We present a 'random ray' model to describe Yes/No reaction times (RTs) and errors. The ray model is analogous to the random walk but it is computationally simpler, requiring only elementary geometry. Ray parameters control the drift rates to the Yes and No decision boundaries, prior bias, and a termination or 'time-out' rule. Rays are normally distributed, but predicted RT distributions are skewed by projection onto the boundaries. Model parameters can be estimated directly from the 16th, 50th, and 84th percentiles of the RT distributions on hit, correct rejection, false alarm, and miss trials, if the data satisfy three easily testable constraints. Examples are given from visual search and object recognition.

Random walk models provide a compelling account of observer's speed and accuracy on signal-present and signal-absent trials (Stone, 1960; Laming, 1968). The observer is assumed to initiate a 'walk' from an origin at the start of each trial. Each step of the walk is directed towards either a Yes boundary or a No boundary. The walk is governed by the *drift rate*, or speed at which evidence towards the signal is accumulated, and the *bias*, or start position relative to the two boundaries. Hits and false alarms occur when the walk crosses the 'Yes' boundary; misses and correct rejections occur when the walk crosses the 'No' boundary. In the discrete walk, one parameter determines both the mean and variance of the drift, but in a continuous random walk, the mean and variance can be controlled by different parameters, making this model flexible enough to fit real data remarkably well (e.g., Ratcliff, 1978, 2002; Ratcliff & Rouder, 1998). However, this leads to fairly complex and unintuitive equations. We therefore investigated a simpler model which we call the 'random ray', which retains the start point and boundary concepts from the walk, but not the diffusion. We do not regard the loss of diffusion as critical, as the actual diffusion of an entity such as a gas is at best a mere analogy.

Like the random walk, the random ray can account for joint variations in speed and accuracy and can handle entire distributions of RTs, not just means and variances. Hit and false alarm rates determine the criterion (x_c) and sensitivity (d'), as in signal-detection theory, after compensation for the fraction of "time-out" trials in which a boundary is not reached within a pre-set period. The ray model is limited to data which satisfy three constraints, concerning *bias*, *slow errors*, and *skew*. The bias constraint is that bias, or *a-priori* information at the start of a trial, is constant in any set of trials (FOOTNOTE 1). The slow error constraint (errors are slower than corrects) and skew constraint (that each RT distribution is sufficiently skewed) are detailed below.

In the discrete random walk illustrated in Fig. 1, the 'Yes' boundary is reached in the presence of a signal (a 'hit'). At each moment in time the evidence towards the signal increases by one upwards step, with probability P , or decreases (step down). The steps are random fluctuations on a steady drift (determined by P) towards the boundary. In the ray model, we simplify this picture by representing the various steps on this trial by a single *ray*, the dotted line from the origin (A) to the point R on the boundary. The decision

("Yes") is made at time $T = y = \text{distance OR}$. The corresponding RT or reaction time is $RT = T + RT_0$, where RT_0 represents the sensory plus motor or "residual" latency (Laming, 1968; Ratcliff, 1978). The distance from the origin to the Yes boundary is denoted c , and the distance to the No boundary is c_1 .

INSERT FIG.s 1,2 ABOUT HERE

Hits

In the ray model the variation of ray angle β over trials is assumed to be Gaussian with mean α and standard deviation s , with the tails truncated so that $0 < \beta < 180$ deg. Gaussian variation will arise if β (processing efficiency) is proportional to the sum of a large number of independent random effects. Fig. 2 illustrates the Gaussian distribution of rays on the slanted x-axis, not on the vertical axis as would be the case for a continuous random walk. The heavy up-sloping line towards the "Yes" boundary indicates the mean ray (m) and the dotted lines show rays at $x = m - \sigma$ and $x = m + \sigma$ on the x-axis. The angle made with the vertical by the mean ray is $\alpha = \text{mean}(\beta)$. The skewed distribution on the upper ("Yes") boundary represents the predicted distribution of the decision times (T) on hit trials. The skew, which is typical of choice RT distributions, is a consequence of a symmetrical (Gaussian) ray-generating process being projected through an angle.

Trigonometry (see Fig. 2) shows that y is a function of x and β , as follows:

$$y = OQ + QR = c \cos(\alpha) + h \tan(\alpha) = x \{ \cos(\alpha) + \sin(\alpha)\tan(\beta) \}, \text{ and}$$

$$c = OA = OB + BA = x \cos(\alpha)/\tan(\beta) + x \sin(\alpha). \quad \text{Eliminating } \tan(\beta),$$

$$y = x \cos(\alpha) c / \{ c - x \sin(\alpha) \} \tag{Eq 1}$$

Re-arranging, one can estimate the x corresponding to a crossing time, y ;

$$x = y_x / \{ \cos(\alpha) + y_x \sin(\alpha) \tan(\alpha) / y_{50} \}$$

where y_x is the x -th percentile of the y distribution.

The 16th, 50th, and 84th percentiles of the underlying Gaussian on the x-axis are $x = m - \sigma$, m , and $m + \sigma$, with m denoting the mean ray. These project to the corresponding percentiles y_{16} , y_{50} , and y_{84} on the upper horizontal axis. Thus

$$\begin{aligned}
m-\sigma &= y_{16} / \{ \cos(\alpha) + y_{16} \sin(\alpha) \tan(\alpha) / y_{50} \} \\
m &= y_{50} / \{ \cos(\alpha) + \sin(\alpha) \tan(\alpha) \} \\
m+\sigma &= y_{84} / \{ \cos(\alpha) + y_{84} \sin(\alpha) \tan(\alpha) / y_{50} \}
\end{aligned}
\tag{Eq 2}$$

Consider the points P, $P_{m-\sigma}$, and $P_{m+\sigma}$ on the x-axis in Fig. 2. Because the Gaussian is symmetrical the distances $(P-P_{m-\sigma})$ and $(P_{m+\sigma}-P)$ on the x-axis must be equal, so

$$\begin{aligned}
&y_{16} / \{ \cos(\alpha) + y_{16} \sin(\alpha) \tan(\alpha) / y_{50} \} - y_{50} / \{ \cos(\alpha) + \sin(\alpha) \tan(\alpha) \} \\
&= y_{50} / \{ \cos(\alpha) + \sin(\alpha) \tan(\alpha) \} - y_{84} / \{ \cos(\alpha) + y_{84} \sin(\alpha) \tan(\alpha) / y_{50} \}.
\end{aligned}$$

Solving for α ,

$$\tan(\alpha) = \sqrt{\{ (y_{84} + y_{16} - 2y_{50}) / (y_{84} + y_{16} - 2y_{84}y_{16}/y_{50}) \}}
\tag{Eq 3.}$$

Eq. 3 provides the constraint on skew. For the square root to be real, its argument must be positive, so if the numerator and denominator of Eq. 3 differ in sign, the ray model is ruled out. The term inside the square-root is independent of scale and origin, so RTo and σ need not be known to apply this check.

We developed the ray model using percentiles, rather than moments, as the higher-order moments of RT distributions tend to be unstable (Luce, 1986). Thus the underlying Gaussian is characterized parametrically, by its mean and variance, whereas the RTs are treated as if they were distribution free. This is a common approach in random walk modeling. Ratcliff (2002) fit five, rather than three, percentiles, and if the data warrant it, this improvement can be made by calculating $m+2\sigma$ and $m-2\sigma$ from y_{05} and y_{95} in Equ. 2.

INSERT FIG. 3 ABOUT HERE

Misses, Correct rejections, and False Alarms.

Misses occur when the random ray reaches the No boundary when the target is present. The geometry involves the distance to the No boundary, $c1$. Fig. 3 diagrams this

case. The distance from the lower left corner (point Q at time zero) to the ray which intersects the No boundary at point R is w , and

$$w = c1 \cdot \cot(\beta-90) \quad \text{Eq 3.}$$

The probability $p(w)$ that a ray intersects the No boundary between w and $w+\delta w$, is determined from the part of the Normal distribution that projects below the horizontal line through the start point. Then $p(w) = N(x_1) - N(x)$, where $\delta x = x_1 - x$ is the step along the x -axis which includes all rays contributing to $p(w)$, and $N(x)$ is the cumulative Normal from $-\infty$ to x . To obtain $p(w)$ as a function of β , one steps β along in small increments from 90 deg to 180 deg and obtains x_1 and x from $c/(\sin(\alpha)+\cos(\alpha)\cdot\cot(\beta))$ (Eq.1.)

Because the x -axis is slanted, each percentile of the miss distribution must be slower than the corresponding percentile of the hit distribution. This constraint (slow errors) on the ray model is independent of the start point ($c1$), RT_0 , and σ , and therefore can also be checked directly from the raw RT distributions.

To treat correct rejections and false alarms we invert the preceding diagrams. The values of c and $c1$ are reversed, so (e.g.) the distance to the upper ('No') boundary is now $c1$. Correct rejections and false alarms take the places of hits and misses respectively. The mean rate is α_N , the mean ray direction in the absence of the target stimulus or signal. The slow-errors constraint implies that each percentile of the false alarm RTs must be slower than the corresponding percentile of the correct rejection RTs.

The effect of the projection on skew

Figure 4 shows the effect of the projection of the normal curve (upper left plot) onto the upper (Yes) boundary (upper right plot) through an angle $\alpha = 62$ deg. The projected distributions are plotted both on linear and log axes. The skew is very evident on the linear time axis. Applying a log transform reduces the skew, making the distribution seem more bell-shaped (lower right plot), as is often the case for empirical RT distributions (e.g. DeCaro & Reeves, 2002). Such log-time symmetry can be expected for all drift angles from 10 to 80 deg, as indicated in the lower left plot, since $\log(\tan(\alpha))$

is almost exactly proportional to α in this range. This is a nice property of the model, although hardly unique.

INSERT Fig. 4 ABOUT HERE

The cut-off, or time-out, parameter Z_c .

The distribution of w includes cases with near-horizontal projections which can make w enormous, generating -violently skewed latency distributions. Estimating the expected value $w.p(w)$ numerically, we found that for a moderate miss rate of 5%, and an unbiased observer ($c=c1$), the *mean* miss RT should be 35 times longer than *mean* hit RT ! Since observers do not wait for such long times to emit a response, we assume they employ a ‘cut-off’ beyond which they guess entirely at random. In this case a cut-off at $Z_c = 3.5$ standard deviations, which eliminates the few very long predicted T's, suffices to bring the means into line. As long as the cut-off is *late*, i.e. delayed enough not to affect the 84th percentile, it will not disturb the latency predictions (FOOTNOTE 2). However, estimates of d' and the response criterion must still be corrected for the fraction of cut-off trials, since error trials in which the observer actually reaches the wrong boundary are now mixed in with error trials in which he or she times out and guesses incorrectly.

Data fitting: visual search and object recognition

We fit data from a visual search experiment that involved the search for an odd-colored target presented among 11, 15 or 20 heterogeneous (multicolored) distractors (Santhi, 2000). Participants were three highly-trained undergraduates who were not aware of the hypotheses of the study. In this experiment, set-size had almost no effect, so we collapsed the data to obtain 1200 target present and 1200 target absent trials for each participant. Data for two participants, HF and MS, satisfied all three constraints of the ray model. There are six model parameters, α , α_N , σ , $c1$, Z_c , and RT_0 ; parameter c is equated to 1 without loss of generality. There are 14 data points to fit per condition; the miss and false alarm rates, and the 12 actual Ts, i.e. the three percentiles from each of the

four types of RT trial, corrected for R_{To} . We estimated R_{To} from the participant's median simple RT obtained to the same target but no distractors, on the assumption that the decision components involved in the latency of interest (T) were not present in the simple RT (FOOTNOTE 3).

In "bias-symmetric" experiments in which signal and noise trials occur equally often and Yes and No responses are equally rewarded, there should be no *a-priori* bias towards one or other response. In this case $c=c1$, and the ray model can be solved directly as one can now obtain α and α_N from Eq.s 1 and 2, the parameter σ is estimated by the slope of the linear regression between the predicted and actual RT's, and Z_c can be found from the error rates as explained below. Fig.5 plots MS's predicted RT's against her observed RTs. The predicted RTs equal R_{To} plus the Ts predicted from the model with $c=c1$. Three RT percentiles are shown for each type of trial (hit, fa, cr, miss), yielding 12 points. The mean ray angles calculated from Equ 2 were $\alpha=77$ (upward, as in Fig. 2) and $\alpha_N=120$ deg (downward). Had these angles summed to 180, the ray-generating processes would have been "process-symmetric", i.e. the information for a correct Yes being identical to the information for a correct No. For what it is worth, we had expected process symmetry in the oddity experiment from the way the stimuli were generated, but MS's median correct Yes RT (544 msec) was slightly faster than her median correct No RT (553 msec), suggesting that her processing (unlike her bias) was not quite symmetrical. Indeed, forcing process-symmetry lowered the goodness of fit by 11%. We conclude that the unsymmetrical ray model is descriptively adequate, but that our expectation of complete symmetry was not met.

We now show how the cut-off, Z_c , can be estimated. The main body of the Gaussian shown in Fig. 2 projects up to generate hits, but the right tail dips below the horizontal and this will generate some misses. With $\alpha=77$ deg (as in MS's case), the right tail dips below the horizontal at a z-score of 1.28, predicting 10.0% misses. The complementary downward-facing Gaussian with $\alpha_N=120$ deg, whose main body accounts for correct rejections, cuts the horizontal at a z-score of 1.71, predicting 4.3% false alarms. In fact MS produced 11.0% misses and 4.0% false alarms, which differ by an average of 0.65% from the predictions. In an experiment with a guessing rate of one

half, Z_c is estimated as the z-score corresponding to twice this difference. In MS's case, a Z_c of 2.49 would time-out 1.3% of trials, and as half of these (0.65%) will become errors this harmonizes the predicted and obtained error rates.

For the entire set of 12 percentiles shown in Fig. 5 and given in Table 1, the averaged absolute error of prediction is 58 msec. The discrepancies are systematic; to predict a large enough inter-quartile range on the error trials, the model over-predicts the inter-quartile range on the correct ones. However, the discrepancies are not large and the correlation between predicted and actual RTs is $r=0.90$. Given that there were no free parameters in the fit, we regard this as a good example of the utility of the ray model.

If bias-symmetry cannot be assumed, a fitting procedure may be used. We chose to minimize the sum of squared errors between the 12 predicted and actual T's, measured in seconds, plus 6 times the sum of squared errors between the predicted and actual error rates, measured as proportions. The factor of 6 gives the error rates about the same weight as the latencies. We developed a MATLAB program which best-fits any sub-set, or all, of the parameters, starting with the initial solutions from Eq.s 1 and 2. In the case of MS it was possible to improve the predictions with all parameters free, but r improved only to 0.94, a marginal gain over the 0.90 obtained with no free parameters.

These results for MS were representative of those for HF in the visual search experiment. We also fit RT and accuracy data from an object recognition study (DeCaro and Reeves, 2002). In this experiment the participant had to decide if a masked picture matched a word or not. We ran two long-term subjects, one of whose data fit well, and the other's (BA) did not. BA's data are interesting in that they illustrate one way in which the ray model can fail. The skew constraint was met, so the mean up ray (26 deg) and down ray (141 deg) could be calculated. However, BA's median error RTs (miss 507 ms, fa 544 ms) were faster than her median correct RTs (hit 553 ms, cr 608 ms), violating the slow error constraint. Her 16th and 84th percentile error RTs were slower than the correct RTs, so only the medians violated the constraint (Table 1; BS). Applying the model anyway, the r for RTs was only 0.65 (Figure 6, lower panel). The error rates were 0.053 (fa's) and 0.064 (misses), which could be matched by the model with a cut-off at $Z_{cut} = 1.76$. However, imposing this cut-off hardly improved matters, and indeed freeing all the

parameters and using the MATLAB program to best-fit the data only improved r to 0.71, so at least in this case the slow error constraint was a useful one. Note that one obvious reason for the ray model to fail, namely the occurrence of fast guesses, cannot be invoked in this case because BA's 16th percentiles did not violate the slow error constraint.

Practiced vs. naïve subjects: the use of Perf.

The ray model requires data which are sufficiently rich to estimate RT percentiles on error trials as well as on correct ones. For short-term participants, it would be nice to have a metric compatible with the ray model from which one could ascertain their performance from sparser data, especially if results from long-term participants already supported the ray model. The median correct RT and the d' can be obtained from relatively few trials. We therefore chose a performance metric we call Perf, with the units of information $(d')^2$ per processing time,

$$\text{Perf} = (d')^2 / T \quad \text{Eq. 6.}$$

where T is the estimated processing time, i.e. $\text{Median}(\text{RT}) - \text{RT}_0$. This metric, $(d')^2 / T$, is a bias-free measure which arises in several signal-processing models in which information accumulates over time (Swensson and Thomas, 1974). We will relate Perf to the ray model, but first we briefly illustrate just how useful it can be. We were able to derive it from basic signal-to-noise considerations in visual search (Santhi & Reeves, 2003), which postulate that

$$d' = \text{signal/noise} = T \cdot c_s / \sqrt{T (m\sigma_E^2 + \sigma_I^2)}. \quad \text{Eq. 7.}$$

The signal ($T \cdot c_s$) equals the signal contrast c_s multiplied by observation time T . The noise terms are m , the number of noise sources (distractors in visual search), σ_E^2 , the noise per distractor, and σ_I^2 , the stimulus-independent noise produced by the rest of the visual system. All sources of noise are uncorrelated, noise is additive, and the noise is observed for the same length of time, T , as the signal.

Squaring Eq. 7 and bringing the *response* measures d' and T together yields the performance index (Perf) and leaves the search model terms behind on the right:

$$\text{Perf} = c_s^2 / (m\sigma_E^2 + \sigma_I^2). \quad \text{Eq. 8.}$$

Equation 8 predicted naïve subjects' search performance (Perf) for a wide variety

of set sizes(m) , stimulus contrasts (c), and variations in distractor noise (σ^2_E) (Santhi & Reeves, 2003), so we have particular confidence in Perf as a measure of performance.

To relate Perf to the ray model, we calculated Perf from the model's predicted error rates and median RTs, assuming symmetry ($c=c1$ and $\alpha_n = 180 - \alpha$) and a late cut-off ($Zc = 3.0$). Perf was proportional to drift rate (α) in the ray model, as shown in Fig. 8 (lower left), for $15 < \alpha < 85$ deg. This linear relation does break down if the predicted d' is very high (>3.5) and the drift angle is small (<15 deg), but since angles close to 0 (instant processing) and 90 (infinite slowness) are not of practical interest, and since some errors are virtually inevitable, this result provides a useful link between Perf and the ray model. Indeed, MS's Perf in the oddity data, for which α was in the safe zone at 77 deg, was actually 29.1 and was predicted to be 30.0 by the ray model, an error of only 3%. Thus less-than extreme variations in the drift rate of the ray model can be estimated fairly well from variations in Perf.

In conclusion, we have presented a simple method for ascertaining the parameters of a random ray model from error rates and percentile RTs in Yes/No experiments. We derived three necessary constraints which data must meet in order for the model to apply. The ray model is like a random walk in that information is accumulated up to a boundary (for Yes, No, or time-out), but it sacrifices the hypothetical diffusion process for relative simplicity. We illustrated the model calculations with data which fit fairly well (MS's) and with data which did not (BA's). We also showed that if processing is known to obey the ray model, variations in ray angle may be tracked by variations in Perf, a useful metric which collapses speed and accuracy into one unique bias-free index (FOOTNOTE 4). What we have not yet done is to apply the ray model to a wide range of experimental data and find out where and when it succeeds.

FOOTNOTE 1. The constant-bias constraint may be violated if temporary shifts in bias generate sequential effects between trials (Laming, 1968). Fortunately *systematic* shifts in bias can be handled by data sorting. For example, if the walk typically starts closer to the Yes boundary after a signal trial than after a noise trial, the data can be conditioned on the previous trial and the ray model applied separately to each partition.

FOOTNOTE 2. Assuming Z_c is late and *constant* implies an abrupt end to the tail of the RT distribution and thus unrealistic predictions of the highest percentiles. Making Z_c a random variable might handle this problem, but complicates matters greatly, and to little purpose since few observations are involved and the predicted error rates and RT percentiles are not affected.

FOOTNOTE 3. Since simple RTs vary from trial to trial, it is better to deconvolve the simple RT distribution from the RT distributions to obtain the T distributions. This requires that the RT moments are larger than the corresponding simple RT moments, and as this was not always so in our data (see Laming, 1968), we took RT to be constant.

FOOTNOTE 4. Many RT papers ignore infrequent errors, but this may be dangerous. Variations in error rates may appear insignificant only because the assumptions of ANOVA are violated. To make this less likely, percentages or frequencies should be converted to proportions of trials, P , and arc-sine transformed to $2 \cdot \sin^{-1}(\sqrt{P})$, when $0 < P < 1$. (If $P=0$ or $P=1$, substitute $1/(2n)$ or $1-1/(2n)$, for n trials.) Also, since using RT exclusively is only justified if errors are constant over conditions, the null hypothesis (that errors are constant) is *avored* by using a nominally 'conservative' alpha level such as 5% or 1%; a higher alpha level such as 10% actually provides a more stringent test. Finally, we note that small variations in error rates have large effects on d' if the error rates are low, so when errors are infrequent, ignoring variations in them is particularly dangerous.

References

- De Caro, S. A., & Reeves, A. (2002) The use of word-picture verification to study basic-level object recognition: further support for view-invariant mechanisms. Memory & Cognition, 30(5), 810-820.
- Laming, D. R. J. (1968). Information theory of choice-reaction times. London: Academic Press.
- Luce, R. D. (1986). Response Times: their role in inferring elementary mental organization. New York: Oxford University Press.
- Ratcliff, R. (1978) A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Ratcliff, R., & Rouder, J. N., (1998). Modeling response times for two-choice decisions. Psychological Science: 9(5), 347-356.
- Ratcliff, R. (2002) A diffusion model account of response time and accuracy in a brightness discrimination task. Psychonomic Bulletin & Review, 9, 278-291.
- Santhi, N. (2000). The role of distractor coherence and target certainty in feature search: a signal detection approach. Unpublished dissertation, Northeastern University.
- Santhi, N. & Reeves, A. (2003) The roles of distractor coherence and target certainty on visual search: a signal detection model. Submitted to Vision Research.
- Stone, M. (1960). Models for choice-reaction time. Psychometrika, 25, 251-260.
- Swensson R. G.& Thomas R E. (1974). Fixed and optional stopping models for two-choice discrimination times. Journal of Mathematical Psychology, 15, 282-291.

Table 1: Percentiles (%) of the RT distributions in signal-present (S) and absent (N) trials.

Values in parentheses are predicted. Participants: MS and BA

Participant	Correct trials			Error trials			
	16%	50%	84%	16%	50%	84%	
MS hit	479 (521)	544 (529)	659 (589)	miss	448 (545)	603 (579)	796 (726)
MS cr	479 (522)	553 (532)	672 (567)	fa	461 (586)	714 (669)	976 (1020)
BA hit	408 (522)	553 (527)	748 (536)	miss	361 (618)	507 (727)	1001 (1032)
BA cr	453 (522)	608 (526)	898 (535)	fa	404 (570)	544 (632)	1149 (886)

A single random walk or ray

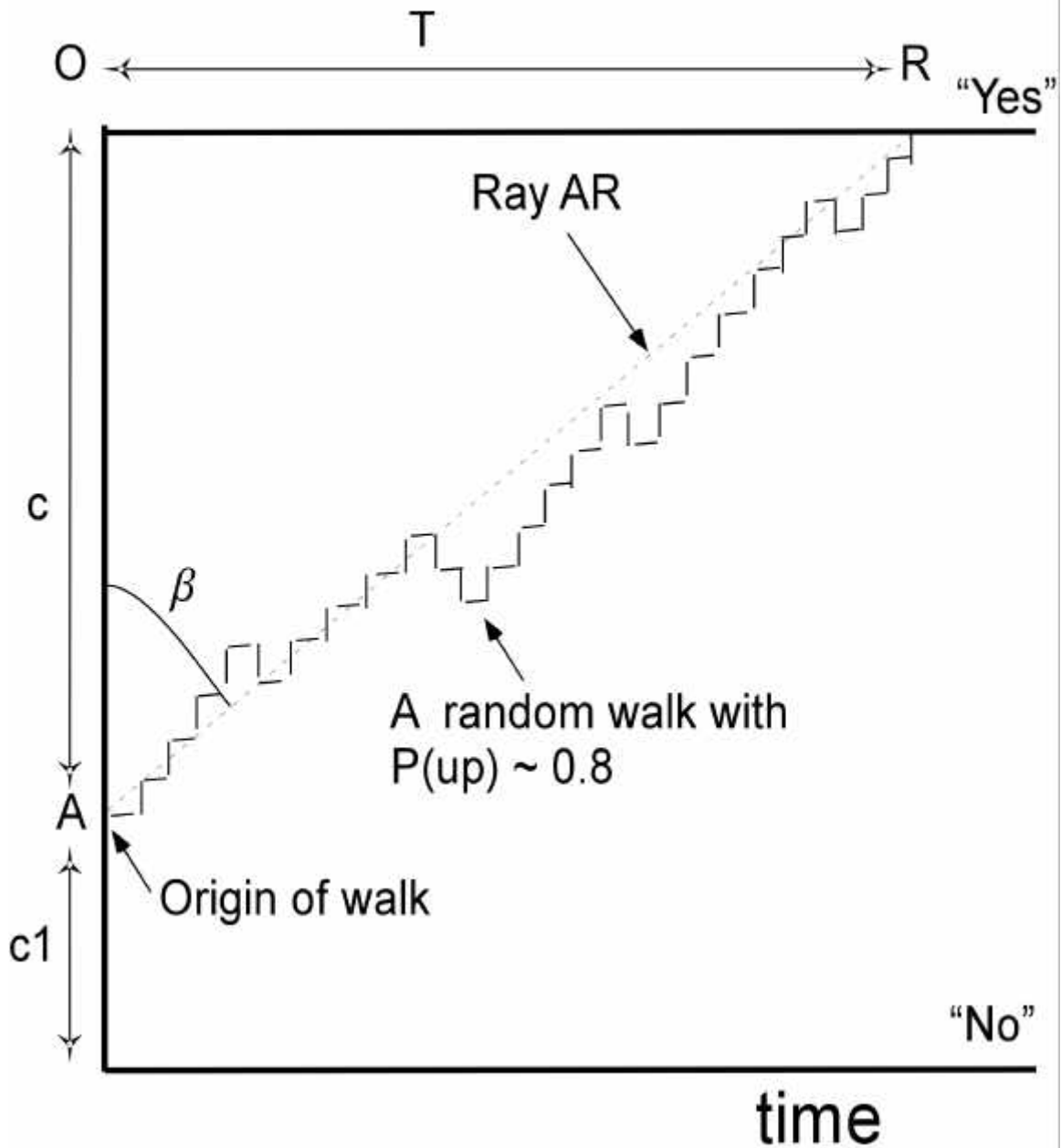


Fig. 1

adam2.cnv 6/7/2002

Deriving x from Percentiles of y

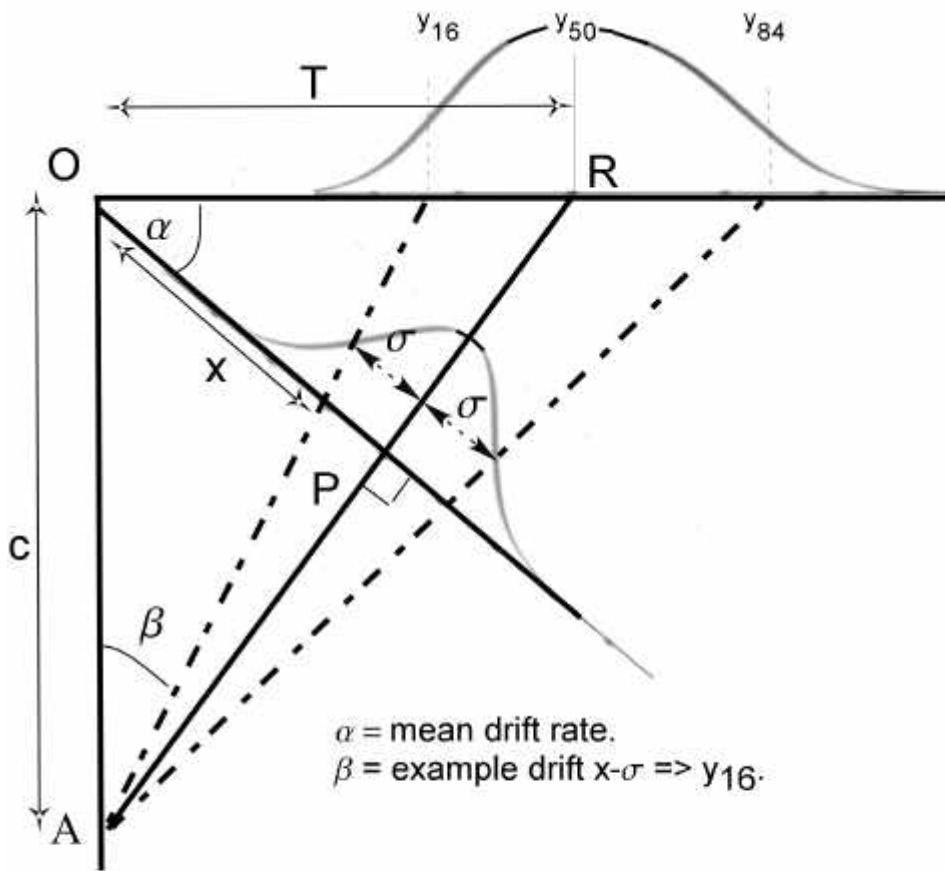


Fig. 2

adam5a.onv

The position (w) at the lower boundary

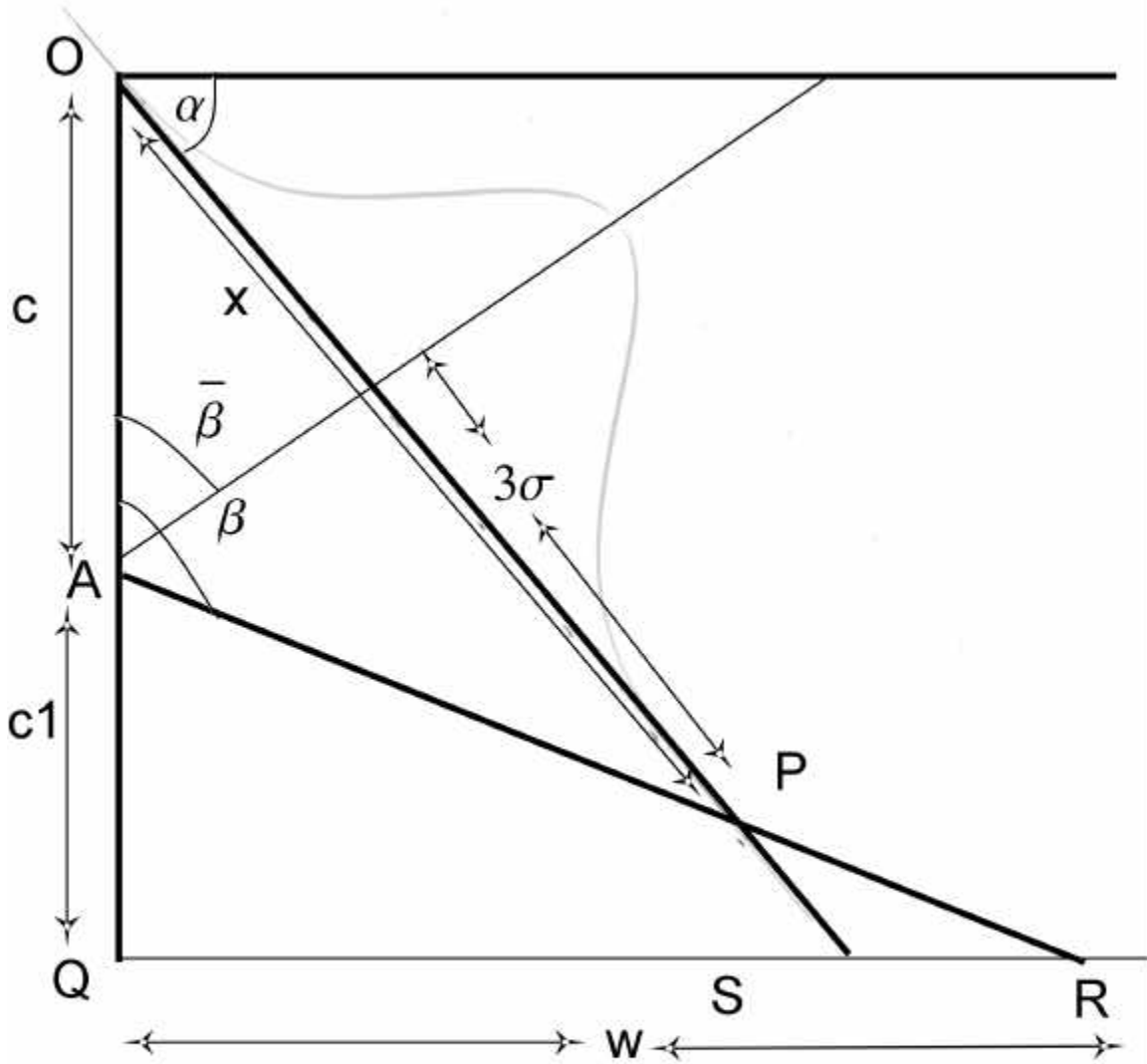


Fig. 3

adam6.cnv Fig 6

Projection of oddity hit RTs

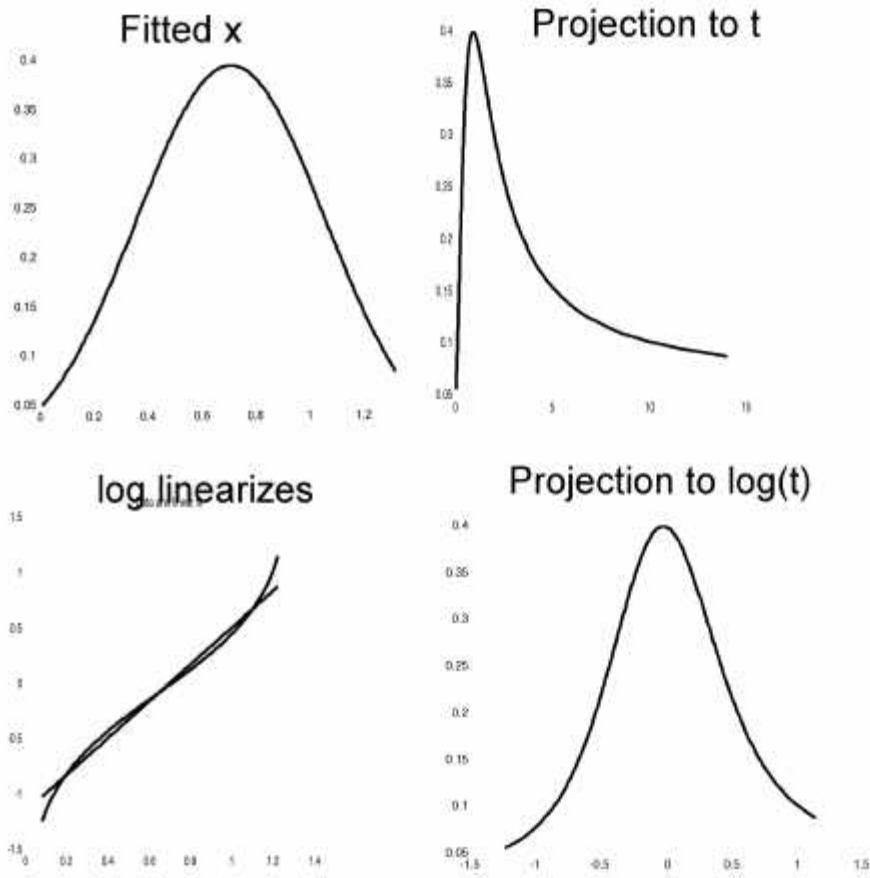


Fig. 4

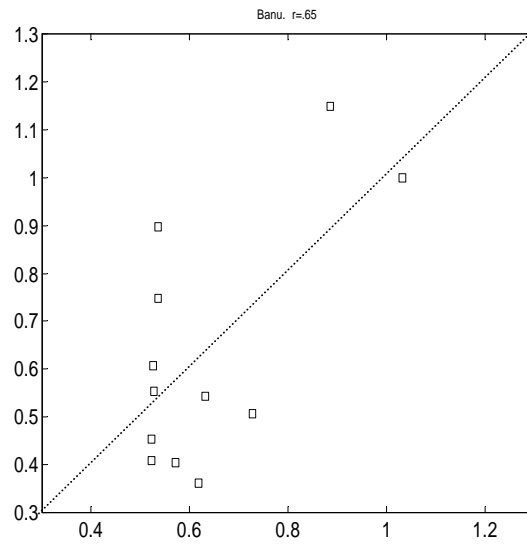
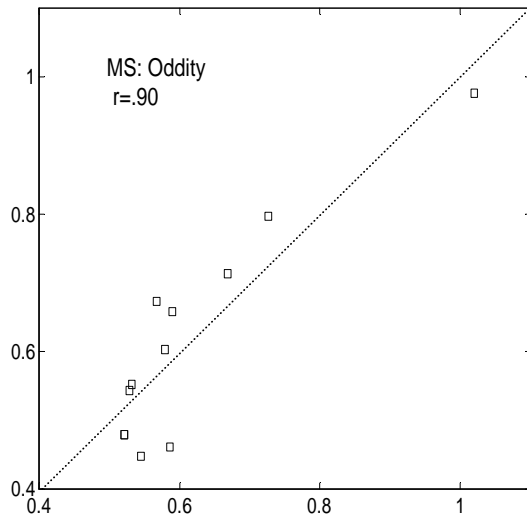


Figure 5: top, participant MS; below; BA.

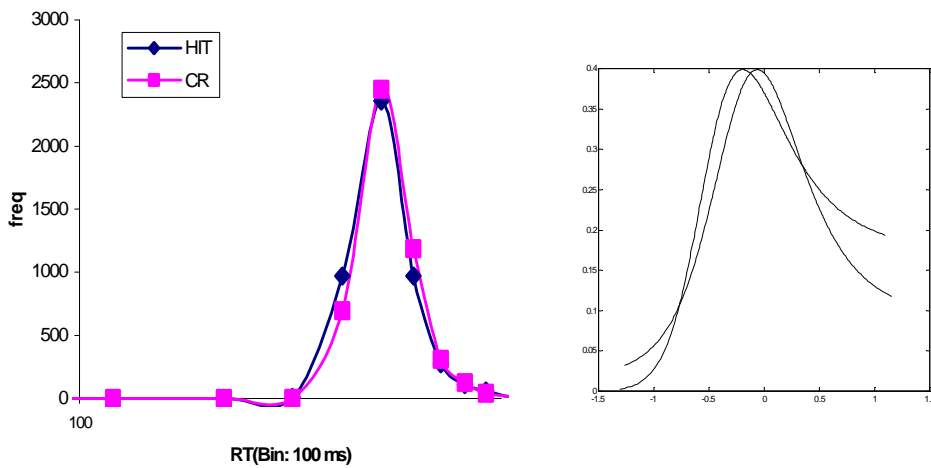


Figure 6 Participant MS (left) and ray model (right). Plots show frequency versus log time for hit RTs and correct rejection RTs. The model predicts some deviation which is not present in the data, but the overall shape and placement of the curves is about right.