Appendix: Model Derivation

mADC decision rule for a multialternative attention task employing the ‘method of constant stimuli’

The conventional attention task shown in Fig.1A involves ‘M+1’ stimulus events: changes at one of ‘M’ locations or no change.

We denote the change event at each location as $S_1, ..., S_M$, and the no change event as S_0. Similarly, we denote the event (or response) locations with indices $1, 2, \ldots, M$ and denote the no change event (or response) with the index 0 (or ϕ). Finally, we denote p_k as the prior probability of change at location k and p_0 as the prior probability of no change.

Extending one dimensional signal detection models, we define a multivariate decision variable, $\boldsymbol{\Psi}$, whose components Ψ_j denote the scalar decision variable at each location j where $j \in \{1, 2, ..., M\}$.

We derive the optimal decision rule for the case where the change at each location k can occur at one of N_k different strengths (change magnitudes). The derivation is based on Bayesian decision theory for the optimal detection of signals in noise (Middleton and van Meter, 1955).

The observer makes a decision/response γ_k to one of M locations k or no response 0. This is denoted by the variable $\delta(\gamma_k|\boldsymbol{\Psi})$, such that

$$\begin{aligned} \delta(\gamma_k|\boldsymbol{\Psi}) &= 1 \qquad \text{if the observer chooses location k} \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{1}$$

$\sum_{k=0}^{M} \delta(\gamma_k|\mathbf{\Psi}) = 1$, since the observer can choose only one of the m locations (or 0 in case of no change).

In this case, the average Bayesian risk is given by:

$$R = \int \sum_{k=0}^{M} \Big[C_k^k \delta(\gamma_k|\mathbf{\Psi}) + \sum_{\substack{j=0 \\ j \neq k}}^{M} C_j^k \delta(\gamma_j|\mathbf{\Psi}) \Big] p_k w_k d\mathbf{\Psi} \tag{2}$$

where C_j^k is the cost of indicating a response to location j when the change occurred at location k, where $k, j = 0, 1, 2, ..., M$ and $w_k = p(\mathbf{\Psi}|S_k)$, i.e. the conditional probability density of the decision variable given that a change occurred at location k. Grouping terms by $\delta(\gamma_k|\mathbf{\Psi})$,

$$R = \int \sum_{j=0}^{M} \Big[\sum_{k=0}^{M} C_j^k p_k w_k \Big] \delta(\gamma_j|\mathbf{\Psi}) d\mathbf{\Psi} \tag{3}$$

Since only one of the $\delta(\gamma_j|\mathbf{\Psi}) \neq 0$, our objective is to find the j that minimizes the sum $\sum_{k=0}^{M} C_j^k p_k w_k$.

Let us consider the N_k magnitudes of changes associated with events S_k at each location k, and posit that each occurs with prior probability p_k^n and with conditional density w_k^n. Thus,

$$p_k w_k = \sum_{n=1}^{N_k} p_k^n w_k^n \tag{4}$$

We assume also that costs are not different for the different magnitudes of change at each location. The sum to be minimized can then be written as:

$$\sum_{k=0}^{M} \sum_{n=1}^{N_k} C_j^k p_k^n w_k^n \tag{5}$$

Let us consider choosing between two alternatives i & j. The choice between i & j depends on the relative average risks, given by:

$$\begin{aligned} &C_j^0 p_0 w_0 + C_j^1 p_1 w_1 + ... + C_j^i p_i w_i + ... + C_j^j p_j w_j + ... + C_j^M p_M w_M \underset{j}{\overset{i}{\gtrless}} \\ &C_i^0 p_0 w_0 + C_i^1 p_1 w_1 + ... + C_i^i p_i w_i + ... + C_i^j p_j w_j + ... + C_i^M p_M w_M \end{aligned} \tag{6}$$

where the notation $\underset{j}{\overset{i}{\gtrless}}$ means choose i if LHS>RHS, and choose j otherwise.

We assume

$$C_j^l = C_i^l \quad \forall l \neq i, j \tag{7}$$

implying that the cost of making an incorrect response to any location (i or j) is the same for changes at a given location (l). Generally, we assume that for every event type S_k, the cost of making an incorrect response is the same, although this cost may be different across different event types.

With this simplification, equation (6) reduces to:

$$(C_j^j - C_i^j) p_j w_j \underset{j}{\overset{i}{\gtrless}} (C_i^i - C_j^i) p_i w_i$$

$$\frac{p_j w_j}{p_i w_i} \underset{j}{\overset{i}{\lessgtr}} \frac{(C_i^i - C_j^i)}{(C_j^j - C_i^j)}$$

where, the reversal of inequality happened because $(C_j^j - C_i^j)$ is defined to be non-zero and negative because the cost for making correct response is less than cost for making an incorrect response.

The left hand side of the previous equation is the generalized likelihood ratio, or the posterior odds ratio (product of the prior odds and the liklihood ratio). Introducing new notation for the cost ratio term on the right hand side,

$$\frac{p_j w_j}{p_i w_i} \underset{j}{\overset{i}{\lessgtr}} \beta_{ji}, \text{ where } \beta_{ji} = \frac{(C_i^i - C_j^i)}{(C_j^j - C_i^j)} \tag{8}$$

Substituting (4), this equation becomes

$$\frac{\sum_{n=1}^{N_j} p_j^n w_j^n}{\sum_{n=1}^{N_i} p_i^n w_i^n} \underset{j}{\overset{i}{\lessgtr}} \beta_{ji}, \text{ where } i, j \in \{0, 1, 2, ..., M\} \tag{9}$$

In our experiment, $N_i = N_j = N$ (6 change angles in our task, Fig. 1A) and $p_i^n = \frac{p_i}{N}$, $p_i^n = \frac{p_i}{N}$ $\forall i, j \in 0, 1, 2, ..., M$ (prior probabilities are evenly divided among the change angles). The only exception to this formula is for the no-change event for which $N_0 = 1$ (no change events do not occur at different change magnitudes).

We first examine the case for making a decision regarding change at location j vs no change.

$$\frac{\frac{p_j}{N}\sum_{n=1}^{N} w_j^n}{p_0 w_0} \underset{\phi}{\overset{j}{\gtrless}} \beta_{j0} \quad \text{or} \quad \frac{\sum_{n=1}^{N} w_j^n}{w_0} \underset{\phi}{\overset{j}{\gtrless}} \beta_{j0}\frac{p_0}{p_j}N \tag{10}$$

where ϕ denotes a no change response. We assume that w_j^ns are multivariate Gaussians with mean $[0, 0, ..., d_j^n, ..., 0]$ (an $m \times 1$ vector and identity covariance matrix. Again, as a special case $d_0^n = 1$ $\forall n$; no change distribution has zero mean. Thus,

$$w_j^n \sim \mathcal{N}^M(d_j^n, I) = \prod_k e^{-(\Psi_k - \mu_k^n)^2}$$

Now, equation (10) above simplifies to

$$\frac{1}{N}\sum_{n=1}^{N} e^{\Psi_j d_j^n - \frac{(d_j^n)^2}{2}} \underset{\phi}{\overset{j}{\gtrless}} \beta_{j0}\frac{p_0}{p_j} \tag{11}$$

Therefore, optimal decision surfaces are hyperplanes of constant Ψ_j even if, in this case there is no evident closed form solution for Ψ_j. These hyperplanes correspond to the criteria, $\Psi_j = c_j$. Consequently, the generalized likelihood ratio measure of bias for choices to location j is defined as:

$$b_j = \frac{1}{\frac{1}{N}\sum_{n=1}^{N} e^{c_j d_j^n - \frac{(d_j^n)^2}{2}}} = \frac{p_j/p_0}{\beta_j} \tag{12}$$

which is the harmonic mean of the biases calculated from individual sensitivity values for each change angle (d_j^n).

We next examine the case for making a decision at location j vs location i.

$$\frac{\frac{p_j}{N}\sum_{n=1}^{N} w_j^n}{\frac{p_i}{N}\sum_{n=1}^{N} w_i^n} \underset{i}{\overset{j}{\gtrless}} \beta_{ji} \tag{13}$$

Dividing the numerator and denominator by $p_0 w_0$ & after some algebra,

$$\frac{\sum_{n=1}^{N} e^{\Psi_j d_j^n - \frac{(d_j^n)^2}{2}}}{\sum_{n=1}^{N} e^{\Psi_i d_i^n - \frac{(d_i^n)^2}{2}}} \underset{i}{\overset{j}{\gtrless}} \frac{\beta_{ji}}{(p_j/p_i)} \tag{14}$$

Thus, the optimal decision surface for choosing change at location i versus change at location j are not hyperplanes. However, planar decision surfaces provide excellent approximations to these optimal surfaces, as shown in Supplementary Fig. S2B, and as derived next.

Demonstration that optimal decision surfaces for the choice between locations i & j can be closely approximated by planar surfaces:

We apply the logarithm on both sides of equation (14):

$$log\left(\frac{\sum_{n=1}^{N} e^{\Psi_j d_j^n - \frac{(d_j^n)^2}{2}}}{\sum_{n=1}^{N} e^{\Psi_i d_i^n - \frac{(d_i^n)^2}{2}}}\right) \underset{i}{\overset{j}{\gtrless}} log\Big(\frac{\beta_{ji}}{(p_j/p_i)}\Big) \tag{15}$$

Since the numerator and denominator are identical in form, except for the location indices, we consider the numerator term alone ($log\big(\sum_{n=1}^{N} e^{\Psi_j d_j^n - \frac{(d_j^n)^2}{2}}\big)$), without loss of generality. We consider two cases, a) for small Ψ_j and b) for large Ψ_j.

a) First, consider small Ψ_j. $\lim_{\Psi_j \to 0} e^{\Psi_j d_j^n} = (1 + \Psi_j d_j^n)$, where, we have ignored terms $O(\Psi_j^2)$ and higher in the expansion for $e^{\Psi_j d_j^n}$. Thus, the numerator simplifies to:

$$\text{Nr.} \approx log\Big(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}.(1 + \Psi_j d_j^n)\Big) \tag{16}$$

Further simplifying,

$$\text{Nr.} \approx log\Big(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}} + \sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}} \Psi_j d_j^n\Big)$$

$$= log\left(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}} \left(1 + \frac{\Psi_j \sum_{n=1}^{N} d_j^n e^{-\frac{(d_j^n)^2}{2}}}{\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}}\right)\right)$$

$$= log\left(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}\right) + log\left(1 + \frac{\Psi_j \sum_{n=1}^{N} d_j^n e^{-\frac{(d_j^n)^2}{2}}}{\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}}\right)$$

$$\approx log\left(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}\right) + \Psi_j\left(\frac{\sum_{n=1}^{N} d_j^n e^{-\frac{(d_j^n)^2}{2}}}{\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}}\right)$$

where, we have used the approximation $log(1+x) \approx x$ for small x. This has the form,

$$\text{Nr.} = B_j + \Psi_j A_j \tag{17}$$

$$\text{where, } A_j = \frac{\sum_{n=1}^{N} d_j^n e^{-\frac{(d_j^n)^2}{2}}}{\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}} \quad B_j = log\left(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}}\right)$$

An identical form holds for the denominator term in equation (15) for small Ψ_i.

b) Next, consider the numerator term from equation (15) for large Ψ_j

$$\text{Nr.} = log\Big(\sum_{n=1}^{N} e^{-\frac{(d_j^n)^2}{2}} e^{\Psi_j d_j^n}\Big)$$

$$= log\left(e^{\Psi_j d_j^N}\Big[\Big(\sum_{n=1}^{N-1} e^{-\frac{(d_j^n)^2}{2}} e^{\Psi_j(d_j^n - d_j^N)}\Big) + e^{-\frac{(d_j^N)^2}{2}}\Big]\right)$$

We consider d_j^ns to be sorted such that $d_j^N > d_j^{N-1} > ... > d_j^1$, which represent to the sensitivities corresponding to different change magnitudes (d_j^N, corresponding to the largest, and d_j^1 to the smallest change magnitude).

As $\lim_{\Psi_j \to \infty} e^{\Psi_j (d_j^n - d_j^N)} \to 0$ ($\because d_j^n - d_j^N < 0,\ \forall n < N$), this simplifies to

$$\text{Nr.} \approx log\big(e^{\Psi_j d_j^N} e^{-\frac{(d_j^N)^2}{2}}\big) \tag{18}$$

$$= \Psi_j d_j^N - \frac{(d_j^N)^2}{2}$$

$$= \Psi_j A_j^{'} + B_j^{'} \tag{19}$$

$$\text{where, } A_j^{'} = d_j^N \quad \& \quad B_j^{'} = -\frac{(d_j^N)^2}{2}$$

Again, an identical form holds for the denominator term in equation (15) for large Ψ_i.

Finally, we compute the decision surfaces for the choice between locations i and j for the different combinations of Ψ_i and Ψ_j magnitudes.

For small Ψ_i, large Ψ_j, the decision surface are given by (equations (17) and (19)):

$$\Psi_j A_j^{'} + B_j^{'} - \Psi_i A_i - B_i = log\eta_{ji} \tag{20}$$

Similarly, for large Ψ_j, large Ψ_i, the decision surfaces are given by (equation (19)):

$$\Psi_j A_j^{'} + B_j^{'} - \Psi_i A_i^{'} - B_i^{'} = log\eta_{ji}, \quad \text{where, } \eta_{ji} = \frac{\beta_{ji}}{(p_j/p_i)} \tag{21}$$

Decisions surfaces for the other two combinations (small Ψ_j, large Ψ_i and small Ψ_j, small Ψ_i) may be similarly derived. In each case, the decision rule is approximated by a planar decision surface (decision hyperplane).