Supplemental Data: Appendix

A Demonstration of model identifiability

A.1 Identifiability of of 2-ADC model

In this section, we demonstrate that there is a one-to-one mapping from the set of parameters $\{d_i, c_i\}, i \in \{1, 2\}$ to the set of response probabilities of the 2-ADC model. As these proofs do not assume a specific form of the decision variable distribution (e.g., Gaussian), we use a slightly different notation for these distributions. Thus, the probability density and the cumulative distribution of the decision variable distribution at location i under noise (catch trials) are given by f_i and F_i, respectively (Methods). Thus, equation system 3 can be rewritten as:

$$p(Y = 1|\mathbf{X}) = \int_{c_1 - d_1 X_1}^{\infty} F_2(e + d_1 X_1 - d_2 X_2 - (c_1 - c_2))\, f_1(e)\, de$$

$$p(Y = 2|\mathbf{X}) = \int_{c_2 - d_2 X_2}^{\infty} F_1(e + d_2 X_2 - d_1 X_1 - (c_2 - c_1))\, f_2(e)\, de \tag{10}$$

$$p(Y = 0|\mathbf{X}) = F_1(c_1 - d_1 X_1)\, F_2(c_2 - d_2 X_2)$$

The proof proceeds in two steps. In the first step, we demonstrate the one-to-one mapping of the choice criteria to the response probabilities. In the second step, we build upon the previous result to demonstrate the one-to-one mapping of the perceptual sensitivities to the response probabilities.

One-to-one mapping of the 2-ADC criteria to the response probabilities during catch trials

First, we demonstrate that there is a one-to-one mapping from the set of criteria, (c_1, c_2), to set of response probabilities during catch trials $p_0^i, i \in \{0, 1, 2\}$ (as mentioned before, we use p_0^i as a notational shorthand for $p(Y = i|X_i = 0, X_j = 0)$).

We consider the system of response probabilities when no stimulus was presented ($X_1 = X_2 = 0$), i.e. false-alarm rates at each location during catch trials:

$$p_0^1 = \int_{c_1}^{\infty} F_2(e - c_1 + c_2)\, f_1(e)\, de$$

$$p_0^2 = \int_{c_2}^{\infty} F_1(e - c_2 + c_1)\, f_2(e)\, de \tag{11}$$

We demonstrate that if set of criteria (c_1, c_2) produces a given set of response probabilities p_0^i, then it is the only set that can produce these probabilties. The analytical proof rests on the following lemmas:

Lemma 1 $p_0^i(c_i, c_j)$ is a monotonically decreasing function of c_i and a monotonically increasing function of c_j, $i, j \in \{1, 2\}, i \neq j$.

Lemma 2 p_0^0 is a monotonically increasing function of both c_1 and c_2. Specifically, $p_0^0 = F_1(c_1)\, F_2(c_2)$.

Simply put, these lemmas assert that response probabilities vary monotonically as a function of choice criteria. The proof of these lemmas is provided in Appendices A.1-A.2 (Supplemental Data). The proof assumes no specific form for the functions f_1 and f_2; only that they are continuous and supported over the entire domain of integration. Upon rearrangement of the identity in Lemma 2:

$$F_i(c_i) = p_0^0 / F_j(c_j) \qquad\qquad i, j \in \{1, 2\}, i \neq j \tag{12}$$

The sequence of arguments for the proof follows:

(i) Let (c_1, c_2) be a set of criteria that produces $\mathcal{P}_0^i$ of p_0^i. Assume that there exists another set (c_1', c_2') that also produces these same probabilities, such that at least one c_i' is distinct from c_i.

(ii) Without loss of generality, let $c_1 > c_1'$.

(iii) From Lemma 1,

$c_1 > c_1' \Rightarrow p_0^1(c_1, c_2) = \mathcal{P}_0^1 < p_0^1(c_1', c_2)$. Similarly,

$c_1 > c_1' \Rightarrow p_0^2(c_1, c_2) = \mathcal{P}_0^2 > p_0^2(c_1', c_2)$.

(iv) Hence, it follows that $c_2' > c_2$ for constant $\mathcal{P}_0^i$. In other words, if one choice criterion *increases*, the other must also *increase* to keep p_0^1 constant.

(v) Being cumulative distribution functions, F_i-s are monotonic functions of their arguments. Thus, $F(c_1) > F(c_1') \Leftrightarrow c_1 > c_1'$.

(vi) From Lemma 2, (equation 12), and point (v) above: $c_1 > c_1' \Rightarrow c_2' < c_2$ for constant $\mathcal{P}_0^0$. In other words, if one choice criterion *increases*, the other must *decrease* to keep p_0^0 constant.

The proof follows by contradiction.

One-to-one mapping of the 2-ADC sensitivities to the response probabilities during stimulus trials

Next, we demonstrate, based on the previous result, that there is a one-to-one mapping from the set of sensitivities, (d_1, d_2) to the set of response probabilities during stimulus trials $p_j^i, i \in \{0, 1, 2\}, j \in \{1, 2\}$.

For a stimulus presented at location i ($X_i = 1$), response probabilities at location i are given by (refer equation system 10):

$$p_1^1 = \int_{c_1 - d_1}^{\infty} F_2(e + d_1 - c_1 + c_2)\, f_1(e)\, de \tag{13}$$

$$p_2^2 = \int_{c_2 - d_2}^{\infty} F_1(e + d_2 - c_2 + c_1)\, f_2(e)\, de \tag{14}$$

The proof rests on the following lemma, which is proved in Appendix B.3 (Supplemental Data):

Lemma 3 $p_j^i(d_i, d_j)$ is a strictly monotonic function of its arguments $(d_i, d_j), i, j \in \{1, 2\}, i \neq j$.

The sequence of arguments for the proof follows:

(i) Based on the previous section, we have already established a one-to-one mapping from the set of criteria (c_1, c_2) to the response probabilities during catch trials. Thus, the criteria are fixed based on the proportion of false-alarms and correct rejections.

(ii) Given a particular (c_1, c_2), each of the probabilities, p_1^1 and p_2^2, in the above system of equations is only a function of its respective $d_i, i \in \{1, 2\}$.

(iii) By Lemma 3, p_i^i is a strictly monotonic function of its respective d_i.

(iv) In other words, there is a one-to-one mapping of the d_i-s to the respective p_i^i-s.

We have shown that given a set of choice criteria, there is a one-to-one mapping of the sensitivities to the response probabilities. In the previous section, we showed a similar one-to-one mapping of the choice criteria to the response probabilities. Thus, there is a one-to-one mapping from the set of parameters $\{d_i, c_i\}, i \in \{1, 2\}$ to a given set of response probabilities (given by model equations 3). This completes the proof.

A.2 Identifiability of of m-ADC model

In the previous section, we showed that the 2-ADC model is identifiable. In this section we extend this result to demonstrate that the m-ADC model is also identifiable. In other words, there is a one-to-one mapping of the model parameters (sensitivities, criteria) to the response probabilities in the m-ADC model.

As before, we rewrite equations 6 with the notation for the more general notation for the probability density and cumulative distribution functions (f_i and F_i, see Methods).

$$p(Y = i|\,\boldsymbol{\xi}) = \int_{c_i - d(\xi_i)}^{\infty} \prod_{k,k\neq i} F_k(e + d(\xi_i) - d(\xi_k) - c_i + c_k)\, f_i(e)\, de \qquad p(Y = 0|\,\boldsymbol{\xi}) = \prod_k F_j(c_k - d_k(\xi_k)) \qquad (15)$$

The demonstration proceeds in two steps. First we demonstrate the following with mathematical induction: if the m-ADC model (for a task with m response alternatives) has a one-to-one mapping of the criteria to the response probabilties, then so does the (m+1)-ADC model (for a task with $(m+1)$ alternatives). Next, we utilize monotonicity to show that there is a one-to-one mapping of the perceptual sensitivities to the response probabilities for the m-ADC model.

One-to-one mapping of the m-ADC criteria to the response probabilities during catch trials

We consider the probabilities of response during catch trials. This is given by setting $d(\xi_k) = 0 \forall k$ in equation system 15:

$$p_0^i = \int_{c_i}^{\infty} \prod_{k,k\neq i} F_k(e - c_i + c_k)\, f_i(e)\, de$$
$$i \in \{0, \ldots, m\} \qquad (16)$$

where $p_0^i = p(Y = i|\,\boldsymbol{\xi};\, \|\boldsymbol{d\xi}\|_1 = 0)$.

Statement Given a set of response probabilities $\mathcal{P}_0^i$ for an m-alternative model, and the ordered set of criteria $C = \{c_i : i \in \{0, \ldots, m\}\}$ that produce these probabilities (according to the system of equations 16 then there is no alternate set of criteria C^* that produces the same probabilities.

Basis There is a one-to-one mapping from the set of criteria $C = \{c_1, c_2\}$ to the response probabilities in a 2-alternative model (m=2).

Inductive step Let there be a one-to-one mapping from the set of criteria to the response probabilities in an *m-alternative* model, i.e., given a set of response probabilities $p_0^i = \mathcal{Q}^i$, there is one, and only one, set of m-ADC criteria $C_m = \{c_i : i \in \{0, \ldots, m\}\}$ that produces these probabilities (from the system of equations 16). Then, given a set of response probabilities $p_0^i = \mathcal{P}_0^i$ for an *m+1-alternative* model, and a set of criteria $C_{m+1} = \{c_j : j \in \{0, \ldots, m + 1\}\}$ that produces these probabilities there is no other set C_{m+1}^* that also produces the same probabilities, i.e., there is a one-to-one mapping from the set of criteria to the response probabilities in an *m+1-alternative* model.

Proof of basis In a previous section we proved the one-to-one mapping of the criteria to the 2-ADC response probabilities during catch trials (Appendix A.1, Supplemental Data). This constitutes the proof of the basis for $m = 2$.

Proof of inductive step The inductive step is proved, as before, in two stages:
The proof rests on the following lemmas, which are proved in Appendices A.4-A.6 (Supplemental Data).

Lemma 4 Given a set of response probabilities $p_0^r = \mathcal{P}_0^r, r \in \{0, \ldots, m+1\}$, and any set of criteria $C = \{c_j : j \in \{0, \ldots, m+1\}\}$ comprising an ordered set that produces these probabilities (according to equations 16). There is a one-to-one correspondence between any choice criterion c_i in C and its complement set $C_i' = \{c_j : j \in \{0, \ldots, m+1\}, j \neq i\}$ that excludes c_i.

Lemma 5 Given a set of response probabilities $\mathcal{P}_0^i, i \in \{1, \ldots, m+1\}$ and the set of all sets of criteria $\{C^k = \{c_j{}^k : j \in \{0, \ldots, m+1\}\}\}$ comprising ordered sets that produce these probabilties (according to equations 16). For any two sets C^1 and C^2, every pair of corresponding elements $(c_j{}^1, c_j{}^2)$ obeys the same order relation, i.e., if any $c_i{}^1 \gtrless c_i{}^2$ then every $c_j{}^1 \gtrless c_j{}^2, i, j \in \{0, \ldots, m+1\}, i \neq j$.

Lemma 6 Given a set of response probabilities $\mathcal{P}_0^0$ and the set of all sets of criteria $\{C^k = \{c_j{}^k : j \in \{0, \ldots, m+1\}\}\}$ comprising ordered sets that produce these probabilties (according to equations 16). For any two sets C^1 and C^2, at least one pair of corresponding elements $(c_j{}^1, c_j{}^2)$ differs in its order relation, i.e., if any $c_i{}^1 \gtrless c_i{}^2$ then at least one $c_j{}^1 \lessgtr c_j{}^2, i, j \in \{0, \ldots, m+1\}, i \neq j$.

Simply put, Lemma 4 states that given set of false-alarm and correct rejection rates, fixing one choice criterion determines all of the other choice criteria. The proof of Lemma 4 utilizes the induction hypothesis (see Appendix B.4, Supplemental Data). Lemma 5 states that if the choice criterion to one location were to increase (decrease), the choice criterion at every location has to also increase (decrease) to maintain the false-alarm rate unchanged at each location. Lemma 6 states that if the choice criterion to one location were to increase (decrease), the choice criterion at least at one location has to decrease (increase) to maintain the correct rejection rate unchanged.

The sequence of arguments for the proof proceeds as follows:

(i) Let $C = \{c_j : j \in \{0, \ldots, m+1\}\}$ be a set of criteria that produce a specific value of $p_0^i = \mathcal{P}_0^i$. Let $C' = \{c_j' : j \in \{0, \ldots, m+1\}\}$ be a different set that produces the same $\mathcal{P}_0^i$.

(ii) By Lemma 4, $c_j \neq c_j' \ \forall \ j$. Without loss of generality, let $c_i > c_i'$.

(iii) By Lemma 5, if $c_i > c_i'$, then $c_j > c_j' \ \forall \ j, j \neq i$.

(iv) By Lemma 6, if $c_i > c_i'$, then at least one $c_j < c_j'$ for some $j \neq i$.

The proof follows by contradiction. Thus, the set of criteria $C = \{c_j : j \in \{0, \ldots, m+1\}\}$, which produces $mathcal P_0^i$ (according to equations 16) is unique. In other words, there is a one-to-one mapping from the set of criteria to the response probabilities during catch trials (false-alarm rates and correct rejections).

One-to-one mapping of the m-ADC sensitivities to the response probabilities during stimulus trials

The proof rests on the following lemma (proved in Appendix B.7, Supplemental Data):

Lemma 7 The response probability $p_j^i(d(\xi_k))$ is a strictly monotonic function of $d(\xi_k)$.

The sequence of arguments proceeds as follows:

(i) By the task specification, no more than one stimulus is presented on a given trial. Thus, for a fixed set of criteria C, the response probabilities p_i^i of equation system 15 are simply a function of their respective perceptual sensitivities d_i.

(ii) From Lemma 7, the response probability $p_i^i(d(\xi_i))$ is a strictly monotonic function of its respective $d_i, i \in \{1, \ldots, M\}$.

(iii) Strict monotonicity implies a one-to-one mapping of the $d(\xi_i)$-s to the respective p_i^i-s.

This completes the proof. Note that the same arguments could be made with other sets of probabilities, such as the false-alarm rates, p_i^j, for reporting a stimulus at location j when a stimulus was presented at location i, which are also monotonic functions of $d(\xi_i)$ (Appendix B.7, Supplemental Data).

B Proof of lemmas on model identifiability

In this section, we demonstrate analytically various lemmas (employed in Appendix A) on model identifiablity.

B.1 Proof of Lemma 1: Monotonic variation of the probability of a NoGo response with choice criteria in 2-ADC catch trials

Assertion: $p(Y = 0|\|\mathbf{X}\|_1 = 0)$ *(or p_0^0) is a monotonically increasing function of both c_1 and c_2. Specifically, $p_0^0 = F_1(c_1)\, F_2(c_2)$.*

Proof: We compute the probability of a NoGo response during catch trials. This happens when Ψ falls below the criterion at both locations ($Y = 0, \text{iff}\, \Psi_1 \leq c_1 \,\cap\, \Psi_2 \leq c_2$). Thus,

$$p(Y = 0|\mathbf{X}) = p(\Psi_1 \leq c_1 \cap \Psi_2 \leq c_2) \tag{17}$$

Upon substitution of the structural model, and noting that the ε_i are independent, this gives:

$$
\begin{aligned}
p(Y = 0|\mathbf{X}) &= p(\varepsilon_1 \leq c_1 \cap \varepsilon_2 \leq c_2) \\
&= p(\varepsilon_1 \leq c_1)p(\varepsilon_2 \leq c_2) \\
&= F_1(c_1)\, F_2(c_2)
\end{aligned}
\tag{18}
$$

Thus, the probability of a correct rejection in the 2-ADC model factors into the product of the 1-ADC correct-rejection probabilities.

$$p_0^0 = F_1(c_1)\, F_2(c_2) \tag{19}$$

As the F_i-s are positive, and monotonically increasing functions of their arguments, p_0^0 is a monotonically increasing function of c_1 and c_2.

B.2 Proof of Lemma 2: Monotonic variation of the probabilities of Go responses with choice criteria in 2-ADC catch trials

Assertion: $p(Y = i|\|\mathbf{X}\|_1 = 0)$ *(or p_0^i) is a monotonically decreasing function of c_i and a monotonically increasing function of c_j.*

Proof: We reproduce equation system 11 here:

$$
\begin{aligned}
p_0^1 &= \int_{c_1}^{\infty} F_2(e + c_2 - c_1)\, f_1(e)\, de \tag{20} \\
p_0^2 &= \int_{c_2}^{\infty} F_1(e + c_1 - c_2)\, f_2(e)\, de \tag{21}
\end{aligned}
$$

With increasing c_1, p_0^1 has to decrease because:

(i) The integrand ($F_2(e + c_2 - c_1)$, specifically) decreases because F_2 is a monotonic function of its arguments ($e + c_2 - c_1$ decreases)

(ii) The domain of integration ($c_1 \to \infty$) decreases as c_1 increases (the integrand is never negative)

With increasing c_2, p_0^1 has to increase because the integrand increases ($F_2(e + c_2 - c_1)$, specifically), and the domain of integration is unaffected by c_2. The (converse) effects of c_1 and c_2 on p_0^2 can be similarly argued.

B.3 Proof of Lemma 3: Monotonic variation of response probabilities with perceptual sensitivities in 2-ADC stimulus trials

Assertion: $p(Y = i|X_i = 1)$ (or p_i^i) and $p(Y = i|X_j = 1)$ (or p_j^i) are both monotonic functions of d_i and $d_j (i, j \in \{1, 2\})$.

Proof: We reproduce part of equation system 10 here for reference.

$$p(Y = i|\mathbf{X}) = \int_{c_i - d_i X_i}^{\infty} F_j(e + d_i X_i - d_j X_j - c_i + c_j) \, f_i(e) \, de \tag{22}$$

$$i, j \in \{1, 2\}, i \neq j$$

where we have dropped the subscript from e_i (a variable of integration).

With increasing d_i, $p(Y = i|X_i = 1)$ or p_i^i has to increase because:

(i) The integrand ($F_j(e + d_i - c_i + c_j)$, specifically) increases because F_j is a monotonic function of its arguments (d_i increases)

(ii) The domain of integration ($c_i - d_i \to \infty$) increases as d_i decreases (the integrand is never negative)

With increasing d_j, $p(Y = i|X_j = 1)$ or p_j^i has to decrease because:

(i) The integrand ($F_j(e - d_j - c_i + c_j)$, specifically) increases because F_j is a monotonic function of its arguments ($-d_j$ decreases)

(ii) The domain of integration is unaffected by d_j.

This completes the proof.

B.4 Proof of Lemma 4: One-to-one correspondence of m-ADC choice criteria

Assertion: *Given a set of response probabilities $p_0^r = \mathcal{P}_0^r, r \in \{0, \ldots, m + 1\}$, and a set of criteria $C = \{c_j : j \in \{0, \ldots, m + 1\}\}$ that produces these probabilities (according to equation system 16). There is a one-to-one correspondence between any choice criterion c_i and its complement set $C_i' = \{c_j : j \in \{0, \ldots, m + 1\}, j \neq i\}$.*

Proof: The proof proceeds in two steps, first demonstrating the mapping $\zeta : C_i' \mapsto c_i$, and then its inverse $\zeta^{-1} : c_i \mapsto C_i'$.

First, consider the probability $p_0^0 = \mathcal{P}_0^0$. A given choice criterion $c_i, i \in \{0, \ldots, m+1\}$ can be expressed in terms of the remaining criteria in the following way.

$$\mathcal{P}_0^0 = \prod_{j=1}^{m+1} F_j(c_j) \tag{23}$$

$$c_i = F_i^{-1}\left(\frac{\mathcal{P}_0^0}{\prod_{j=1, j \neq i}^{m+1} F_j(c_j)}\right) \tag{24}$$

where F_i is invertible, being a cumulative distribution function. Given a particular $p_0^0 = \mathcal{P}_0^0$, and a set of m criteria $\{c_j : j \in \{0, \ldots, m + 1\}, j \neq i\}$ the remaining criterion c_i is uniquely determined, thus demonstrating the mapping $\phi : C_i' \mapsto c_i$.

Next, consider the set of probabilities $\mathcal{P}_0^i$. From system 16, these can be written as:

$$\mathcal{P}_0^i = \int_{c_i}^{\infty} \prod_{k=1, k \neq i}^{m+1} F_k(e - c_i + c_k) \, f_i(e) \, de \tag{25}$$

With the variable substitution $e' = e - c_i$, and following some algebra, this set of equations can be rewritten as:

$$\mathcal{P}_0^i = \int_0^\infty \prod_{k=1,k\neq i}^{m+1} F_k(e + c_k)\, f_i(e + c_i)\, de \tag{26}$$

Let C be a set of criteria $\{c_j : j \in \{0,\ldots,m+1\}\}$ that produces the probabilities on the left hand side of this equation. Let us assume that one of the criteria in this set, say c_{m+1} (without loss of generality) has a known value.

Define the following functions (for $i \in \{0,\ldots,m\}$).

$$F_\mu(e; c_{m+1}) = [F_{m+1}(e + c_{m+1})]^{\frac{1}{m}} \tag{27}$$

$$G_i(e + c_i; c_{m+1}) = F_i(e + c_i)\, F_\mu(e; c_{m+1}) \tag{28}$$

We note that both F_μ and G_i are parameterized by c_{m+1}. F_μ, the m-th root of a cumulative distribution function, and G, the product of F_μ and F_i are both monotonic, continuous functions, and it is easy to see that $\lim_{e\to-\infty} G_i = 0$; $\lim_{e\to+\infty} G_i = 1$. Thus, G_i is itself a cumulative distribution function with the following probability density:

$$g_i(e + c_i; c_{m+1}) = \frac{\partial G}{\partial e} \tag{29}$$

$$= F_i(e + c_i)\, \frac{\partial F_\mu(e; c_{m+1})}{\partial e} + F_\mu(e; c_{m+1})\, f_i(e + c_i) \tag{30}$$

Now, let us consider the following system of equations:

$$\mathcal{Q}^i = \int_0^\infty \prod_{k=1,k\neq i}^{m} G_k(e + c_k; c_{m+1})\, g_i(e + c_i; c_{m+1})\, de \tag{31}$$

With some algebra, we can show that $\mathcal{Q}^i = \mathcal{P}_0^i + (\mathcal{P}_0^{m+1}/m)$.

By the induction hypothesis for m-equations, given a set of q^i-s, and the parameter c_{m+1}, all of the c_k-s are uniquely determined. Because c_{m+1} was an arbitrarily chosen criterion, the result can be generalized as follows: given a set of $\mathcal{Q}^i$-s, and any choice criterion c_i, all of the other choice criteria in $C_i' = \{c_j : j \in \{0,\ldots,m+1\}, j \neq i\}$-s are uniquely determined, thus demonstrating the inverse mapping $\phi^{-1} : c_i \mapsto C_i'$.

Thus, for a given set of response probabilities $\mathcal{P}_0^i$ and a set of criteria C that produces these probabilities, we have shown a one-to-one correspondence among any one choice criterion, and the remaining criteria in the set $c_i \leftrightarrow C_i'$.

B.5 Proof of Lemma 5: Direct variation among all criteria in the m-ADC model

Assertion: *Given a set of response probabilities $\mathcal{P}_0^i, i \in \{1,\ldots,m+1\}$ and the set of all sets of criteria $\{C^k = \{c_j{}^k : j \in \{0,\ldots,m+1\}\}\}$ comprising ordered sets that produce these probabilties (according to equations 16). For any two sets of criteria C^1 and C^2, every pair of corresponding elements $(c_j{}^1, c_j{}^2)$ obeys the same order relation, i.e., if any $c_i{}^1 \gtreqless c_i{}^2$ then every $c_j{}^1 \gtreqless c_j{}^2, i, j \in \{0,\ldots,m+1\}$.*

Proof: Given set of response probabilities $\mathcal{P}_0^i$. Let $C^1 = \{c_j{}^1 : j \in \{0,\ldots,m+1\}\}$ be a set of criteria the produce these probabilities, and let $C^2 = \{c_j{}^2 : j \in \{0,\ldots,m+1\}\}$ be another, distinct (not identical) set that also produces the same probabilities. Also let all choice criteria from set C^1, except that corresponding to choice i (c_i), be greater (or lesser) in value than the corresponding criteria in set C^2. We demonstrate that in this case, the criterion c_i in set C^1 must also be greater (or lesser) in value than the corresponding criterion in set C^2.

The $\mathcal{P}_0^i$ are given by (refer equation 26):

$$\mathcal{P}_0^i = \int_0^\infty \prod_{j=1, j\neq i}^{m+1} F_j(e + c_j{}^1) \, f_i(e + c_i{}^1) \, de \tag{32}$$

$$= \int_0^\infty \prod_{j=1, j\neq i}^{m+1} F_j(e + c_j{}^2) \, f_i(e + c_i{}^2) \, de$$

Note that if $c_j{}^1 \gtrless c_j{}^2$,

$$\prod_{j=1, j\neq i}^{m+1} F_j(e + c_j{}^1) \gtrless \prod_{j=1, j\neq i}^{m+1} F_j(e + c_j{}^2) \quad \forall e \tag{33}$$

as the F_j-s are monotonically increasing functions of their arguments. Hence, for the right hand sides of equation 33 to be equal to each other (and each equal to $\mathcal{P}_0^i$) $c_i{}^1 \gtrless c_i{}^2$. The latter result is confirmed by inspecting the integrands of equation 33, and is also evident from the following lemma.

> **Lemma 8** The response probability p_0^i is a strictly (monotonically) increasing function of c_j and a strictly (monotonic) decreasing function of c_i.

The lemma is proved in a subsequent section. Thus, if every $c_j{}^1 \gtrless c_j{}^2, j \in \{0, \ldots m+1\}, j \neq i$ then, $c_i{}^1 \gtrless c_i{}^2$. However, we have just shown that there is a one-to-one correspondence between each c_i and its complement set $C_i' = \{c_j : j \in \{0, \ldots, m+1\}, j \neq i\}$. Thus, the converse statement must also hold: that is, if $c_i{}^1 \gtrless c_i{}^2$, then every $c_j{}^1 \gtrless c_j{}^2, j \in \{0, \ldots m\}, j \in \{0, \ldots, m+1\}, j \neq i$. This completes the proof.

B.6 Proof of Lemma 6: Inverse variation among at least a pair of criteria in the m-ADC model

Assertion: *Given a set of response probabilities $\mathcal{P}_0^0$ and the set of all sets of criteria $\{C^k = \{c_j{}^k : j \in \{0, \ldots, m+1\}\}\}$ comprising ordered sets that produce these probabilties (according to equations 16). For any two sets C^1 and C^2, at least one pair of corresponding elements $(c_j{}^1, c_j{}^2)$ differs in its order relation, i.e., if any $c_i{}^1 \gtrless c_i{}^2$ then at least one $c_j{}^1 \lessgtr c_j{}^2, i, j \in \{0, \ldots, m+1\}, i \neq j$.*

Proof: Given set of response probabilities $\mathcal{P}_0^i$. Let $C^1 = \{c_j{}^1 : j \in \{0, \ldots, m+1\}\}$ and $C^2 = \{c_j{}^2 : j \in \{0, \ldots, m+1\}\}$ be distinct sets of criteria that produce these probabilities (given by equation system 16). Also let any one choice criterion from set C^1, corresponding to choice i (c_i), be greater in value than the corresponding criterion in set C^2 i.e. $c_i{}^1 > c_i{}^2$.

We prove the result by contradiction. Assume that *none* of the other criteria in set C^1 is lesser than the corresponding criteria in set C^2. In other words, every $c_j{}^1 \geq c_j{}^2, j \in \{0, \ldots m+1\}$.

Given the probability of a NoGo response during catch trials, this can be written as:

$$\mathcal{P}_0^0 = \prod_{j=1}^{m+1} F_j(c_j{}^1) \qquad\qquad = \prod_{j=1}^{m+1} F_j(c_j{}^2) \tag{34}$$

The functions F_j are monotonic functions of their arguments. If every $c_j{}^1 \geq c_j{}^2$ equality of the right hand side expressions holds only if $c_j{}^1 = c_j{}^2$, which violates the assumption that C^1 and C^2 are non-identical sets. Thus, if any one $c_i{}^1 > c_i{}^2$, the assumption that *none* of the other criteria in set C^1 is lesser than the corresponding criteria in set C^2 leads to a contradiction. Hence, if any $c_i{}^1 > c_i{}^2, i \in \{0, \ldots, m+1\}$ then at least one criterion in set C_1 has to be lesser than the corresponding criterion in set C_2.

It is easy to see that the converse is also true, i.e. if any $c_i{}^1 < c_i{}^2, i \in \{0, \ldots, m+1\}$ then at least one criterion in set C_1 has to be greater than the corresponding criterion in set C_2. This completes the proof.

B.7 Proof of Lemma 7: Monotonic variation of m-ADC response probabilities with perceptual sensitivity

Assertion: The response probability p_j^i is a strictly monotonic (increasing) function of $d(\xi_i)$ and a strictly monotonic (decreasing) function of $d(\xi_j)$.

Proof: We reproduce the system of equations 15 for reference:

$$p(Y = i|\ \boldsymbol{\xi}) = \int_{c_i - d(\xi_i)}^{\infty} \prod_{k=1, k \neq i}^{m} F_k(e + d(\xi_i) - d(\xi_k) - c_i + c_k)\ f_i(e)\ de \tag{35}$$

Consider the probability of response to location i when the stimulus is presented at the same location ($d(\xi_k) = 0 \ \forall\ k \neq i$).

$$p_i^i = \int_{c_i - d(\xi_i)}^{\infty} \prod_{k=1, k \neq i}^{m} F_k(e + d(\xi_i) - c_i + c_k)\ f_i(e)\ de \tag{36}$$

With increasing $d(\xi_i)$, the response probability p_i^i has to increase, as the integrand increases with $d(\xi_i)$ (each F_k is a monotonically increasing function of its argument), and the integration (positive integrand) occurs over a larger domain ($c_i - d(\xi_i)$ decreases).

Next, consider the probability of response to location i when the stimulus is presented at location $j, j \neq i$ ($d(\xi_k) = 0 \ \forall\ k \neq j$).

$$p_j^i = \int_{c_i}^{\infty} \prod_{k=1, k \neq i,j}^{m} F_k(e - c_i + c_k) F_j(e - d(\xi_j) - c_i + c_j)\ f_i(e)\ de \tag{37}$$

Again, it is apparent that with increasing $d(\xi_j)$ the response probability p_j^i has to decrease, as the integrand ($F_j(e - d(\xi_j) - c_i + c_j)$, specifically) decreases with increasing $d(\xi_j)$ (the domain of integration is unaffected by $d(\xi_j)$).

This completes the proof.

B.8 Proof of Lemma 8: Monotonic variation of m-ADC response probabilities with choice criteria

Assertion: The response probability p_0^i is a strictly monotonic (decreasing) function of c_i and a strictly monotonic (increasing) function of c_j.

Proof: Consider the probability of response to location i when the no stimulus is presented ($X_k = 0 \forall k$).

$$p_0^i = \int_{c_i}^{\infty} \prod_{k=1, k \neq i}^{m} F_k(e - c_i + c_k)\ f_i(e)\ de \tag{38}$$

With increasing c_i, the response probability p_0^i has to decrease as the integrand decreases with c_i (each F_k a monotonically decreases with c_i), and the integration (positive integrand) occurs over a smaller domain. Similarly, with increasing c_j, p_0^i has to increase as the integrand ($F_j(e - c_i + c_j)$, specifically) increases with c_j.

This completes the proof.

C Concavity of the log-likelihood function

In this section, we identify key challenges with demonstrating the concavity of the log-likelihood function. While there are many ways to demonstrate concavity analytically, perhaps the most conceptually straightforward is the second derivative test.

We consider the multinomial log-likelihood function for the 2-ADC model.

$$\mathcal{L} = \sum_{i=0}^{2} \sum_{j=0}^{2} \mathcal{O}_j^i \log(p_j p_j^i) + \mathcal{C} \tag{39}$$

where $\mathcal{O}_j^i$ represents the observed number of responses to location i for the stimulus event j, p_j represents the prior probability of stimulus event j and p_j^i denotes the probability of response to location i for the stimulus event j (a conditional probability), and $\mathcal{C}$ represents an additive constant associated with the multinomial coefficient $\binom{N}{\mathcal{O}_j^i}$, where N is the total number of observations.

The prior probability of each stimulus event, p_j, is generally constant across an experimental session. With this assumption, the p_j terms can be factored out of $\mathcal{L}$, as they contribute only an additive term to $\mathcal{L}$ without affecting its shape. Thus, to prove convexity, let us consider a simplified function that excludes all of the additive terms:

$$\mathcal{L}' = \sum_{i=0}^{2} \sum_{j=0}^{2} \mathcal{O}_j^i \log(p_j^i) \tag{40}$$

In order to prove concavity of this function, we need to show that its Hessian is negative semidefinite (non-positive eigenvalues). Note that $\mathcal{L}'$ is a function of the four 2-ADC model parameters (d_1, d_2, c_1, c_2). Thus, its Hessian is given by:

$$\tilde{H} = \begin{pmatrix} \frac{\partial^2 \mathcal{L}'}{\partial d_1^2} & \frac{\partial^2 \mathcal{L}'}{\partial d_1 \, \partial d_2} & \frac{\partial^2 \mathcal{L}'}{\partial d_1 \, \partial c_1} & \frac{\partial^2 \mathcal{L}'}{\partial d_1 \, \partial c_2} \\ \frac{\partial^2 \mathcal{L}'}{\partial d_2 \, \partial d_1} & \frac{\partial^2 \mathcal{L}'}{\partial d_2^2} & \frac{\partial^2 \mathcal{L}'}{\partial d_2 \, \partial c_1} & \frac{\partial^2 \mathcal{L}'}{\partial d_2 \, \partial c_2} \\ \frac{\partial^2 \mathcal{L}'}{\partial c_1 \, \partial d_1} & \frac{\partial^2 \mathcal{L}'}{\partial c_1 \, \partial d_2} & \frac{\partial^2 \mathcal{L}'}{\partial c_1^2} & \frac{\partial^2 \mathcal{L}'}{\partial c_1 \, \partial c_2} \\ \frac{\partial^2 \mathcal{L}'}{\partial c_2 \, \partial d_1} & \frac{\partial^2 \mathcal{L}'}{\partial c_2 \, \partial d_2} & \frac{\partial^2 \mathcal{L}'}{\partial c_2 \, \partial c_1} & \frac{\partial^2 \mathcal{L}'}{\partial c_2^2} \end{pmatrix} \tag{41}$$

Each term of the Hessian may be represented as $\frac{\partial^2 \mathcal{L}'}{\partial \theta_1 \, \partial \theta_2}$, where θ_1 and θ_2 represent a two of the four parameters (not necessarily distinct). With some algebra, this generic term evaluates to:

$$\begin{aligned} \frac{\partial^2 \mathcal{L}'}{\partial \theta_1 \, \partial \theta_2} &= \sum_{i=0}^{2} \sum_{j=0}^{2} \mathcal{O}_j^i \frac{\partial}{\partial \theta_1} \left(\frac{1}{p_j^i} \frac{\partial p_j^i}{\partial \theta_2} \right) \\ &= \sum_{i=0}^{2} \sum_{j=0}^{2} \mathcal{O}_j^i \frac{1}{p_j^i} \left(\frac{\partial^2 p_j^i}{\partial \theta_1 \, \partial \theta_2} - \frac{1}{p_j^i} \frac{\partial p_j^i}{\partial \theta_1} \frac{\partial p_j^i}{\partial \theta_2} \right) \end{aligned} \tag{42}$$

These represent 10 distinct terms in the Hessian, since it is a symmetric matrix. However, certain simplifications are possible. For example, we can show that:

$$\frac{\partial p_j^i}{\partial d_k} = -\delta_{jk} X_k \frac{\partial p_j^i}{\partial c_k} \tag{43}$$

where δ_{jk} is the Kronecker delta function, $i, j \in \{0, 1, 2\}$ and $k \in \{1, 2\}$. In other words, the partial derivative of the response probabilities with respect to the sensitivities, are numerically equal to the partial derivatives with respect to the corresponding criteria (demonstrated analytically, below). This result is not surprising, per se, given that each response probability, p_j^i (given by equation 10) is invariant on the surface $c_i - d_i X_i = \text{const}$ and $c_j - d_j X_j = \text{const}$, i.e., increasing c_i or decreasing d_i by the same value (or vice

versa) causes $p_{\mathbf{X}}^i$ to change by the same amount (for a graphical intuition, see Figure 5C).

Based on these relations, we can define the second partial derivatives of the response probabilities as:

$$\frac{\partial^2 p_j^i}{\partial d_k \, \partial d_l} = \delta_{kl}\delta_{jk} X_k \, \frac{\partial^2 p_j^i}{\partial c_k \, \partial c_l}$$

$$\frac{\partial^2 p_j^i}{\partial d_k \, \partial c_l} = -\delta_{jk} X_k \, \frac{\partial^2 p_j^i}{\partial c_k \, \partial c_l} \tag{44}$$

From equations 42, 43 and 44, the Hessian is populated by second partial derivatives, of the form $\dfrac{\partial^2 p_j^i}{\partial c_1 \, \partial c_2}$, as well as products of the first partial derivatives, of the form $\dfrac{\partial p_j^i}{\partial c_1} \dfrac{\partial p_j^i}{\partial c_2}$. The product with the Kronecker delta functions and the X_j-s renders some of these terms zero.

We have previously shown (Appendix B.1-B.2) that the p_j^i in the 2-ADC model are monotonic functions of the criteria. Thus, the first partial derivatives $\dfrac{\partial p_j^i}{\partial c_k}$ are either always positive or negative for a given i, j, k. No such generalization can be made on the second partial derivatives (e.g., Figure S1A); these appear to vary with the actual value of c_1 and c_2.

We analytically evaluate the first order partial derivatives of p_j^i with respect to the four parameters (d_1, d_2, c_1, c_2), to verify if further simplification of the Hessian (equation 42) may be achieved.

We reproduce part of equation system 10 for the 2-ADC task here for reference.

$$p(Y = i|\mathbf{X}) = \int_{c_i - d_i X_i}^{\infty} F_j(e + d_i X_i - d_j X_j - c_i + c_j) \, f_i(e) \, de \tag{45}$$

$$i, j \in \{1, 2\}, i \neq j$$

We rewrite the above equations with the following transformation $e' = e - c_i + d_i X_i$; with this transformation, the criterion and sensitivity are eliminated from the limits of integration. The system may then be rewritten as:

$$p(Y = i|\mathbf{X}) = \int_0^{\infty} F_j(e' + c_j - d_j X_j) \, f_i(e' + c_i - d_i X_i) \, de' \tag{46}$$

$$i, j \in \{1, 2\}, i \neq j$$

Computing the partial derivative of system 46 with respect to d_i (and replacing the dummy variable of integration e' with e):

$$\begin{aligned}
\frac{\partial p(Y = i|\mathbf{X})}{\partial d_i} &= \frac{\partial}{\partial d_i}\left(\int_0^{\infty} F_j(e + c_j - d_j X_j) \, f_i(e + c_i - d_i X_i) \, de\right) \\
&= \int_0^{\infty} \frac{\partial}{\partial d_i}\left(F_j(e + c_j - d_j X_j) \, f_i(e + c_i - d_i X_i)\right) \, de \\
&= \int_0^{\infty} F_j(e + c_j - d_j X_j)\left(\frac{\partial f_i(e + c_i - d_i X_i)}{\partial d_i}\right) \, de
\end{aligned} \tag{47}$$

Integrating by parts, and noting that $\frac{\partial}{\partial d_i}(d_k X_k) = \delta_{ik} X_k$, we have:

$$
\begin{aligned}
\frac{\partial p(Y = i|\mathbf{X})}{\partial d_i} &= \left[F_j(e + c_j - d_j X_j) \frac{\partial F_i(e + c_i - d_i X_i)}{\partial d_i} \right]_0^\infty \\
&\quad - \int_0^\infty f_j(e + c_j - d_j X_j) \frac{\partial F_i(e + c_i - d_i X_i)}{\partial d_i} \, de \\
&= -F_j(c_j - d_j X_j) \, f_i(c_i - d_i X_i)(-X_i) \\
&\quad - \int_0^\infty f_j(e + c_j - d_j X_j) \, f_i(e + c_i - d_i X_i)(-X_i) \, de \\
&= X_i \{ F_j(c_j - d_j X_j) \, f_i(c_i - d_i X_i) \\
&\quad + \int_0^\infty f_j(e + c_j - d_j X_j) \, f_i(e + c_i - d_i X_i) \, de \}
\end{aligned}
\tag{48}
$$

Computing the partial derivative of system 46 with respect to d_j:

$$
\begin{aligned}
\frac{\partial p(Y = i|\mathbf{X})}{\partial d_j} &= \int_0^\infty \left(\frac{\partial F_j(e + c_j - d_j X_j)}{\partial d_j} \right) f_i(e + c_i - d_i X_i) \, de \\
&= \int_0^\infty f_j(e + c_j - d_j X_j)(-X_j) f_i(e + c_i - d_i X_i) \, de \\
&= -X_j \int_0^\infty f_j(e + c_j - d_j X_j) f_i(e + c_i - d_i X_i) \, de
\end{aligned}
\tag{49}
$$

Similarly, computing the partial derivative of system 46 with respect to c_i and c_j (noting that $\frac{\partial}{\partial c_i}(c_k) = \delta_{ik}$):

$$
\begin{aligned}
\frac{\partial p(Y = i|\mathbf{X})}{\partial c_i} &= -(F_j(c_j - d_j X_j) \, f_i(c_i - d_i X_i) \\
&\quad + \int_0^\infty f_j(e + c_j - d_j X_j) \, f_i(e + c_i - d_i X_i) \, de)
\end{aligned}
\tag{50}
$$

$$
\frac{\partial p(Y = i|\mathbf{X})}{\partial c_j} = \int_0^\infty f_j(e + c_j - d_j X_j) f_i(e + c_i - d_i X_i) \, de
\tag{51}
$$

From these equations, it is clear that:

$$
\frac{\partial p_{\mathbf{X}}^i}{\partial d_i} = -X_i \frac{\partial p_{\mathbf{X}}^i}{\partial c_i} \qquad\qquad\qquad \frac{\partial p_{\mathbf{X}}^i}{\partial d_j} = -X_j \frac{\partial p_{\mathbf{X}}^i}{\partial c_j}
\tag{52}
$$

where we have used the notation $p_{\mathbf{X}}^i$ for $p(Y = i|\mathbf{X})$. These equations are equivalent to equations 43, above.

These first partial derivatives do not appear to permit further simplification. Computing the second partial derivatives from equations 48 and 49 and, subsequently, demonstrating that the eigenvalues of the Hessian (equation 42) are non-positive, appears to require considerable further algebraic manipulation. These results highlight the challenges with demonstrating the concavity of the log-likelihood function.

D Proof of results on model optimality

In this section, we derive various results regarding optimal decision surfaces for the m-ADC model.

D.1 Optimal decision surfaces are hyperplanes of constant posterior odds ratio

We define the following relations (to be used in our derivation) for a stimulus detection task (e.g.,Figure 1A):

$$\lambda_l^k = C_l^k - C_l^0 \tag{53}$$

$$\mathcal{L}_{i0}(\boldsymbol{\Psi}) = \frac{\mathcal{N}_m(\boldsymbol{\Psi}|X_i = 1)}{\mathcal{N}_m(\boldsymbol{\Psi}|X_k = 0 \,\forall k)} \tag{54}$$

$$\Lambda_{i0}(\boldsymbol{\Psi}) = \frac{p_i}{p_0}\mathcal{L}_{i0} \tag{55}$$

where C_l^k represents the cost of responding to location k when a stimulus occurred at location l; C_l^0 represents the cost of giving a NoGo response, when a stimulus occurred at location l; λ_l^k represents the cost of responding to location k relative to giving a NoGo response, when a stimulus occurred at location l; $\mathcal{L}_{i0}(\boldsymbol{\Psi})$ is the likelihood ratio corresponding to a stimulus at location i relative to no stimulus, and $\Lambda_{i0}(\boldsymbol{\Psi})$ is the posterior odds ratio given by multiplying the likelihood ratio by the prior odds ratio of a stimulus at location i relative to no stimulus: $p_i/p_0 = p(X_j = 1)/p(X_i = 0 \,\forall\, i)$. We note that, for variables C and λ, non-zero subscripts refer to the location of stimulus, and non-zero superscripts to the location of response. A subscript of zero (e.g., C_0^k) denotes the no-stimulus event (catch trial), whereas a superscript of zero (e.g., C_k^0) denotes a NoGo response.

The general form of optimal decision surfaces for maximizing average utility (or minimizing average risk), for additive signals and noise, obey the following relations (Middleton & Meter, 1955, equations 16-17):

$$L_k(\boldsymbol{\Psi}) = L_l(\boldsymbol{\Psi}); \qquad\qquad L_k(\boldsymbol{\Psi}) = 0 \qquad\qquad \forall\, k,l \in \{1,\ldots,m\}, k \neq l$$

where:

$$L_k = \lambda_0^k + \sum_{i=1}^m \lambda_i^k \Lambda_i(\boldsymbol{\Psi}) \tag{56}$$

These hypersurfaces enclose $m + 1$ distinct decision domains corresponding to each of the m response alternatives and the NoGo response.

We can rewrite these equations as:

$$\lambda_0^k + \sum_{i=1}^m \lambda_i^k \Lambda_{i0}(\boldsymbol{\Psi}) = \lambda_0^l + \sum_{i=1}^m \lambda_i^l \Lambda_{i0}(\boldsymbol{\Psi}) \quad k \neq l \tag{57}$$

$$\lambda_0^k + \sum_{i=1}^m \lambda_i^k \Lambda_{i0}(\boldsymbol{\Psi}) = 0 \tag{58}$$

Substituting for the relative costs, λ, in terms of the absolute costs, C, yields:

$$C_0^k - C_0^0 + \sum_{i=1}^m (C_i^k - C_i^0)\Lambda_{i0}(\boldsymbol{\Psi}) = C_0^l - C_0^0 + \sum_{i=1}^m (C_i^l - C_i^0)\Lambda_{i0}(\boldsymbol{\Psi}) \quad k \neq l \tag{59}$$

$$C_0^k - C_0^0 + \sum_{i=1}^m (C_i^k - C_i^0)\Lambda_i(\boldsymbol{\Psi}) = 0 \tag{60}$$

These equations can be further simplified as:

$$C_0^k + \sum_{i=1}^{m} C_i^k \Lambda_{i0}(\mathbf{\Psi}) = C_0^l + \sum_{i=1}^{m} C_i^l \Lambda_{i0}(\mathbf{\Psi}) \tag{61}$$

$$C_0^k + \sum_{i=1}^{m} C_i^k \Lambda_{i0}(\mathbf{\Psi}) = C_0^0 + \sum_{i=1}^{m} C_i^0 \Lambda_{i0}(\mathbf{\Psi}) \tag{62}$$

Our assumption regarding costs (Methods) can be expressed analytically as:

$$C_j^k = C_j^l \; \forall j, k, l \in \{0, \dots, m\}; k \neq j, l \neq j \tag{63}$$

Incorporating this assumption, equation 61 simplifies to:

$$(C_k^k - C_k^l)\Lambda_{k0}(\mathbf{\Psi}) = (C_l^l - C_l^k)\Lambda_{l0}(\mathbf{\Psi}) \tag{64}$$

$$\frac{\Lambda_{k0}(\mathbf{\Psi})}{\Lambda_{l0}(\mathbf{\Psi})} = \frac{C_l^l - C_l^k}{C_k^k - C_k^l} \tag{65}$$

$$\Lambda kl(\mathbf{\Psi}) = \boldsymbol{\beta}_{kl} \tag{66}$$

where $\Lambda_{kl}(\mathbf{\Psi}) = \frac{\Lambda_{k0}(\mathbf{\Psi})}{\Lambda_{l0}(\mathbf{\Psi})} = \frac{p_k}{p_l} \frac{\mathcal{N}_m(\mathbf{\Psi}|X_k=1)}{\mathcal{N}_m(\mathbf{\Psi}|X_l=1)}$ is the posterior odds ratio of a stimulus at location k relative to a stimulus at location l and $\boldsymbol{\beta}_{kl} = (C_l^l - C_l^k)/(C_k^k - C_k^l)$.

Incorporating the assumption from equation 63, equation 62 simplifies to:

$$C_0^k + C_k^k \Lambda_{k0} = C_0^0 + C_k^0 \Lambda_{k0} \tag{67}$$

$$\Lambda_{k0} = \frac{C_0^0 - C_0^k}{C_k^k - C_k^0} \tag{68}$$

$$\Lambda_{k0} = \boldsymbol{\beta}_{k0} \tag{69}$$

Thus, optimal decision surfaces (equations 66 and 69) are the surfaces (hyperplanes) of constant posterior odds ratio (isosurfaces) for each pair of stimulus events (Λ_{kl}) and for each stimulus vs. the no-stimulus event (Λ_{k0}). The values of the constants ($\boldsymbol{\beta}_{kl}$ or $\boldsymbol{\beta}_{k0}$) that define the optimal decision surfaces for reporting a stimulus at a location k vs. one at another location l, depend on the relative costs (or benefits) of correctly reporting a stimulus (hit) at that location, k (or l), vs. incorrectly reporting (misidentification) the other location l (or k).

D.2 Optimal decision surfaces intersect at a point

Consider the optimal decision surfaces for detecting a stimulus at location i or location j vs. no stimulus. These are given by:

$$\Psi_i^* d_i = \log \boldsymbol{\beta}_{i0} - \log \frac{p_i}{p_0} + \frac{d_i^2}{2}$$

$$\Psi_j^* d_j = \log \boldsymbol{\beta}_{j0} - \log \frac{p_j}{p_0} + \frac{d_j^2}{2}$$

where Ψ_i^* and Ψ_j^* represent optimal values of Ψ_i and Ψ_j specified by equation 7.

Subtracting the two equations yields:

$$\Psi_i^* d_i - \Psi_j^* d_j = \log \boldsymbol{\beta}_{i0} - \log \boldsymbol{\beta}_{j0} - \log \frac{p_i}{p_0} + \log \frac{p_j}{p_0} + \frac{d_i^2}{2} - \frac{d_j^2}{2}$$

$$= \log \frac{\boldsymbol{\beta}_{i0}}{\boldsymbol{\beta}_{j0}} - \log \frac{p_i}{p_j} + \frac{d_i^2 - d_j^2}{2} \tag{70}$$

The ratio of $\boldsymbol{\beta}$-s term (first term on the right-hand-side of equation 70) can be expanded as:

$$\frac{\boldsymbol{\beta}_{i0}}{\boldsymbol{\beta}_{j0}} = \left(\frac{C_0^0 - C_0^i}{C_i^i - C_i^0} \right) \left(\frac{C_j^j - C_j^0}{C_0^0 - C_0^j} \right) \tag{71}$$

$$= \left(\frac{C_j^j - C_j^0}{C_i^i - C_i^0} \right) \left(\frac{C_0^0 - C_0^i}{C_0^0 - C_0^j} \right) \tag{72}$$

From equation 63, $C_0^i = C_0^j$, so that $(C_0^0 - C_0^i)/(C_0^0 - C_0^j) = 1$. In addition, from equation 63, $C_j^0 = C_j^i$ and $C_i^0 = C_i^j$. Thus, $(C_j^j - C_j^0)/(C_i^i - C_i^0) = (C_j^j - C_j^i)/(C_i^i - C_i^j) = \boldsymbol{\beta}_{ij}$.

Thus, equation 70 becomes:

$$\Psi_i^* d_i - \Psi_j^* d_j = \log \boldsymbol{\beta}_{ij} - \log \frac{p_i}{p_j} + \frac{d_i^2 - d_j^2}{2} \tag{73}$$

Notice that the right-hand-side of this equation is identical with the right-hand-side of equation 8. Thus, these optimal values of Ψ_i^* and Ψ_j^* for detecting a stimulus at location i or location j, respectively, in noise (equation 7) also lie on the optimal surfaces for reporting a stimulus at location i versus a stimulus at location j (equation 8). This completes the proof demonstrating that optimal decision surfaces defined by equation 7 and 8 intersect at a point.

E Model for a discrimination task with a NoGo response

In this section, we derive the model equations for two-alternative discrimination task that incorporates a NoGo response.

In the conventional 2-AFC discrimination task, the observer indicates, for example, how a test (target) stimulus differs from a standard stimulus (e.g., brighter versus dimmer, longer versus shorter, clockwise versus counterclockwise direction of rotation or movement).

On the other hand, in a ternary choice 2-ADC discrimination task, the observer must, in addition, indicate if she/he perceives the target stimulus to be the same as the standard by giving a NoGo response. For example, in a 2-ADC orientation discrimination task, the observer must not only indicate whether a target stimulus differs in orientation from a standard, with a Go/NoGo response, but must also indicate the direction of the difference as clockwise or counterclockwise (from the standard), with different Go responses.

We describe the 2-ADC discrimination model based on the length discrimination task of García-Pérez and Alcalá-Quintana (2011a) (see main text for description), although, the model is generally applicable to other discrimination tasks (like the orientation discrimination task just described) as well.

In this model, independent decision variables Ψ_A and Ψ_B encode sensory evidence for the stimulus above and the stimulus below, respectively. In the "L-configuration" the stimulus above is the test (vertical) stimulus, and the stimulus below is the standard (horizontal). In the "inverted-L" configuration, the stimulus above is the standard (horizontal) and the stimulus below is the test (vertical) stimulus.

Just as in the conventional 2-AFC design, the observer reports the longer stimulus (above or below) by comparing the perceived lengths of each (relative values of Ψ_A- and Ψ_B; biases in this decision are captured by the respective choice criteria (relative values of c_A and c_B). The key exception to this rule is that the observer gives a NoGo response if the perceived length of both stimuli are within a certain range of the point of subjective equality (gray region defined by the criteria c_A and c_B in Figure 7A).

Thus, the decision rule for the 2-ADC (discrimination) model is:

$$
\begin{aligned}
Y = 1 \ \text{if} \ &(\Psi_A < -c_B \ \cap \ \Psi_A - c_A > \Psi_B - c_B) \cup \\
&(-c_B \leq \Psi_A \leq c_A \ \cap \ \Psi_B < -c_A) \cup (\Psi_A > c_A \ \cap \ \Psi_A - c_A > \Psi_B - c_B) \\
Y = 2 \ \text{if} \ &(\Psi_B < -c_A \ \cap \ \Psi_B - c_B > \Psi_A - c_A) \cup \\
&(-c_A \leq \Psi_B \leq c_B \ \cap \ \Psi_A < -c_B) \cup (\Psi_B > c_B \ \cap \ \Psi_B - c_B > \Psi_A - c_A) \\
Y = 0 \ \text{if} \ &-c_B \leq \Psi_A \leq c_A \ \cap \ -c_A \leq \Psi_B \leq c_B
\end{aligned}
\tag{74}
$$

where the designations $Y = 0, 1, 2$ correspond to the NoGo, A>B (above > below) and B>A (below > above) responses, respectively, in Figure 7A. Such a decision rule implies that observers have internalized the point of subjective equality of the test stimulus to the standard, a plausible assumption when the standard stimulus remains fixed throughout the experiment, and is well-known to the observers beforehand (the length of the standard stimulus was fixed at 104 pixels in this task, and the authors, who were also the observers, were presumably familiar with the standard stimulus).

In this model, the psychophysical function (perceived length) is a linear function of stimulus strength (physical length). Thus, the sensitivity (d_z) is linearly related to the physical length (x) of the stimulus as: $d_z(x) = \beta_z x$, where $z = \{s, t\}$ represent, respectively the standard (horizontal) and test (vertical) stimuli. The point of subjective equality (PSE, origin of the coordinate axes), is the physical length of the test stimulus at which its perceived length becomes equal to that of the standard stimulus. Thus, $d_s(x_s) = d_t(PSE)$, where x_s is the length of the standard stimulus (104 pixels). Hence, $\beta_s x_s = \beta_t PSE$ or $PSE = \beta_s x_s / \beta_t$.

The structural model is conceptually identical with that of the 2-ADC detection model.

$$
\Psi_A = d(\xi_A) + \varepsilon_A \qquad\qquad \Psi_B = d(\xi_B) \, X_B + \varepsilon_B
\tag{75}
$$

where $d(\xi_A)$ and $d(\xi_B)$ represent the difference between the perceived length of the stimulus above and below, respectively, from the perceived length of the standard ($d_s(x_s)$) or, equivalently, at the point of subjective equality ($d_t(PSE)$). Thus, when the (vertical) test stimulus is presented *above* the (horizontal) standard (L configuration), $\xi_A = x_t, \xi_B = x_s$, $d(\xi_A) = d_t(x_t) - d_t(PSE) = d_t(x_t) - d_s(x_s)$ and $d(\xi_B) = d_s(x_s) - d_s(x_s) = 0$ (because the horizontal/standard is always presented at its standard length). Similarly, when the (vertical) test stimulus is presented *below* the (horizontal) standard (inverted-L configuration), $\xi_A = x_s, \xi_B = x_t$, $d(\xi_A) = 0$, and $d(\xi_B) = d_t(x_t) - d_s(x_s)$. For conciseness, we introduce the notation: $\Delta d = d_t(x_t) - d_s(x_s)$.

The model may be described, equivalently, as follows: The joint distribution of Ψ_A and Ψ_B are represented by a bivariate Gaussian random variable $\mathbf{\Psi} = [\Psi_A, \Psi_B]$ with mean $\mathbf{\Delta d}$ and identity covariance matrix. The magnitude of $\mathbf{\Delta d}$ varies with the length of the test stimulus, x_t, as $|\mathbf{\Delta d}| = \Delta d$; the change occurs along the x-axis when the the test stimulus is presented above the standard (L), or along the y-axis, when the test stimulus is presented below the standard (inverted-L). Thus, $\mathbf{\Delta d} = \mathbf{0}$ when $x_t = PSE$.

The response probabilities in this task may be computed by inspection of Figure 7A. For example, the probability of the response $Y = 1$ (A>B) is the integral of the distribution of $\mathbf{\Psi}$ over the red region: the latter is all of the region below the oblique line $\Psi_B - \Psi_A = c_B - c_A$, except for the area overlapping the NoGo response (gray) region. Similarly, the probability of the response $Y = 2$ (B>A) is the integral over the blue region, which is all of the region above the oblique line, except for the area overlapping the NoGo response (gray) region. The probability of a NoGo response ($Y = 0$) is simply the integral over the gray (rectangular) region.

$$p(Y = 1|\mathbf{\Delta d}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\Psi_A + c_B - c_A} \varphi(\mathbf{\Psi}; \mathbf{\Delta d})\, d\Psi_A d\Psi_B - \int_{-c_B}^{c_A} \int_{-c_A}^{\Psi_A + c_B - c_A} \varphi(\mathbf{\Psi}; \mathbf{\Delta d})\, d\Psi_A d\Psi_B$$

$$p(Y = 2|\mathbf{\Delta d}) = \int_{-\infty}^{\infty} \int_{\Psi_A + c_B - c_A}^{\infty} \varphi(\mathbf{\Psi}; \mathbf{\Delta d})\, d\Psi_A d\Psi_B - \int_{-c_B}^{c_A} \int_{\Psi_A + c_B - c_A}^{c_B} \varphi(\mathbf{\Psi}; \mathbf{\Delta d})\, d\Psi_A d\Psi_B \tag{76}$$

$$p(Y = 0|\mathbf{\Delta d}) = \int_{-c_A}^{c_B} \int_{-c_B}^{c_A} \varphi(\mathbf{\Psi}; \mathbf{\Delta d})\, d\Psi_A d\Psi_B \tag{77}$$

where we have used the notation φ to represent the bivariate normal distribution of the decision variable $\mathbf{\Psi}$.

These equations may be readily modified for the 2-ADCX task, which incorporates an interaction among Ψ_A and Ψ_B.

In this case, the mean of $\mathbf{\Psi}$ varies with the length of the test stimulus as: $\mathbf{\Delta d_X} = [\Delta d, \eta\, \Delta d]$, when the test stimulus is above the standard (L configuration) or $= [\eta\Delta d,\ \Delta d]$ when the test stimulus is below the standard (inverted-L configuration), where, as before $\Delta d = |\mathbf{\Delta d}| = \beta_t x_t - \beta_s x_s$, and η is the parameter that captures the interaction. η is numerically equal to $\arctan(\alpha)$, α being the angle in the graphical illustration in Figure 7C, and reported in Table 1. Incorporating this value of $\mathbf{\Delta d_X}$ into equations 76 gives the response probabilities in the 2-ADCX model.

There is considerable scope for future work, including extending the model to the multialternative case, and demonstrating optimality and identifiability of the parameters. In addition, the assumption regarding the subjects internalizing the point of subjective equality works well for discrimination tasks with a well-known, fixed standard (as demonstrated by model fits), but would need to be amended for more general tasks involving roving designs with interleaved standards of different strengths.

F The m-AFC model with bias as a special case of the m-ADC model

In this section, we demonstrate that the m-AFC model with bias is a special case of the m-ADC model.

In the m-ADC model, if the decision variable never falls below the criterion at any location, the observer never provides a NoGo response. This can be achieved by setting the criteria to very low (large negative) values. In this case the m-ADC model is identical with an m-AFC model.

We denote as b_i, the difference $c_m - c_i$, which we term the bias for location i relative to location m; by this definition, $b_m = 0$. Introducing these terms into equation system 15:

$$p(Y = i|\boldsymbol{\xi}) \quad = \quad \int_{c_i - d(\xi_i)}^{\infty} \prod_{k, k \neq i} F_k(e + d(\xi_i) - d(\xi_k) + b_i - b_k) \, f_i(e) \, de \tag{78}$$

Formally, the m-ADC model reduces to the m-AFC model as the criteria are reduced to very low values ($c_i \to -\infty$), while keeping b_i constant. Applying this limit to the above equation:

$$p(Y = i|\boldsymbol{\xi}) \quad = \quad \int_{-\infty}^{\infty} \prod_{k, k \neq i} F_k(e + d(\xi_i) - d(\xi_k) + b_i - b_k) \, f_i(e) \, de \tag{79}$$

and $p(Y = 0|\boldsymbol{\xi}) = \lim_{c_k \to -\infty} F_k(c_k - d(\xi_k)) = 0$, or, the probability of a NoGo response is zero.

These equations describe a recently developed m-AFC model formulation that incorporates bias (DeCarlo, 2012). Thus, the m-ADC model is a more general form of the m-AFC model.

Notice that such a model describes the behavior of an ideal observer (one who seeks to maximize success), when no catch trials are incorporated into the task design. In this case, the prior probability of a catch trial is zero ($p_0 = 0$) and according to equation 9, $\lim_{p_0 \to 0} \Lambda_{j0} \to -\infty$, so that $c_j \to -\infty$, whereas Λ_{ij}, which does not depend on p_0, remains unchanged, as do the differences $c_i - c_j$ (and, hence, the b_i-s). Thus, an ideal observer's behavior switches naturally from an m-ADC model to an m-AFC model when catch trials are excluded from the task design.

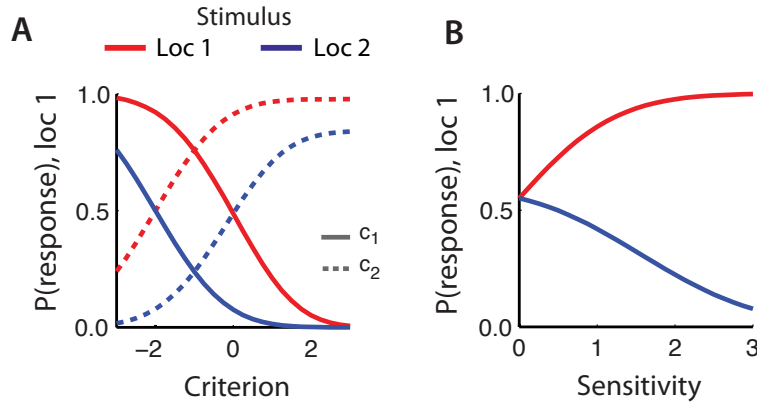# Supplemental Data: Figures and Tables

Figure S1: **Effect of varying sensitivities and criteria on 2-ADC response probabilities** (A) Variation of the probability of response at location 1 with the criterion at each location (for constant sensitivities, Table S2A). The probability of response to location 1, for a stimulus presented at location 1, decreases monotonically with an increasing choice criterion (c_1) at location 1 (solid red line) and increases monotonically with an increasing choice criterion (c_2) at location 2 (dashed red line). The same monotonic trends are observed when a stimulus is presented at location 2 (blue curves). (**B**) Variation of the probability of response at location 1 with the sensitivity at each location (for constant criteria, Table S2A). The probability of response to location 1 increases monotonically with increasing sensitivity (d_1) to a stimulus at location 1 (red), and decreases monotonically with increasing sensitivity (d_2) to a stimulus at location 2 (blue).
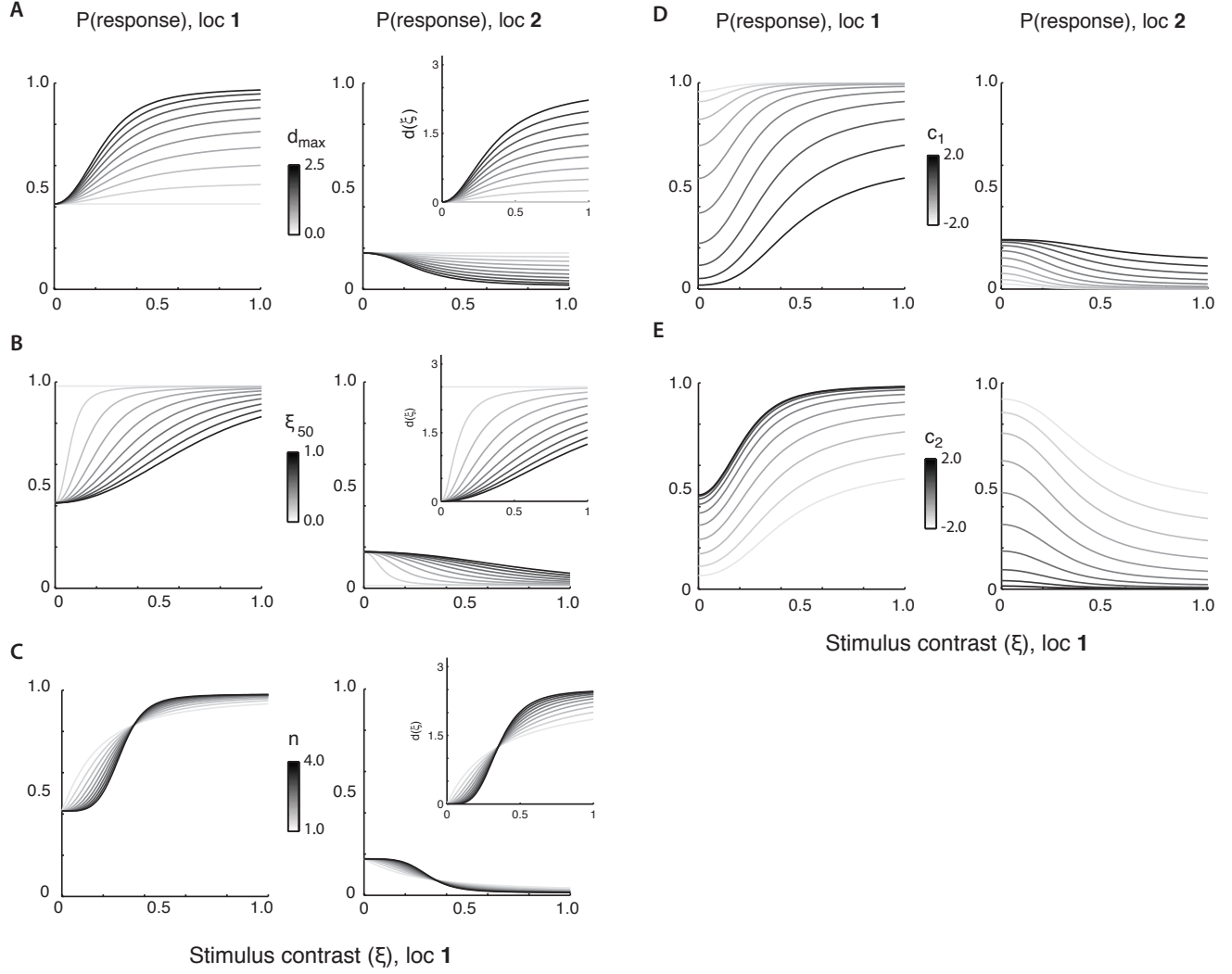
Figure S2: **Effect of varying psychophysical parameters and criteria on 2-ADC psychometric functions** (**A**) (*Left*) Psychometric functions $p(\xi)$ at location 1 as a function of stimulus contrast ξ at location 1. The family of curves (light gray to black) correspond to increasing values of asymptotic sensitivity d_{max} at location 1. (*Right*) Same as in left panel, but psychometric functions at location 2 as a function of stimulus contrast at location 1. (*Inset*) Psychophysical functions ($d(\xi)$) for increasing d_{max} (scale parameter). (**B**) Same as in (A), but psychometric functions for increasing values of half-max contrast ξ_{50} (shift parameter). (**C**) Same as in (A), but psychometric functions for increasing values of the exponent n (slope parameter). (**D**) Same as in (A), but psychometric functions for increasing values of the criterion at location 1, c_1. (**E**) Same as in (A), but psychometric functions for increasing values of the criterion at location 2, c_2.
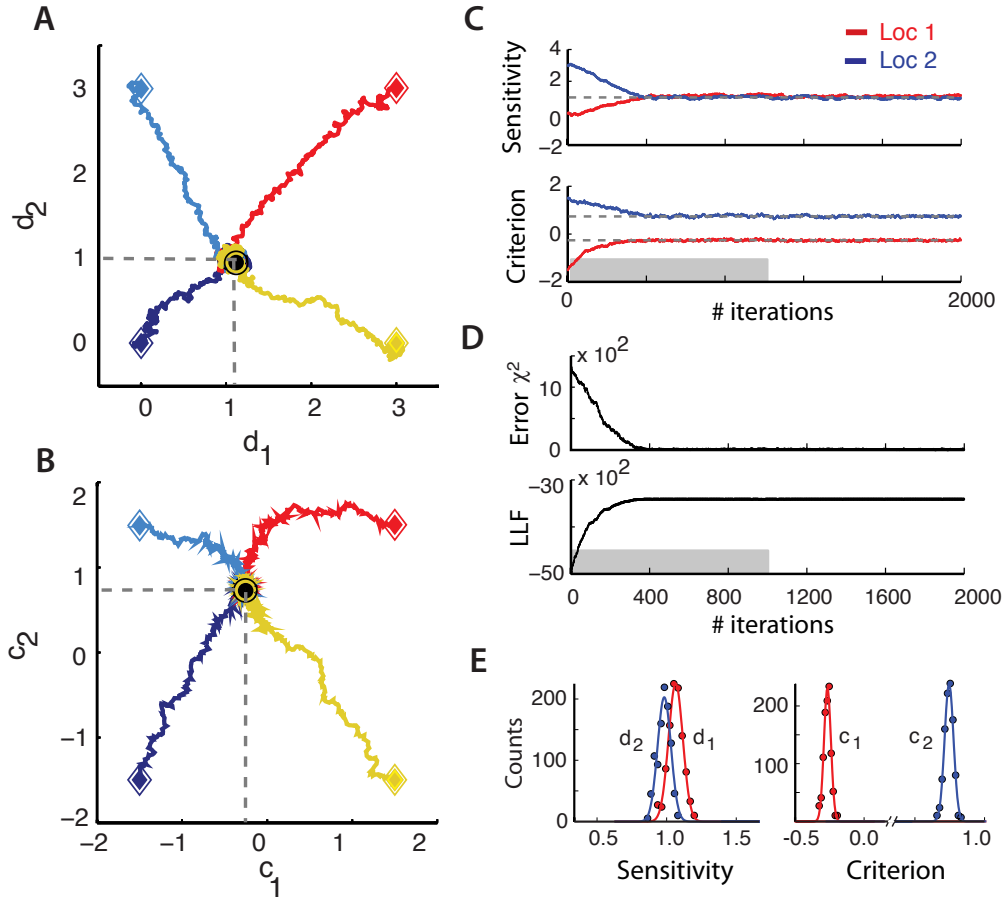
Figure S3: **Bayesian estimation of model parameters with the Markov-Chain Monte Carlo algorithm** (**A-B**) Markov-chain Monte-Carlo (MCMC, Metropolis sampling) algorithm for estimating perceptual sensitivity (**A**) and choice criterion (**B**) at each location from simulated response counts in the two-alternative detection task (Table S2B). For various initial guesses (colored diamonds-s), the Markov chain converged reliably to identical values of sensitivity and criterion at each location (black circles). Colored lines: Markov chains during MCMC runs for different initial guesses. (**C**) Evolution of the values of sensitivity (upper panel) or criterion (lower panel) at each location during a particular MCMC run (magenta data in panels C-D) for location 1 (red) or location 2 (blue). Gray bar: burn-in period (1000 iterations). (**D**) The chi-squared error function (upper panel) decreased steadily, and the log-likelihood increased (lower panel) over successive iterations of the MCMC run. (**E**) Stationary (posterior) distributions (circles) of the sensitivity (left panel) and criterion (right panel) values at each location for the MCMC run (panel E). These distributions were used to construct standard errors and 95% credible intervals for the parameters (Table S2C). Red data: location 1; blue data: location 2. Lines: Gaussian fits to each distribution.
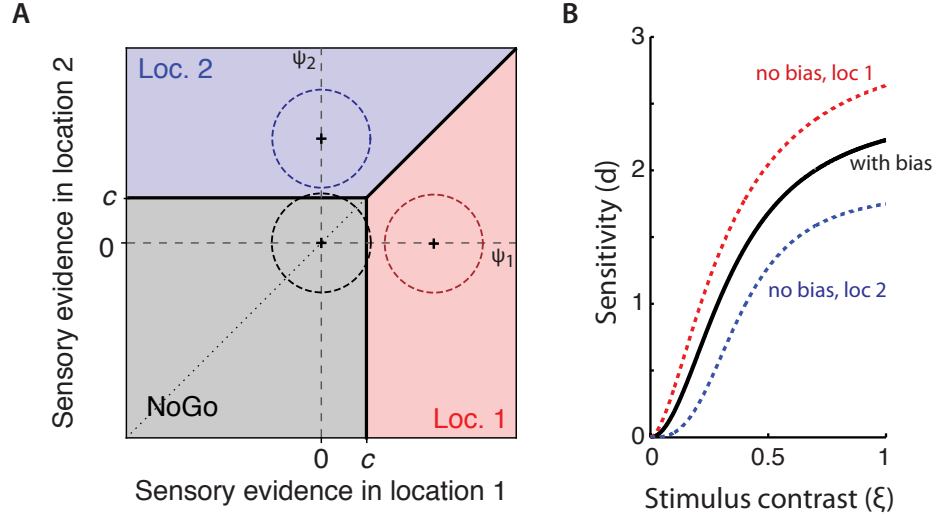
Figure S4: **Fitting a 2-ADC model with a symmetric decision rule** (**A**) Schematic 2-ADC model, with a symmetric decision rule, that does not incorporate bias ($c_1 = c_2 = c$). (**B**) Maximum likelihood estimates of the psychophysical function of stimulus contrast, $d(\xi)$. Black curve: Psychophysical function estimated with a 2-ADC model that incorporated bias. Red and blue dashed curves: Psychophysical functions at locations 1 and 2, respectively, estimated with a model that did not incorporate bias.

Table S1. Stimulus-response contingency table for 2-AFC and 2-ADC tasks.

A. 2x2 stimulus-response contingency table for a 2-AFC (Yes/No) task.

Response / Stimulus	Go response @ Loc 1	NoGo response
Stimulus @ Loc 1	Hit (HR)	Miss (MR)
Catch (No stimulus)	False-alarm (FA)	Correct rejection (CR)

B. 3x3 stimulus-response contingency table for a 2-ADC task.

Response / Stimulus	Go response @ Loc 1	Go response @ Loc 2	NoGo response
Stimulus @ Loc 1	Hit (HR$_1$)	*Misidentification (incorrect)*	Miss (MR$_1$)
Stimulus @ Loc 2	*Misidentification (incorrect)*	Hit (HR$_2$)	Miss (MR$_2$)
Catch (No stimulus)	False-alarm (FA$_1$)	False-alarm (FA$_2$)	Correct rejection (CR)

In 2-AFC (Yes/No) tasks there is only one false-alarm: a Go response during catch trials (FA). In addition to this, another type of false-alarm response can occur in 2-ADC tasks: a Go response at a location when a stimulus was presented at the opposite location (gray shaded cells).

Table S2. **Simulated parameter recovery with MLE and MCMC.**

A. Parameters used in the simulation.

Parameter	Location 1	Location 2
sensitivity	$d_1 = 1.0$	$d_2 = 1.0$
criterion	$c_1 = -0.25$	$c_2 = 0.75$
noise	$\varepsilon_1 = N(0, 1)$	$\varepsilon_2 = N(0, 1)$
stimulus prior probability	$p_1 = 0.25$	$p_2 = 0.25$

B. Simulated contingency table of response counts (N = 4000 trials from 20 simulated runs).

Response / Stimulus	Go response @ Loc 1	Go response @ Loc 2	NoGo response
Stimulus @ Loc 1	871	*66*	63
Stimulus @ Loc 2	*422*	414	164
Catch (No stimulus)	1122	263	615

C. Sensitivities and criteria recovered with maximum likelihood (MLE) and Bayesian (Markov Chain Monte Carlo) estimation procedures.

Parameter	MLE (mean ± SE)	Bayesian (mean ± SE)	95% CI
sensitivity	$d_1 = 1.07 \pm 0.06$	$d_1 = 1.07 \pm 0.08$	$d_1: 0.98 - 1.16$
	$d_2 = 0.99 \pm 0.06$	$d_2 = 0.98 \pm 0.08$	$d_2: 0.89 - 1.08$
criterion	$c_1 = -0.27 \pm 0.03$	$c_1 = -0.27 \pm 0.03$	$c_1: -0.30 - -0.21$
	$c_2 = 0.75 \pm 0.04$	$c_2 = 0.75 \pm 0.05$	$c_2: 0.67 - 0.81$

SE: standard error, CI: credible intervals

Table S3. **Maximum likelihood (ML) estimates of the psychometric function with and without accounting for bias.**

A. Parameters used in the simulation of a 2-ADC model with bias ($c_1 \neq c_2$).

Parameter		Location 1		Location 2
sensitivity	$\mathbf{d_1(\xi_1)}$	$d_{max} = 2.5$ $n = 2.0$ $c_{50} = 0.35$	$\mathbf{d_2(\xi_2)}$	$d_{max} = 2.5$ $n = 2.0$ $c_{50} = 0.35$
criterion		$\mathbf{c_1} = 0.1$		$\mathbf{c_2} = 0.7$
noise		$\varepsilon_1 = N(0, 1)$		$\varepsilon_2 = N(0, 1)$
stimulus prior probability		$p_1 = 0.25$		$p_2 = 0.25$

B. ML estimates of 2-ADC psychometric parameters with and without accounting for bias.

Parameter		MLE <u>with bias</u> (mean ± SE)		MLE <u>without bias</u> (mean ± SE)
sensitivity	$\mathbf{d_1(\xi_1)}$	$d_{max} = 2.48 \pm 0.03$ $n = 2.01 \pm 0.05$ $c_{50} = 0.34 \pm 0.006$	$\mathbf{d_1(\xi_1)}$	$d_{max} = 3.03 \pm 0.05$ $n = 1.69 \pm 0.04$ $c_{50} = 0.33 \pm 0.008$
	$\mathbf{d_2(\xi_2)}$	$d_{max} = 2.49 \pm 0.03$ $n = 2.03 \pm 0.04$ $c_{50} = 0.35 \pm 0.005$	$\mathbf{d_2(\xi_2)}$	$d_{max} = 2.49 \pm 0.03$ $n = 2.95 \pm 0.07$ $c_{50} = 0.38 \pm 0.004$
criterion		$\mathbf{c_1} = 0.10 \pm 0.002$ $\mathbf{c_2} = 0.71 \pm 0.003$		$\mathbf{c_1} = 0.35 \pm 0.002$ $\mathbf{c_2} = 0.35 \pm 0.002$