Supporting Information

Sridharan et al. 10.1073/pnas.1316824111

SI Methods

Stimulus Presentation and Data Acquisition. Stimulus presentation and data acquisition were automated by custom Matlab (MathWorks) scripts using the Psychophysics Toolbox extensions (1). Stimuli were presented on a touch-sensitive computer screen, which also recorded the bird's peck responses. During testing, birds were placed inside a custom apparatus (2) in a sitting position facing the screen. The bird's torso was wrapped in a cloth sleeve with the torso's midsagittal axis aligned with the center of the screen and orthogonal to the screen. To begin a trial, a small (±5 mm) zeroing cross appeared at the center of the screen. The chicken had to peck accurately on the cross (within 10 mm) to initiate a trial. This action forced the head and eyes into a standard position relative to the screen for the next ∼150 to 200 ms (e.g., Fig. S3*A*) (3). The stimulus array was presented on the screen for 50 ms immediately following a peck on the cross. The array disappeared before the chicken had a chance to move from the standard position; the earliest movement onsets occurred at latencies of ∼200 ms (Fig. S3 *B* and *C*, reaction times for highest relative target strengths).

Pecking includes a stereotyped action pattern of binocular, frontal fixation during which the eyes assume a standard position in the head (3, 4). Therefore, following a peck, the orientation of the eyes relative to the stimulus can be inferred with high reliability from the position (location and orientation) of the head relative to the stimulus. By monitoring head orientation in our experiments, stimuli could be presented at consistent, defined locations in the bird's visual field. Head position and orientation were monitored in real time with submillimeter spatial and high temporal (120 Hz) resolution with an infrared-based tracking system (Natural Point; OptiTrack Systems). In addition, the timing and location of each peck on the touch screen, as well as details regarding the stimulus configuration, were stored offline for post-hoc analyses.

Training and Testing. Training was accomplished in four stages. In the first stage, the birds were trained to peck at the location of a briefly flashed target. In the second stage, response boxes were introduced on the side of the target and the bird was rewarded for pecking at the location of the response box nearest the target. In the third stage, the distracter was introduced on the side opposite the target, and birds were rewarded for pecking on the response box nearest the target (stimulus on that side) while ignoring the distracter (stimulus on the opposite side) and were punished with a time-out (10–15 s) for pecking on the side of the distracter. In the final stage, cues were added to the task. Birds were punished (time-outs) for pecking at the location of the cue before the appearance of the target. During the testing phase, a peck at any location on the screen (other than the zeroing cross) before the target was presented, terminated the trial. A movie of a well-trained bird (bird 1) performing the task is provided in Movie S1.

For each bird on each day, data were acquired in two successive sessions of ∼100 trials each, with a brief (10-min) break in between. The two test sessions were preceded by a warm-up session (∼50 trials) to verify that the bird was engaged in the task, and to permit behavior to stabilize before beginning data collection. Data from these warm-up sessions were excluded from the analysis. For the cued–uncued measurements, data were acquired from three birds (n = 199 experimental sessions). For the validly–invalidly cued measurements, data were acquired from these same three birds plus a fourth bird (n = 211 experimental sessions). During the testing phase, pecks to the incorrect side (side opposite the response boxes) almost never occurred; when these occurred the corresponding trials were excluded from analysis. Overall, such pecks occurred on less than 0.1% of all trials; the proportions were similar for each bird for the cued-uncued and validly-invalidly cued tasks.

Analysis of Percent Correct (Hit Rate) and Response Times. We analyzed the data first, as if they were acquired with a two-alternative forced choice (2-AFC) design (choice of upper vs. lower box). Percent correct (hit rate) was computed as the proportion of correctly localized targets as a fraction of the total number of correctly and incorrectly localized targets (Go response). For these analyses, peck responses above the horizon were scored as signifying selection of the upper target, and responses below the horizon as signifying selection of the lower target. Excluding rare, ambiguous responses (<±5 mm from the horizon) from the analysis provided similar results.

Percent correct was computed individually for each target and distracter contrast (9 × 9 matrix in Fig. 1 *B–E*) after combining data across the left and right hemifields (except for head position analysis where each hemifield was treated separately; Fig. S6) and across experimental sessions. Percent correct was plotted as a function of relative target strength by pooling data from trials with equivalent target-to-distracter contrast ratios (the diagonally aligned entries in the performance matrix, orthogonal to the arrow; Fig. 1 *B–E*).

Response times were computed as the time from stimulus array presentation until the peck on the response box, as acquired from the touch screen. Only correct trials (hits) were included in computing response times. In addition, we also computed the reaction time, based on the time from stimulus array presentation until the initiation of a head orienting response (Fig. S3). These orienting movements were apparent in the lateral (azimuthal) direction (left vs. right), and the onsets were detected as an abrupt change in slope by a custom temporal edge-detection algorithm implemented in Matlab; representative onsets detected by the algorithm are shown in Fig. S3*A* (arrows).

Curve Fits for Percent Correct and Response Time Data. Percent correct data as a function of relative target strength were fit with a cumulative Gaussian function of the form

$$p = p_{\max} F\left(m\left(T_{rel} - T_{rel\text{-}50}\right)\right)$$

where p represents the percent correct, T_{rel} represents the relative target strength, i.e, the target-to-distracter contrast ratio (log units), F the cumulative Gaussian function, $p_{\max}$ the asymptotic performance, m a slope parameter, and $T_{rel\text{-}50}$ the relative target strength at which p reaches half of its maximum value. In addition, we defined the parameter r_{50} as the value of T_{rel} at which performance reached 50% of its maximal range across all values of tested relative target strengths. All fits were performed after subtracting baseline chance performance (50%).

Response and reaction time data were fit with a three-parameter power law function of the form

$$MRT = \gamma + \alpha c^{-\beta}$$

where MRT represents the mean response (or reaction) time, c the stimulus contrast, γ an offset parameter, and α and β the scaling and exponent (slope) parameters, respectively. Following

a logarithmic transformation, β represents the slope of a linear fit of the response time as a function of contrast [$\log(MRT\text{-}\gamma)$ vs. $\log(c)$]. Such a power law dependence of response times on stimulus strength is well documented in human psychophysics, and is referred to as Pieron's law (5). The offset parameter (γ), which represents the asymptotic response time, could not be reliably estimated for the uncued and invalidly cued trials because the response time data did not asymptote even at the highest T_{rel} values tested. Hence, γ, for the uncued (or invalidly cued) trials was assumed to be identical to that of the interleaved cued (or validly cued) trials, and only α and β were permitted to vary (Fig. S2).

Unless noted otherwise, error bars represent SEs across experimental sessions, computed with the jack-knife procedure (described in detail in ref. 2). All curve fitting was done with a nonlinear, least squares curve-fitting algorithm as implemented by the *lsqcurvefit* function in Matlab.

Signal Detection Analysis with the 2-AFC Model. We used a signal detection model for the 2-AFC task that incorporates bias (6). Hit rates (h) were arbitrarily assigned as the proportion of correctly localized upper targets, and false-alarm rates (f) arbitrarily assigned as the proportion of incorrectly localized lower targets (by construction of the 2-AFC signal detection model, the converse assignations for hit and false-alarm rates should provide identical d' values). Localization accuracy (d') and bias (b) were computed from the following relations (7):

$$d' = \Phi^{-1}(h) - \Phi^{-1}(f)$$

$$b = \frac{1}{2}\left(\Phi^{-1}(h) + \Phi^{-1}(f)\right)$$

where Φ^{-1} denotes the inverse of the cumulative unit normal (Gaussian) distribution function. Accuracy (d') and bias (b) were computed independently for each combination of target and distracter contrasts; for stimuli of matching strengths, an equal-variance signal detection model (as used here) is appropriate (8). Average d' and b were computed for each relative target strength (by averaging across the diagonal entries) or absolute target contrast (by averaging across distracter contrasts) to produce the psychometric functions shown in Fig. 2.

Psychometric functions describing the variation of d' with relative target strength were fit with a three-parameter cumulative Gaussian function (as described above), whereas those describing the variation of d' with absolute target contrast were fit with a three-parameter Naka–Rushton function of the form

$$d' = d_{\max}\frac{c^n}{c^n + c_{50}^n}$$

where c denotes the stimulus contrast, d' denotes the localization accuracy, $d_{\max}$ denotes asymptotic accuracy, n denotes the slope parameter, and c_{50} denotes the contrast at which accuracy reached 50% of maximum. The claim regarding the effects of spatial cueing as a contrast gain (reduction in c_{50}) was also confirmed with the parameters derived from the psychometric fits to d'.

Testing for Significant Differences Between Cued and Uncued Performance (Bootstrap Analysis). We used a bootstrap procedure (9) to estimate significant differences between cued vs. uncued (or validly vs. invalidly cued) localization accuracy values, mean response times, or parameters of the psychometric function. The bootstrap analysis was performed by randomly interchanging trial labels (cued vs. uncued or validly vs. invalidly cued) for all trials in an experiment, and independently for all experiments. We repeated this procedure 1,000–10,000 times, and the differences between the quantities of interest (accuracy, MRT, $d_{\max}$, etc.) were computed for each permutation to create a bootstrap (null) distribution. The number of potential permutations was several orders of magnitude greater (2^N, where N, the number of experiments, was ~200) than the number of repetitions (10^3 or 10^4), thereby permitting unique permutations. One-tailed P values reported represent the proportion of values in the bootstrap distribution that were smaller (or larger) than the observed differences between the quantities or computed parameters, depending on the a priori hypothesis for the direction of the change (cueing was hypothesized to improve, and distracters to degrade, all performance metrics). For each analysis, P values were Bonferroni corrected for multiple comparisons.

Signal Detection Analysis with the Two-Alternative Unforced Choice Model. In our tasks, birds could provide a NoGo response (two pecks on the zeroing cross) following the appearance of the response boxes, and receive an uncertain reward (reward probability = 0.33). Such tasks have been referred to as "unforced" choice or "uncertain option" tasks in the literature (10). Such a two-alternative unforced-choice (2-AUFC) design permitted us to estimate the birds' choice certainty in this task.

The conventional 2-AFC signal detection model (described above) is not sufficient to analyze NoGo responses. Therefore, we developed a signal detection model for the 2-AUFC task to analyze the NoGo responses. In this model, a NoGo response occurs whenever the evidence does not strongly favor the presence of a stimulus at either of the two locations (upper vs. lower) (Fig. 3*A*). This happens on trials in which the decision variable (*SI Appendix*) falls between two criterion values (biases, b_1 and b_2; Fig. 3*A*, thick vertical lines). In trials for which the decision variable falls outside these values, Go responses are generated to either location depending on the domain of decision space in which the decision variable falls (less than $-b_1$ or greater than $+b_2$). The biases b_1 and b_2 are, in principle, independent of d' and we hypothesized that the values of these biases (their range or the sum of their absolute values: $|b_1| + |b_2|$) were produced by a process that reflected the bird's choice certainty with generating a Go response.

The model permitted establishing an analytical relationship between the localization accuracy (d'); biases (b_1 and b_2); and the proportion of NoGo responses (p_N), correct Go responses (hits, p_H), and incorrect Go responses (errors, p_E):

$$p_H = \frac{1}{2}\left[\Phi\left(\frac{d'}{2} - b_1\right) + \Phi\left(\frac{d'}{2} - b_2\right)\right]$$

$$p_E = 1 - \frac{1}{2}\left[\Phi\left(\frac{d'}{2} + b_1\right) + \Phi\left(\frac{d'}{2} + b_2\right)\right]$$

$$p_N = \frac{1}{2}\left[\Phi\left(\frac{d'}{2} + b_1\right) + \Phi\left(\frac{d'}{2} + b_2\right)\right] - \frac{1}{2}\left[\Phi\left(\frac{d'}{2} - b_1\right) + \Phi\left(\frac{d'}{2} - b_2\right)\right]$$

where Φ denotes the cumulative unit normal (Gaussian) distribution function; we have assumed equal prior probabilities for targets at the two locations (upper vs. lower). The derivation of these relationships is provided in *SI Appendix*, and is based on a latent variable formulation similar to the one used in ref. 7.

Based on the observed data, these analytical relationships were used to derive maximum likelihood (ML) estimates for the accuracy and bias parameters (d', b_1, b_2). ML estimates were generated using constrained optimization (line search), as implemented by the *fmincon* function in Matlab, with the constraint that all of the parameters had to be positive. Randomization tests based on the χ^2

statistic were used to evaluate the goodness-of-fit of the 2-AUFC model to behavioral data.

In addition to fitting the model to the entirety of the observed responses, we evaluated the ability of the model to accurately describe behavior by fitting the model to a subset of stimulus-response contingencies, and generating predictions for the remaining stimulus-response contingencies (not used in fitting the model). Such a procedure provides an internal test of model validity in describing behavior (ref. 6, p. 193). Model predictions were generated as follows. The 2-AUFC task provides a 2×3 stimulus–response contingency table: For each of two stimulus events (target upper vs. lower), one of three response types can occur (Go response upper, Go response lower, or NoGo). Of the three (conditional) response probabilities for each of the stimulus events, only four (two for each contingency) are free to vary independently, as the three conditional response probabilities for each stimulus event are mutually exclusive and exhaustive (have to sum to 1). On the other hand, the 2-AUFC model is a three-parameter model (d', b_1, b_2) that requires only three independent observations for parameter estimation. Thus, there is a surplus of independent observations (one extra degree of freedom) relative to the number of free parameters.

By estimating 2-AUFC parameters from observed response proportions for only three (of six) contingencies, we predicted the value of the remaining one independent, and two dependent, contingencies. The three contingencies used to fit the model were repeatedly sampled from among the various possible combinations of the six stimulus–response contingencies (with at least one from each stimulus event), and predictions were, in turn, generated for the remaining contingencies (Fig. S4*A*). The model fit was performed separately for each target contrast, averaging across distracter contrasts.

Control for Overt Biases in Head Position at the Time of Target Presentation. Head position (*x*/*y*/*z* corresponding to azimuth, elevation, and distance from the screen, respectively) and orientation (roll/pitch/yaw) were acquired in real time based on an infrared-based motion capture system (Natural Point; Optitrack V120 Duo). The system has a 120-Hz sampling rate (8.33-ms sampling period), and submillimeter spatial resolution in all three dimensions. Measurements of head position were based on three to four infrared reflective markers in a triangular/quadrangular configuration (~1.5 to 2.0-cm spacing between markers) placed on a mount attached to the head (Movie S1). Motion capture data were analyzed (with custom software) offline to identify and exclude trials in which the head position was systematically different between cued and uncued (or between validly and invalidly cued) trials at the time of stimulus presentation. Specifically, we looked for horizontal head displacement (Δ) and rotation (yaw, θ; Fig. 5*A*) during the stimulus presentation period, in a window from 0 to 75 ms following the peck that triggered the stimulus array. We identified trials with closely matching Δ and θ values (tolerance: $\Delta = \pm 0.5$ mm; $\theta = \pm 1°$, the linear and angular resolution, respectively, of our motion capture technique) between cued and uncued (or validly and invalidly cued) trials for each bird (including all trials of each type), and separately for stimuli on either side (left vs. right). The very stringent tolerance limits for matching Δ and θ values resulted in a large proportion of the trials being excluded following the matching procedure: ~50% for the cued/uncued task and ~75% for the valid/invalid cue task. The proportion of excluded trials is greater in the valid/invalid cued task due to the very small proportion of invalidly cued trials (~10%) to which Δ and θ in the validly cued trials were stringently matched. The stringent tolerance for Δ and θ ensured that following the matching procedure, distributions of Δ and θ values closely overlapped across cued and uncued (or validly and invalidly cued) trials (Fig. S6 *A* and *B*, green; data from bird 1). The data were reanalyzed after excluding trials that fell outside this region of overlap (Fig. 5 *B*–*E*, solid lines); despite the removal of a significant proportion of trials, the results resembled the original results based on the entire dataset (Fig. 5 *B*–*E*, dashed lines).

SI Appendix: 2-AUFC Model

We derive a signal detection model for a 2-AUFC task. In such tasks, a subject can make either a Go response to report one of two mutually exclusive stimuli, or a NoGo (or an opt-out or "Don't know") response. NoGo responses are rewarded with a smaller reward (or on a smaller proportion of trials) than accurate Go responses.

We base the model on a 2-AUFC spatial localization task in which the subject is required to discriminate between two potential target locations (e.g., upper vs. lower), although the model is generally applicable to a variety of 2-AUFC discriminations (color, orientation, etc.).

The 2×3 contingency table for a 2-AUFC task consists of two stimulus events (a stimulus at one of two locations) and three response types (Go response at either one of the two locations or a NoGo response) (see table below). We use the following notation for the stimulus and response events (ref. 7): X_i is a random variable that denotes the location of the stimulus on each trial, $X_i = 1$ indicates the presentation of a stimulus at location i, and $X_i = 0$, otherwise. In our 2-AUFC localization task, a stimulus is always presented at one of two locations on each trial. Hence, the following condition is necessarily fulfilled: $X_1 + X_2 = 1$. Y is a variable that denotes the response on each trial; $Y = 1$ indicates a Go response to location 1, $Y = 2$ indicates a Go response to location 2, and $Y = 0$ indicates a NoGo response.

We develop a latent variable formulation that relates the conditional probability of each type of response for each stimulus event to the underlying psychophysical parameters: accuracy and bias.

As in conventional signal detection theory (SDT), we define a decision variable, Ψ_i, that encodes sensory evidence at each location, i ($i \in \{1, 2\}$). For the task described in the text, indices 1 and 2 represent upper and lower locations, respectively.

The decision variable distribution at each location i is defined as follows:

$$\Psi_i = dX_i + \varepsilon_i \qquad \textbf{[S1]}$$

where ε_i represents the distribution of the decision variable when no stimulus is presented (noise distribution). In other words, the mean of the decision variable distribution at a location is proportional to the strength of the stimulus at that location (d), and the variance of the decision variable distribution represents the noisiness of the sensory evidence. In line with conventional SDT, the ε_i follows a unit normal distribution, $\mathcal{N}(\mu = 0, \sigma = 1)$ (zero-mean Gaussian distribution with unit variance). In addition, the decision variable at each location is assumed to be independently distributed (no interaction terms). Note that for unit normal noise, the metric d also represents the discriminability (or overlap) of the Ψ distributions for the two stimulus events (Fig. 3*A*).

The subject's choice of response (Y) depends on the following decision rule:

$$\begin{array}{llll} Y = 1 & \text{if} & \Psi_1 > \Psi_2 + b_2 & \\ Y = 2 & \text{if} & \Psi_2 > \Psi_1 + b_1 & \\ Y = 0 & \text{if} & \Psi_1 \le \Psi_2 + b_2 \quad \text{and} & \Psi_2 \le \Psi_1 + b_1 \end{array} \qquad \textbf{[S2]}$$

where the two criterion values, b_1 and b_2 (biases), determine cutoff values of the Ψ_1 and Ψ_2 decision variables, respectively. This decision rule is used by the subject for reporting a Go response to location 1 ($Y = 1$) or location 2 ($Y = 2$), or a NoGo response ($Y = 0$) (Fig. 3*A*, $b_1, b_2 \ge 0$).

With this formulation, we derive the hit rate at location 1, i.e., the probability of a response to location 1 when a stimulus was presented at location 1.

$$\begin{aligned} p(Y=1|X_1=1,X_2=0) &= p(\Psi_1>\Psi_2+b_2|X_1=1,X_2=0)\\ &= p(dX_1+\varepsilon_1>dX_2+\varepsilon_2+b_2|X_1=1,X_2=0)\\ &= p(d+\varepsilon_1>\varepsilon_2+b_2)\\ &= p(\varepsilon_2-\varepsilon_1<d-b_2). \end{aligned}$$

Because ε_1 and ε_2 are zero-mean, independent, Gaussian-distributed random variables, their difference is also a zero-mean Gaussian-distributed variable with a variance that is the sum of the two individual variances. We scale either side of the inequality with the SD of $\varepsilon_2-\varepsilon_1$ ($\sigma=\sqrt{2}$), to make the distribution unit normal.

$$\begin{aligned} p(Y=1|X_1=1,X_2=0) &= p\left(\frac{\varepsilon_2-\varepsilon_1}{\sqrt{2}}<\frac{d-b_2}{\sqrt{2}}\right)\\ &= \Phi\left(\frac{d-b_2}{\sqrt{2}}\right) \end{aligned} \quad \textbf{[S3]}$$

where Φ denotes the cumulative distribution function of the unit normal distribution.

Next, we derive the false-alarm (or error) rate at location 1, i.e., the probability of a response at location 1 when a stimulus was presented at location 2.

$$\begin{aligned} &p(Y=1|X_1=0,X_2=1)=p(\Psi_1>\Psi_2+b_2|X_1=0,X_2=1)\\ &=p(dX_1+\varepsilon_1>dX_2+\varepsilon_2+b_2|X_1=0,X_2=1)\\ &=p(\varepsilon_1>d+\varepsilon_2+b_2)\\ &=p(\varepsilon_2-\varepsilon_1<-(d+b_2))\\ &=\Phi\left(\frac{-(d+b_2)}{\sqrt{2}}\right)\\ &=1-\Phi\left(\frac{d+b_2}{\sqrt{2}}\right) \end{aligned} \quad \textbf{[S4]}$$

where we have used the identity, $\Phi(-x)=1-\Phi(x)$.

Finally, we derive the "miss" rate for location 1, i.e., the probability of a NoGo response when a stimulus was presented at location 1.

$$\begin{aligned} &p(Y=0|X_1=1,X_2=0)\\ &=p(\Psi_1\le\Psi_2+b_2\cap\Psi_2\le\Psi_1+b_1|X_1=1,X_2=0)\\ &=p(dX_1+\varepsilon_1\le dX_2+\varepsilon_2+b_2\cap\\ &\quad dX_2+\varepsilon_2\le dX_1+\varepsilon_1+b_1|X_1=1,X_2=0)\\ &=p(d+\varepsilon_1\le\varepsilon_2+b_2\cap\varepsilon_2\le d+\varepsilon_1+b_1)\\ &=p(\varepsilon_2-\varepsilon_1\ge d-b_2\cap\varepsilon_2-\varepsilon_1\le d+b_1)\\ &=\int_{(d-b_2)/\sqrt{2}}^{(d+b_1)/\sqrt{2}} f(\hat{\varepsilon})d\hat{\varepsilon}\\ &=\Phi\left(\frac{d+b_1}{\sqrt{2}}\right)-\Phi\left(\frac{d-b_2}{\sqrt{2}}\right) \end{aligned} \quad \textbf{[S5]}$$

where $\hat{\varepsilon}=(\varepsilon_2-\varepsilon_1)/\sqrt{2}$ and we have used the identity, $\int_a^b f(\hat{\varepsilon})d\hat{\varepsilon}=\Phi(b)-\Phi(a)$ (as before, Φ denotes the cumulative distribution function of the unit normal distribution).

In a similar fashion, we derive the hit rate, false-alarm (or error) rate, and miss rate (NoGo response probability) for location 2.

$$p(Y=2|X_1=0,X_2=1)=\Phi\left(\frac{d-b_1}{\sqrt{2}}\right) \quad \textbf{[S6]}$$

$$p(Y=2|X_1=1,X_2=0)=1-\Phi\left(\frac{d+b_1}{\sqrt{2}}\right) \quad \textbf{[S7]}$$

$$p(Y=0|X_1=0,X_2=1)=\Phi\left(\frac{d+b_2}{\sqrt{2}}\right)-\Phi\left(\frac{d-b_1}{\sqrt{2}}\right) \quad \textbf{[S8]}$$

This completes the model formulation, except for a scaling factor. This factor arises because the parameters d,b_1,b_2 are all represented along orthogonal decision axes in the model. On the other hand, the axis of discrimination (connecting the centers of the Ψ distributions for each stimulus event) follows an angle of 45° with respect to these decision axes. Hence, upon projecting each parameter onto the discrimination axis (decision axis of Fig. 3*A*) a scaling factor of $1/\sqrt{2}$ must be applied to each parameter. In addition, as $d/\sqrt{2}$ represents the displacement of each distribution from their midpoint, it must be doubled to obtain the distance between their centers (d' in Fig. 3*A*).

Thus, we compute these scaled parameters as

$$\hat{d}=\frac{2d}{\sqrt{2}}=\sqrt{2}d \quad \widehat{b_1}=\frac{b_1}{\sqrt{2}} \quad \widehat{b_2}=\frac{b_2}{\sqrt{2}}. \quad \textbf{[S9]}$$

These parameters, $\hat{d},\widehat{b_1}$, and $\widehat{b_2}$, are referred to as d', b_1, and b_2, respectively in the main text.

The following 2 × 3 contingency table summarizes the response probabilities and their relationship to the parameters d',b_1,b_2, as defined in the main text (Fig. 3*A*).

Response probability	Go response at location 1, Y = 1	Go response at location 2, Y = 2	NoGo response, Y = 0
Stimulus at location 1 ($X_1=1,X_2=0$)	$\Phi\left(\frac{d'}{2}-b_2\right)$	$1-\Phi\left(\frac{d'}{2}+b_1\right)$	$\Phi\left(\frac{d'}{2}+b_1\right)-\Phi\left(\frac{d'}{2}-b_2\right)$
Stimulus at location 2 ($X_1=0,X_2=1$)	$1-\Phi\left(\frac{d'}{2}+b_2\right)$	$\Phi\left(\frac{d'}{2}-b_1\right)$	$\Phi\left(\frac{d'}{2}+b_2\right)-\Phi\left(\frac{d'}{2}-b_1\right)$

The proportion of correct (hits, p_H), incorrect (error, p_E), and NoGo (miss, p_N) responses are given by multiplying the conditional probabilities with the prior probability of each stimulus event (e.g., upper or lower):

$$\begin{aligned} p_H &= p(Y=1|X_1=1,X_2=0)\,p(X_1=1,X_2=0)\\ &\quad +p(Y=2|X_1=0,X_2=1)\,p(X_1=0,X_2=1)\\ p_E &= p(Y=2|X_1=1,X_2=0)\,p(X_1=1,X_2=0)\\ &\quad +p(Y=1|X_1=0,X_2=1)\,p(X_1=0,X_2=1)\\ p_N &= p(Y=0|X_1=1,X_2=0)\,p(X_1=1,X_2=0)\\ &\quad +p(Y=0|X_1=0,X_2=1)\,p(X_1=0,X_2=1). \end{aligned}$$

For equal prior probabilities for the two stimulus events $[p(X_1=1,X_2=0)=p(X_1=0,X_2=1)=0.5]$, these relationships are

$$\begin{aligned} p_H &= \frac{1}{2}\left[\Phi\left(\frac{d'}{2}-b_1\right)+\Phi\left(\frac{d'}{2}-b_2\right)\right]\\ p_E &= 1-\frac{1}{2}\left[\Phi\left(\frac{d'}{2}+b_1\right)+\Phi\left(\frac{d'}{2}+b_2\right)\right]\\ p_N &= \frac{1}{2}\left[\Phi\left(\frac{d'}{2}+b_1\right)+\Phi\left(\frac{d'}{2}+b_2\right)\right]-\frac{1}{2}\left[\Phi\left(\frac{d'}{2}-b_1\right)+\Phi\left(\frac{d'}{2}-b_2\right)\right]. \end{aligned} \quad \textbf{[S10]}$$

These represent the variation of correct responses, errors and misses with d', b_1, and b_2 described in the main text and *SI Methods*.

1. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10(4):433–436.
2. Sridharan D, Ramamurthy DL, Knudsen EI (2013) Spatial probability dynamically modulates visual target detection in chickens. *PLoS ONE* 8(5):e64136.
3. Schwarz JS, Sridharan D, Knudsen EI (2013) Magnetic tracking of eye position in freely behaving chickens. *Front Syst Neurosci* 7:91.
4. Martinoya C, Le Houezec J, Bloch S (1984) Pigeon's eyes converge during feeding: Evidence for frontal binocular fixation in a lateral-eyed bird. *Neurosci Lett* 45(3):335–339.
5. Luce RD (1986) *Response Times* (Oxford Univ Press, New York).
6. Macmillan NA, Creelman DC (2005) *Detection Theory: A User's Guide* (Lawrence Erlbaum Associates Inc., Mahwah, NJ).
7. DeCarlo LT (2012) On a signal detection approach to m-alternative forced choice with bias, with maximum likelihood and bayesian approaches to estimation. *J Math Psychol* 56:196–207.
8. Smith PL, Ratcliff R (2009) An integrated theory of attention and decision making in visual signal detection. *Psychol Rev* 116(2):283–317.
9. Efron B, Tibshirani RJ (1993) An introduction to the bootstrap (Chapman and Hall, New York).
10. Kepecs A, Mainen ZF (2012) A computational framework for the study of confidence in humans and animals. *Philos Trans R Soc Lond B Biol Sci* 367(1594): 1322–1337.

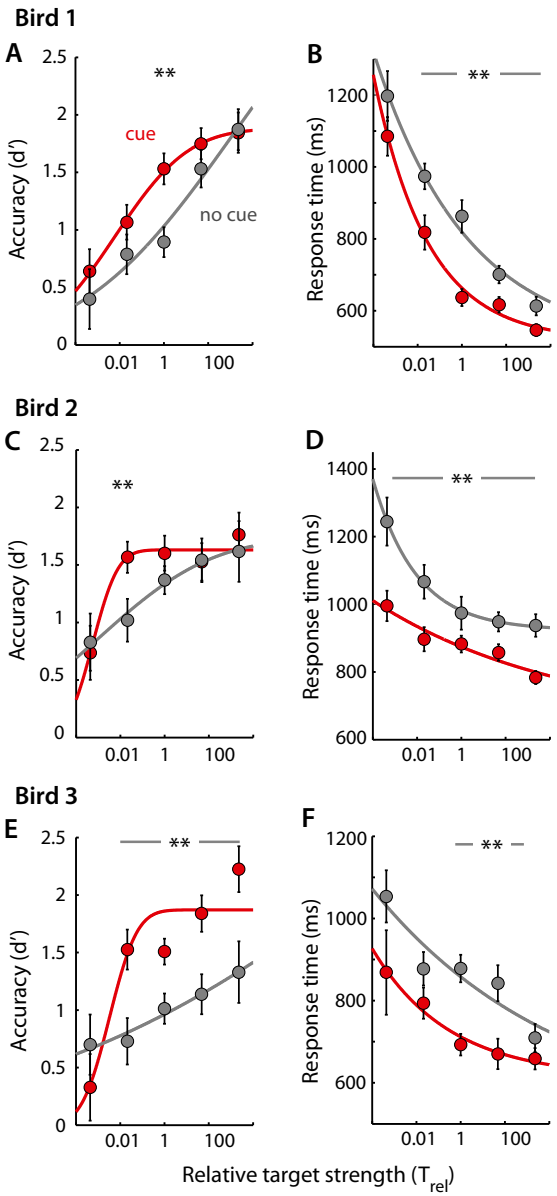


Fig. S1. Spatial cueing effects for individual birds. (*A*) Psychometric functions of localization accuracy (d', 2-AFC) without (gray) and with (red) the cue for an individual bird (bird 1), as function of relative target strength (T_{rel}). Other conventions are the same as in Fig. 2*B* (n = 87 experiments). (*B*) Psychometric functions of response time without (gray) and with (red) the cue for an individual bird (bird 1). Other conventions are the same as in Fig. 1*G*. (*C* and *D*) Same as in *A* and *B* except for bird 2 (n = 44 experiments). (*E* and *F*) Same as in *A* and *B* except for bird 3 (n = 68 experiments).

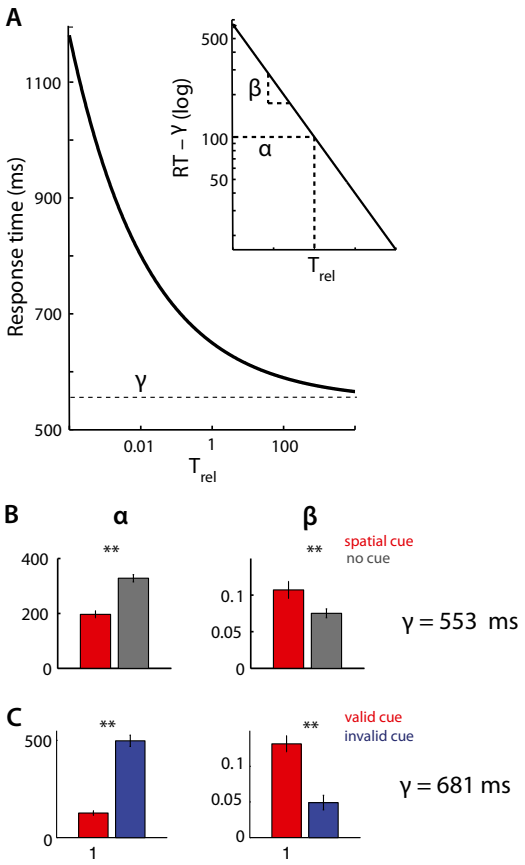


Fig. S2. Spatial cueing effects on response times. (*A*) Schematic showing parameters of the power law (Pieron's law) fits to psychometric functions of response times, as a function of relative target strength (log units). γ is the offset parameter (asymptote), α is the scaling parameter, and β is the slope parameter (exponent). (*Inset*) Offset subtracted response time (RT-γ) as a function of T_{rel} (both in log units). α is the value of RT-γ at $T_{rel} = 1$ (dashed lines) and β is the slope of the linear fit of (RT-γ) vs. T_{rel} (on a log–log scale). (*B*) α and β estimates from psychometric functions of response times (Fig. 1*G*) in cued (red) and uncued (gray) trials). Error bars represent SEM (jack-knife). γ was determined from the cued data alone (*SI Methods*). Other conventions are as in Fig. 1*H* (*Right*). (*C*) Same as in *B*, but from psychometric functions of response times (Fig. 4*C*) in validly cued (red) and invalidly cued (blue) trials. Other conventions are as in Fig. 4*D*.

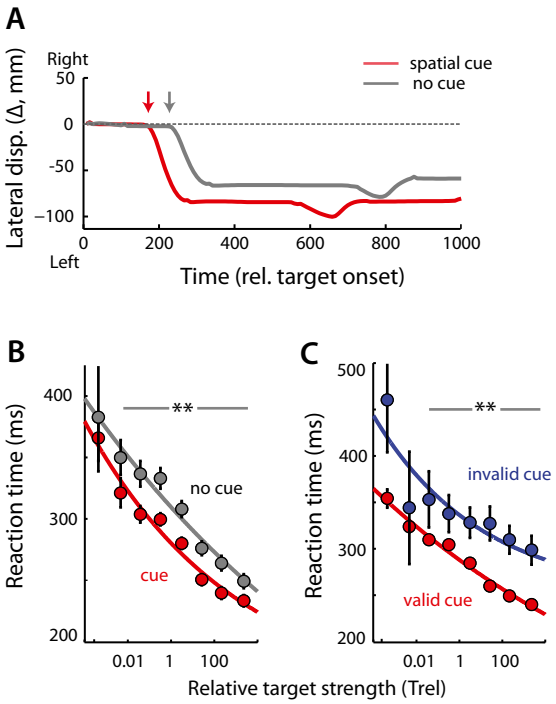


Fig. S3. Spatial cueing effects on reaction times. (*A*) Representative traces of lateral displacement of head position following target array onset (t = 0). Reaction times were computed based on the onset of the initial movement (arrows) toward the side of the target. Gray trace represents an uncued trial and red, a cued trial. (*B*) Reaction times, without (gray) and with (red) a spatial cue, for correctly localized targets as a function of relative target strength (population data). Other conventions are the same as in Fig. 1*G*. (*C*) Reaction times, with valid (red) and invalid (blue) spatial cues, for correctly localized targets as a function of relative target strength (population data). Other conventions are the same as in Fig. 4*C*.

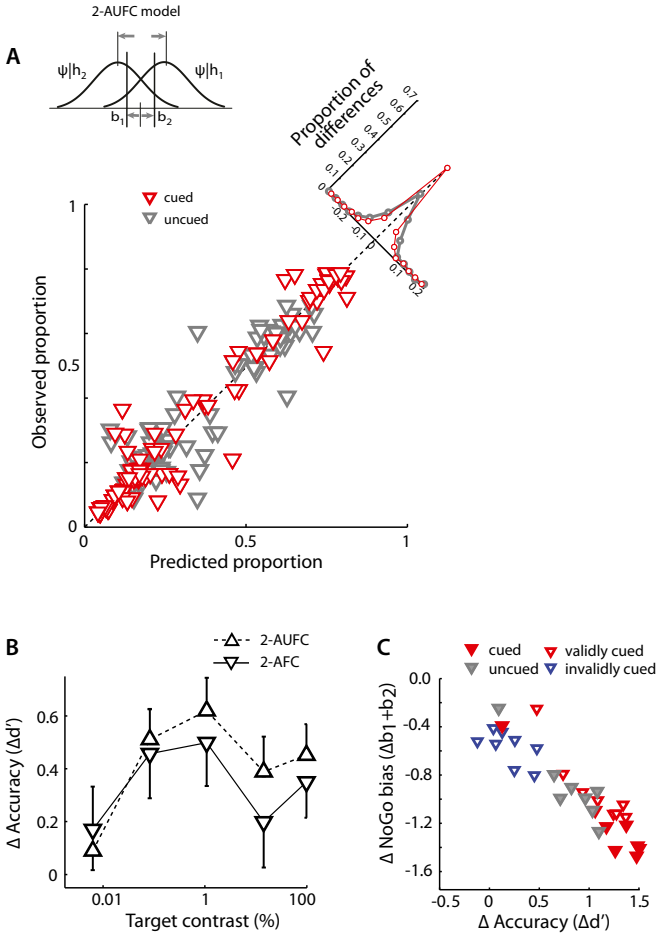


Fig. S4. Signal detection analysis with the 2-AUFC model. (*A*) Observed (*y* axis) response proportions vs. predictions (*x* axis) from the 2-AUFC model incorporating NoGo (opt-out) responses for performance without (gray) or with (red) a spatial cue. Accuracy and bias parameters (d', b_1, b_2) were estimated from only a subset of stimulus-response contingencies from the 2×3 contingency table (separately for each target contrast, averaged across distracter contrasts). Based on these parameters, 2-AUFC model predictions for response proportions for the remaining contingencies were generated (*SI Methods*). (*Inset, Upper Right*) Histogram showing the distribution of differences between observed and predicted response proportions. (*B*) Difference in accuracy ($\Delta d'$) between the cued and uncued trials computed with the 2-AFC (inverted triangles, solid line) and 2-AUFC (upright triangles, dashed line) models as a function of target contrast (population data). Error bars represent SEM (jack-knife). (*C*) Increase in $\Delta d'$ (relative to that for the lowest contrast targets, *x* axis) plotted against the decrease in the ΔNoGo bias $\Delta(b_1 + b_2)$ (relative to that for the lowest contrast targets, *y* axis). Individual points represent individual target contrast values (data pooled across birds). Filled symbols represent cued trials (red) interleaved with uncued trials (gray). Open symbols represent the validly cued trials (red) interleaved with invalidly cued trials (blue).

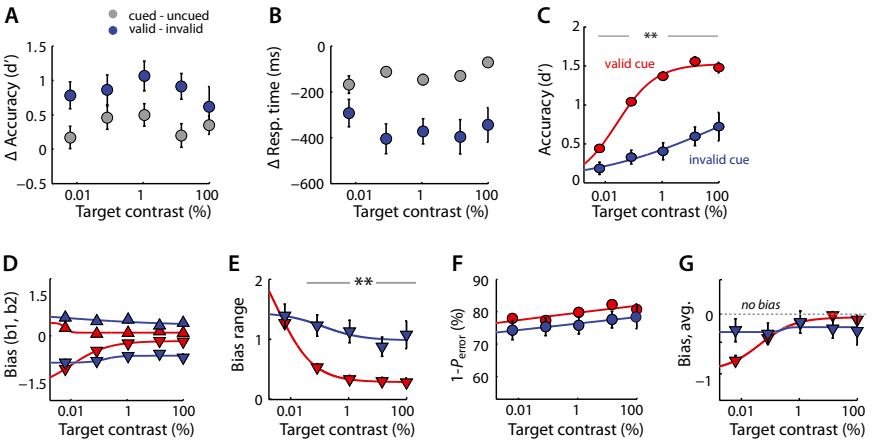


Fig. S5. Effects of invalid cueing. (*A*) Difference in accuracy (Δ*d*′, 2-AFC) between cued and uncued trials (gray) or between validly and invalidly cued trials (blue) as a function of target contrast. Error bars represent SEM (jack-knife). (*B*) Difference in response times (ΔResp. time) between cued and uncued trials (gray) or between validly and invalidly cued trials (blue) as a function of target contrast. Other conventions as in *B*. (*C*) Same as in Fig. 3*C*, but 2-AUFC localization accuracy (*d*′) for validly (red) vs. invalidly (blue) cued trials. Other conventions are as in Fig. 4*B*. (*D*) Same as in Fig. 3*D* (2-AUFC, biases), but for validly (red) vs. invalidly (blue) cued trials. (*E*) Same as in Fig. 3*E* (2-AUFC bias range or NoGo bias), but for validly (red) vs. invalidly (blue) cued trials. (*F*) Same as in Fig. 3*F* (error-aversion certainty), but for validly (red) vs. invalidly (blue) cued trials. (*G*) Same as in Fig. 3*G* (median 2-AUFC bias), but for validly (red) vs. invalidly (blue) cued trials.

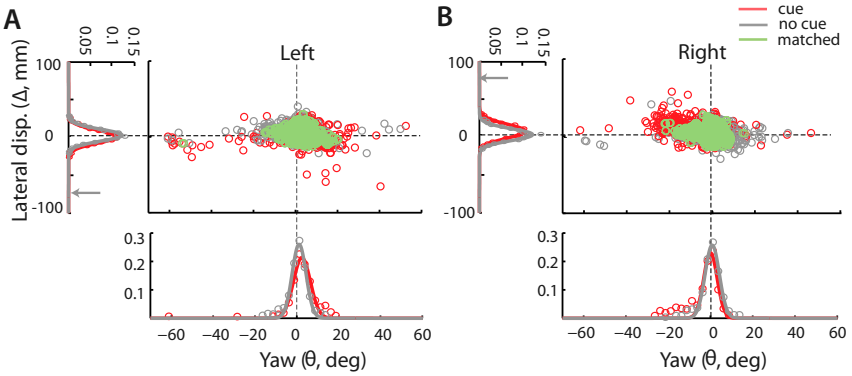
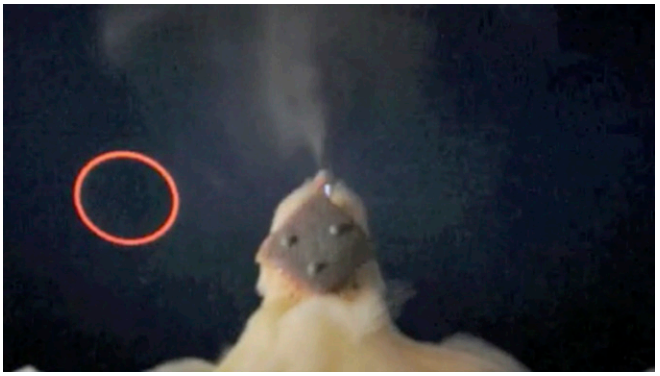


Fig. S6. Joint distribution of head lateral displacement and yaw in cued and uncued trials. (*A*) Joint distribution of Δ and θ (Fig. 5*A*) during target array presentation (0–75 ms following array onset) on trials without (gray circles) or with (red circles) a spatial cue (data from an individual bird). In these trials, the target was presented on the left side. Green circles are the Δ and θ values matched across uncued and cued trials (tolerance window: Δ = ±0.5 mm; θ = ±1°). Red circles (outside the green region) represent trials with a head position bias during cued trials that was not matched in uncued trials (or vice versa, for gray circles). Dashed lines represent Δ = 0 mm and θ = 0°. (*Left Inset*) Marginal distribution of Δ. The gray arrow is the target location. (*Lower Inset*) Marginal distribution of θ. Curves represent Gaussian fits (gray is uncued and red is cued). (*B*) Same as in *A*, but joint distribution of Δ and θ for targets presented on the right side.



Movie S1. Chicken performing a cued target localization (filtering) task. The sequence of events is described in *Results* and *Methods* and depicted in Figs. 1*A* and 4*A*. The movie here shows a modified version of the task described in the main text, in that the stimuli are presented further away from the zeroing cross and response boxes are presented away from the stimuli, and closer to the zeroing cross.

Movie S1