Supporting Information

Barthelmé and Mamassian 10.1073/pnas.1007704107

SI Materials and Methods

Experiment 1. Observers ran five sessions of the experiment in total. The first session was a training session, included so that orientation discrimination performance would be more stable in the following sessions. The first session involved a familiarization phase (500 trials total) and a pretest phase (400 trials), which we used to set the contrast levels in the main experiment.

Familiarization. For the familiarization period we used an easy version of the orientation discrimination task. Observers were instructed to fixate on a cross at the center of the screen and attend to a stimulus that would appear briefly to the left or to the right of the fixation cross. They had to report the perceived orientation of the target (CW or CCW), using the keyboard. The task is summarized in Fig. S1.

The observers performed the task over five blocks of 100 trials, set up so that difficulty increased gradually over time. In the first block the distractors were horizontal and the eccentricity was small (3.5°). In the second block eccentricity was increased to 10.7°, and in the third block the orientation of the distractors was changed to vertical, inducing crowding. In the last two blocks stimuli were presented at eccentricity of 12.5°, which was the eccentricity for the rest of the experiment, and distractors were horizontal in the fourth block and vertical in the fifth.

We varied the contrast of the central target in each block to accustom the observer to the task. Contrast levels were chosen using the same algorithm used in the pretest (see below). Observers were given feedback (correct/incorrect) on every trial. None reported difficulties in understanding or performing the task.

Pretest. The task in the pretest was identical to the familiarization task. We varied the contrast of the target adaptively to measure observers' discrimination performance as a function of contrast. Contrast levels were chosen to maximize the information about the parameters of the psychometric function, as in Kontsevich and Tyler (1).

Because performance is affected by the orientation of the distractors, the pretest was divided into four blocks of 100 trials: In two of those blocks the distractors were horizontal and in the other two they were vertical. Feedback was given every 5 trials only: Observers were told on how many trials they had responded correctly (of the previous 5).

Main experiment. The next four sessions occurred on different days and lasted <1 h. We used the data from the pretest to choose appropriate levels of contrast for the main experiment (i.e., levels that spanned the range of the psychometric functions). If these levels were found to be inappropriate, they were adjusted between sessions.

In the main experiment two stimuli were displayed in every trial instead of one, so the observer was briefly introduced to the new task at the beginning of the first session. The instructions given were to pick the stimulus for which they felt more confident, to maximize the number of correct answers. To provide motivation for doing so, feedback was given every five trials, as in the pretest.

Observers ran 800 trials, with obligatory 10-s breaks at the end of every 100-trial block.

Experiment 2. Experiment 1 was identical to experiment 2, except when indicated otherwise.

Familiarization. Observers performed 30 trials of orientation discrimination. The stimuli were displayed at a random eccentricity, but remained for 250 ms so that the task was extremely easy. Observers received feedback on every trial.

Pretest. We checked for the presence of a masking effect before running the main experiment. Observers performed 300 trials of orientation discrimination, and we interleaved trials with long and short ISIs. Eccentricities were chosen at random from six preset levels (method of constant stimuli). The resulting psychometric functions were inspected to make sure masking did occur with short ISIs. They were also used to assess the observer's performance at high eccentricities: The experiment requires performance to be low for unmasked stimuli at high eccentricities (otherwise there would be no benefit to ever choosing masked stimuli). If the observer's performance was too high for high eccentricities, the contrast of the stimulus was adjusted down for the main experiment and up otherwise. The contrast level used for every observer is listed in Table S1. Observers received feedback on every trial.

Main experiment. Observers performed 500 trials in the first session and 800 trials in the four other sessions, except for observer AD who due to an oversight performed 500 trials in the first two sessions and 800 trials in the rest. They received feedback every 5 trials.

Control Conditions. *Baseline.* In baseline trials observers chose between stimuli that had the same contrast and were both crowded or both uncrowded (in experiment 1) or had the same eccentricity and were both masked or both unmasked (in experiment 2). The benefits of picking one over the other were minimal: The observer could compensate only for eventual lapses in attention—e.g., missing the first stimulus entirely—or possible assymetries in the visual field. Performance in orientation discrimination as measured in baseline trials was used to infer the expected performance ratio, as explained in *Data Analysis*.

Performance in the baseline condition in experiment 1 is plotted in the main text (Fig. 2). The results for experiment 2 are plotted in Fig. S6: We summarize performance using the thresholds of the psychometric functions and verify that thresholds were higher in the masked condition.

Contrast only/eccentricity only. In experiment 1, in contrast-only trials the two stimuli had same orientation distractors (i.e., both stimuli were crowded or both were uncrowded), but the contrast of the two targets was different (resp. in experiment 2 the two stimuli were both masked or both unmasked but their eccentricity differed). They serve as a control condition because here observers should favor the stimulus with the highest contrast, showing that observers can use contrast or eccentricity when relevant.

To check that observers do pick the higher-contrast stimulus with probability higher than chance, we fit a logistic model:

$$\mathrm{logit}(\pi_R) = \beta_0 + \beta_1(\log x_R - \log x_L). \qquad \textbf{[S1]}$$

The model tries to predict the probability π_R that the observer will pick the right-hand stimulus, given that the contrast of the left-hand one is x_L and that of the right-hand one is x_R. β_0 and β_1 are free parameters, with β_0 controlling the bias and β_1 the effect of the difference in contrast. This is equivalent to fitting a logistic psychometric function to $\log \frac{x_R}{x_L}$, with β_1 determining the slope.

If observers indeed favor stimuli with higher contrast, we expect that $\beta_1 > 0$.

To check that β_1 was reliably measured above 0, we fit model 1 repeatedly to parametric bootstrap replicates. The distribution of replicates for each observer for crowded and uncrowded stimuli is shown in Fig. S3. The distributions do not overlap 0 except in the case of observer ZDC and NM for crowded stimuli (respectively

4% and 31% of replicates below 0). In all other cases the effect is extremely robust.

We ran a similar analysis for experiment 2, replacing the difference in log contrast with the difference in log proximity. The results appear in Fig. S4. We find no overlap of the bootstrap distributions with 0.

Data Analysis. ***Psychometric functions and the expected performance ratio.*** We fit psychometric functions by maximum likelihood to observers' measured performance in the baseline condition. We used the following functional form:

$$\Psi(x) = \frac{\Phi(\log x; \mu, \sigma^2) + 1}{2}. \qquad \textbf{[S2]}$$

Here $\Phi(x; \mu, \sigma^2)$ is the cumulative Gaussian distribution with mean μ and variance σ^2. μ and σ are the parameters of the psychometric function, controlling the threshold and slope, respectively. In experiment 1, we used two free parameters per subject and per distractor orientation, to obtain a set of curves $\{\Psi_{C,i}(x), \Psi_{U,i}(x)\}$, where i indexes the observer, C and U stand for crowded and uncrowded, and x stands for stimulus contrast. The same analysis holds for experiment 2, with crowded/uncrowded replaced with masked/unmasked and contrast replaced with proximity (the opposite of eccentricity). The expected performance ratio of observer i for a pair of stimuli of contrast (x_C, x_U) is defined as

$$\rho_i(x_C, x_U) = \frac{\Psi_{C,i}(x_C)}{\Psi_{U,i}(x_U)}. \qquad \textbf{[S3]}$$

A quantity of particular interest is the iso-performance curve, plotted as the black solid curve in Figs. 4 and 6 of the main text. This curve represents the set of stimuli that yield the same expected performance and is defined as

$$\frac{\Psi_{C,i}(x_C)}{\Psi_{U,i}(x_U)} = 1.$$

Because the $\Psi(x)$ are invertible functions, we can express the iso-performance curve as a function of x_C:

$$\phi(x_C) = \Psi_{U,i}^{-1}(\Psi_{C,i}(x_C)). \qquad \textbf{[S4]}$$

Note that $\phi(x_C)$ depends on observers' measured performance in the baseline condition and so is not a fixed experimental parameter. To obtain the confidence intervals on $\phi(x_C)$ plotted in Figs. 4 and 6 of the main text, we generated parametric bootstrap samples using the maximum-likelihood estimates of the psychometric function parameters (2), which in turn yield a bootstrap distribution over $\phi(x_C)$.

Response surfaces. The response surfaces for individual observers shown in Figs. 4 and 6 of the main text are a smoothed version of the raw response data obtained via multivariate adaptive response splines (MARS) (3). We modeled the probability of choosing the crowded stimulus as a logit transformation of a latent spatial function $f(x_C, x_U)$,

$$\text{logit}\ \pi(x_C, x_U) = f(x_C, x_U), \qquad \textbf{[S5]}$$

with $\text{logit}\ \pi = \log(\frac{\pi}{1-\pi})$. [Again, the exact same analysis holds for experiment 2, with crowded/uncrowded replaced with masked/unmasked and contrast replaced with proximity (the opposite of eccentricity).] We make no parametric assumptions about f. MARS models f in a spline basis:

$$f(x_C, x_U) = \sum_i w_i f_i(x_C, x_U). \qquad \textbf{[S6]}$$

The index i is over the basis set $\{f_1, \ldots, f_n\}$, which is made up of spline functions of an increasingly local nature. MARS is an adaptive procedure that tries to introduce only as many functions in the basis set as required by the data, avoiding overfitting. It makes for a useful visualization tool for our purposes, smoothing out the noise in the raw data but keeping the main features of the latent response surface. We used the *earth* package in R to obtain the plots (4). The surfaces obtained when using penalized thin plate splines (5) are extremely similar, so that the response surfaces we obtain are robust to the method used.

Assessing the reliability of the findings. Our main claim is that the response surfaces presented in Figs. 4 and 6 of the main text are incompatible with cue-based strategies. The reliability of that conclusion rests entirely on that of the response surfaces, which are estimated from the data. Because reliability in a (2D) response surface cannot be evaluated by the traditional techniques of confidence intervals, we first summarize the response surface via the *indifference line*, defined as the set of all stimuli (x_C, x_U) such that the observer is indifferent between the two options. This indifference line corresponds to the white regions in the response surfaces. Because we do not make any parametric assumptions about the response surface, nothing guarantees that the indifference line actually is a line, or is even continuous, but inspection of Figs. 4 and 6 suggests that it is roughly so. The following model,

$$\text{logit}\ \pi\,(x_C, x_U) = \alpha_0 + \alpha_1 x_C + \alpha_2 x_U, \qquad \textbf{[S7]}$$

provides a reasonable summary of the response surfaces, and the indifference line is defined by

$$I_\beta = \{(x_C, x_U) |\, \alpha_0 + \alpha_1 x_C + \alpha_2 x_U = 0\}. \qquad \textbf{[S8]}$$

Importantly, the parameters of the indifference line provide all we need to know to decide whether the observer is following a cue. If the observer follows the contrast cue (equivalently, the eccentricity cue in experiment 2), then we expect the indifference line to be the main diagonal, which implies $\alpha_1 = \alpha_2$ and $\alpha_0 = 0$. If the observer behaves according to the crowdedness (maskedness in experiment 2), we expect the indifference line to be the y axis, which implies $\alpha_2 = 0$.

We produced nonparametric bootstrap replicates of the indifference line for each observer by resampling the data and fitting model 7 to the resampled data. We plot the resulting bootstrap distributions in Figs. S5 and S6. The bootstrap distributions do not overlap the main diagonal, nor do they overlap the y axis (with the exception of subject ZDC in experiment 1). Our conclusions appear to be reliably supported by the data.

1. Kontsevich LL, Tyler CW (1999) Bayesian adaptive estimation of psychometric slope and threshold. *Vision Res* 39:2729–2737.
2. Wichmann FA, Hill NJ (2001) The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept Psychophys* 63:1314–1329.
3. Friedman JH (1991) Multivariate adaptive regression splines. *Ann Stat* 19:1–67.
4. Milborrow S, derived from code by T. Hastie, Tibshirani R (2007) Earth: Multivariate adaptive regression spline models R package. Available at http://cran.r-project.org/web/packages/earth/index.html. Accessed November 1, 2010.
5. Wood SN (2003) Thin plate regression splines. *J R Stat Soc Ser B Stat Methodol* 65: 95–114.

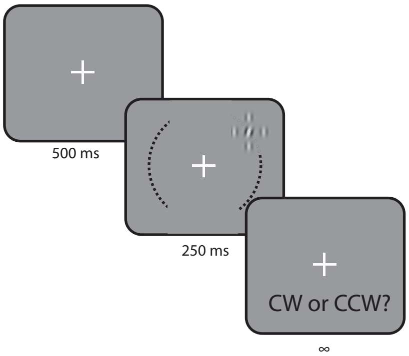

Fig. S1. Time course of familiarization and pretest trials in experiment 1. A cross appears for fixation, and then a stimulus is displayed in a random location along two semicircles (dotted curves). The stimulus disappears and the observer must indicate the perceived orientation of the target patch in the stimulus (here, it is tilted clockwise of the vertical). The ∞ symbol indicates the observer had unlimited time to respond.

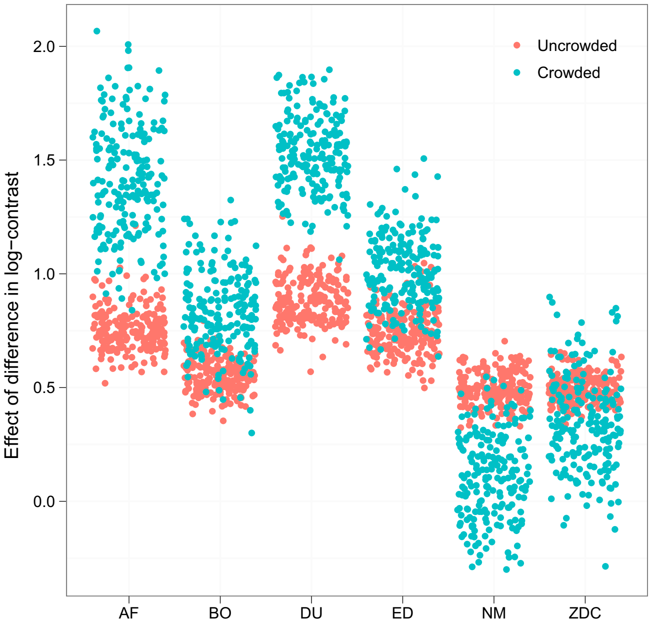

Fig. S2. β_1, effect of the difference in log-contrast on observers' choice of stimuli in the contrast-only condition of experiment 1. The effect of a difference in log-contrast is estimated from model 1. We plot bootstrap replicates of the estimated effect for all observers and for uncrowded and crowded stimuli. Random jitter is added in the *x*-direction to facilitate visualization.

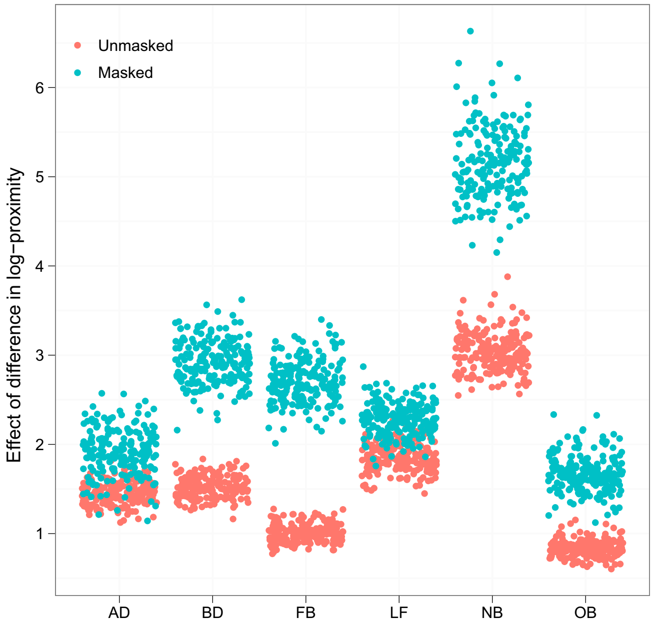

Fig. S3. Effect of the difference in log-contrast on observers' choice of stimuli in the contrast-only condition of experiment 2.

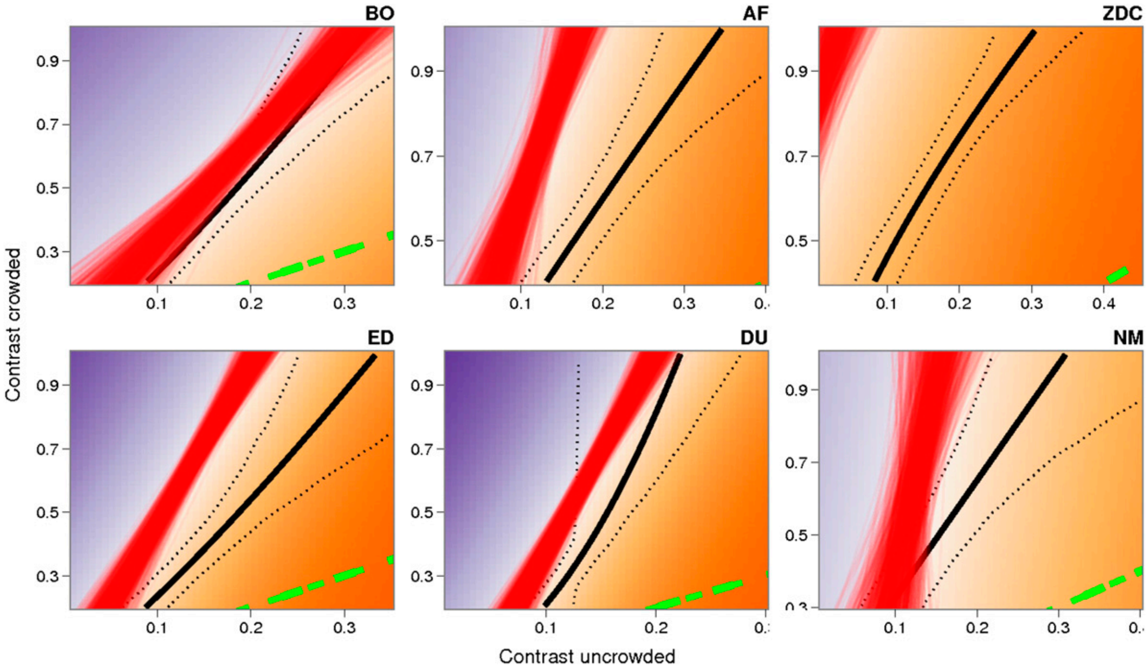

Fig. S4. Bootstrap distribution of the indifference line for the mixed condition in experiment 1. The same general format as that of Fig. 4 in the main text is shown. We plot linear response surfaces (obtained from fitting model 7) along with 1,000 bootstrap samples of the indifference line (in red).

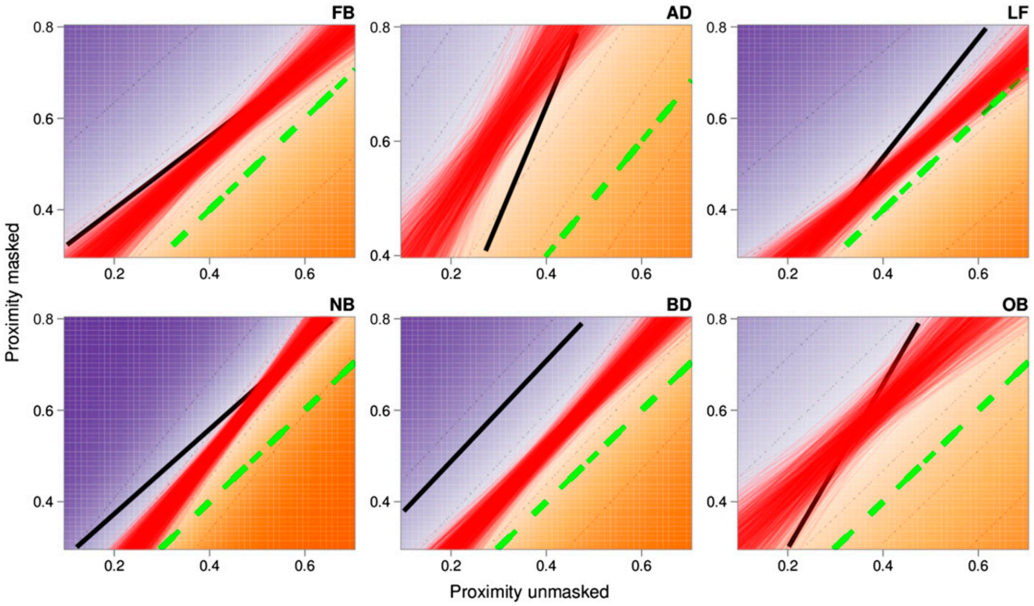

Fig. S5. Bootstrap distribution of the indifference line for the mixed condition in experiment 2.

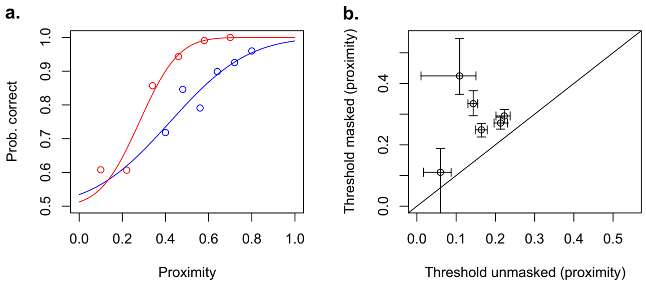

Fig. S6. Results for the baseline condition in experiment 2. We summarize performance data via a psychometric function linking stimulus *proximity* to probability correct. Proximity is defined as the opposite of eccentricity, with a value of 0 for the highest eccentricity displayable on the monitor and a value of 1 for the minimum eccentricity. (*A*) Psychometric functions in the masked (blue) and unmasked (red) conditions for observer AD. (*B*) Observers' thresholds in conditions unmasked and masked, with parametric bootstrap confidence intervals (10% and 90% quantiles). Higher performance in the unmasked condition is indicated by points falling above the main diagonal.

Table S1. Contrast levels used for each observer in experiment 2

Observer	FB	LF	NB	AD	BD	OB
Contrast, %	7	10	7.5	7	7.5	7