

How Selective Are V1 Cells for Pop-Out Stimuli?

Jay Hegdé^{1,2} and Daniel J. Felleman¹

¹Department of Neurobiology and Anatomy, University of Texas Medical School, Houston, Texas 77030, and ²Vision Center Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037

Many neurons in visual area V1 respond better to a pop-out stimulus, such as a single vertical bar among many horizontal bars, than to a homogeneous stimulus, such as a stimulus with all vertical bars. Many studies have suggested such cells represent neural correlates of pop-out, or more generally figure–ground segregation. However, preference for pop-out stimuli over homogeneous stimuli could also arise from a nonspecific selectivity for feature discontinuities between the target and the background, without any specificity for pop-out per se. To distinguish between these two confounding scenarios, we compared the responses of V1 neurons to pop-out stimuli with the responses to “conjunction-target” stimuli, which have more complex feature discontinuities between the target and the surround, as in a stimulus with a blue vertical bar among blue horizontal bars and yellow vertical bars. The target in conjunction-target stimuli does not pop out, which we psychophysically verified. V1 cells in general responded similarly to pop-out and conjunction-target stimuli, and only a small minority of cells (~2% by one measure) distinguished the pop-out and conjunction-target stimuli from each other and from homogeneous stimuli. Nevertheless, the responses of approximately 50% of the cells were significantly modulated across all center-surround stimuli, indicating that V1 cells can convey information about the feature discontinuities between the center and the surround as part of a network of neurons, although individual cells by themselves fail to explicitly represent pop-out. In light of our results, unambiguous pop-out selectivity at the level of individual cells remains to be demonstrated in V1 or elsewhere in the visual cortex.

Key words: binding; center-surround summation; contextual modulation; feature integration; figure–ground segregation; serial search; striate cortex; surround modulation; visual search

Introduction

A blue vertical bar is easy to recognize among a background of yellow horizontal bars. It “pops out”. In general, pop-out occurs when a target substantially differs from the background, or distractors, in terms of one or more visual features. But when the target is defined by a unique conjunction of features, such as a blue vertical bar target among yellow vertical and blue horizontal distractors, the time it takes to find the target, or reaction time, generally depends on the number of distractors (Treisman, 1980, 1988; Wolfe, 1994). By contrast, reaction time for pop-out is generally short, on the order of a few hundred milliseconds, and is independent of the number of distractors in the stimulus (Treisman, 1980, 1988). Psychophysical studies of pop-out and the “conjunction-target” stimuli have yielded important insights into the perceptual mechanisms of figure–ground segregation and feature integration (or binding), and are central to many influential models of visual object recognition (Julesz, 1984; Treisman, 1988; Wolfe, 1994; Palmer et al., 2000; Hochstein and Ahissar, 2002) (also see September 1999 special issue of *Neuron*).

Neural mechanisms of the pop-out phenomenon have been

studied in many visual areas, and especially intensively in visual area V1 (for review, see Albright and Stoner, 2002). Because V1 classical receptive fields (CRFs) tend to be small with relatively large nonclassical surrounds, the selectivity for pop-out has generally been studied in terms of center-surround modulation, with the target centered on the CRF (or the “center”) and the distractors in the surround. Previous studies have shown that many V1 cells distinguish pop-out stimuli from the corresponding homogeneous stimuli (i.e., stimuli with the same texture elements in the center and the surround), when the pop-out is based on center-surround differences in orientation (Knierim and Van Essen, 1992; Lamme, 1995; Zipser et al., 1996; Li et al., 2000; Nothdurft et al., 2000), motion (Lamme, 1995), color, stereoscopic disparity (Zipser et al., 1996), or luminance (Zipser et al., 1996; Levitt and Lund, 1997; Polat et al., 1998; Lee et al., 2002). Some studies have argued that such patterns of surround modulation represent neural correlates of the pop-out phenomenon (Kastner et al., 1997; Nothdurft et al., 1999) or, more generally, figure–ground segregation (Lamme, 1995; Zipser et al., 1996).

However, selectivity for pop-out stimuli relative to homogeneous stimuli could also arise from a nonspecific selectivity for feature discontinuities between the target and the background, rather than a selectivity for pop-out per se. Because pop-out stimuli necessarily contain the center-surround feature discontinuity and homogeneous stimuli do not, it is impossible to distinguish selectivity for the feature discontinuity from genuine selectivity for pop-out stimuli using these two types of stimuli alone. Additional stimuli that dissociate the existence of feature discontinuities from pop-out are needed.

Conjunction-target stimuli are useful for this purpose, because

Received May 25, 2003; revised Sept. 8, 2003; accepted Sept. 11, 2003.

This work was supported in part by a Fight for Sight Postdoctoral Fellowship to J.H., by National Institutes of Health Grant EY 08372 to D.J.F., and by Vision Core Grant P30-EY-10618 to the University of Texas Health Sciences Center. We are grateful to Dr. John Maunsell for advice and help throughout this project. We thank Drs. Robert Desimone and Andrew Mitz for useful software, Dr. Thomas Albright for the use of psychophysical experimental facilities, Dr. Jonathan Hill for useful discussions, and Drs. John Maunsell, Lawrence Snyder, and Gene Stoner for helpful comments on this manuscript.

Correspondence should be addressed to Jay Hegdé, Vision Center Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037. E-mail: jay@salk.edu.

Copyright © 2003 Society for Neuroscience 0270-6474/03/239968-13\$15.00/0

they also contain feature discontinuities between the center and the surround, but do not pop out. Because pop-out and conjunction-target stimuli differ in the precise nature of the discontinuities and the perceptual effects they elicit, neurons that are genuinely selective for pop-out stimuli or the pop-out percept should distinguish between the two types of stimuli. On the other hand, cells that are merely selective for the existence of center-surround discontinuities regardless of the nature of the discontinuities should respond similarly to pop-out and conjunction-target stimuli, but distinguish them both from the homogeneous stimuli.

We compared the responses of V1 cells to the three types of stimuli. We found that most V1 cells failed to explicitly distinguish among the three stimulus types by any of the response measures used. Our results suggest that area V1 is unlikely to contain explicit representations of visual pop-out, or even the precise nature of center-surround feature discontinuities. Nonetheless, V1 cells do signal the existence of feature discontinuities between the center and the surround and may play an important role in the analysis of center-surround feature differences as a part of a distributed network.

Materials and Methods

Neurophysiology

Animal subjects and surgical procedures. The neurophysiological experiments were performed in alert, fixating macaques. Two adult macaques (*Macaca mulatta*, one male and one female) were used as subjects. Before fixation training, each animal was implanted with a headpost and a scleral search coil using sterile surgical procedures. After recovery, the animals were trained in the fixation task, after which another surgical procedure was performed in which a craniotomy 2.5 cm in diameter was made over opercular V1, and a recording chamber was implanted over the craniotomy. All animal-related protocols were approved by The University of Texas Health Sciences Center Animal Welfare Committee in accordance with National Institutes of Health guidelines.

Stimuli. The stimulus set consisted of 36 bar array stimuli (Fig. 1, bottom), constructed using a repertoire of four bar types (Fig. 1, top): *A*, the bar with the most effective, or preferred, orientation and color of the cell (preferred bar); *B*, the bar with the preferred color but the least preferred (null) orientation of the cell; *C*, the bar with the preferred orientation but the null color of the cell; and *D*, the bar with the null color and orientation of the cell (i.e., null bar). Thirty-two of the stimuli were center-surround stimuli, with a single bar in the center (i.e., the CRF), and many additional bars (i.e., “distractors”) in the nonclassical surround. The center-surround stimuli included many pop-out stimuli in which the center bar, or the “target”, was distinguishable from the distractors in terms of orientation alone (orientation pop-out; e.g., stimuli A3, A5, B3, B5, etc.), color alone (color pop-out; e.g., A4, A6, B4, B6, etc.), or both (color-orientation pop-out, e.g., A7, B7, etc.). Note that the surrounds of the pop-out stimuli were uniform in some cases (e.g., A5–A7), and nonuniform in others (i.e., A3, A4). In some stimuli, which we will refer to as conjunction-target stimuli, the target was defined by a unique conjunction of color and orientation (e.g., A1–D1; A2–D2). Four stimuli (A8–D8) were homogeneous, with the same bar type in the center and the surround, and four were “center-alone” with only a bar in the center and none in the surround.

Collectively, our stimuli allowed us to directly compare the psychophysical and neurophysiological results from our current experiments, plus those from many previous studies. Also, the stimulus set included many novel stimuli the neuronal responses to which have never been studied to our knowledge (e.g., conjunction-target stimuli, pop-outs of both color and orientation, and pop-out stimuli with mixed surrounds). Our stimuli also allowed us to explore whether selectivity of V1 cells for pop-out versus conjunction-search stimuli varied as a function of the effectiveness of the center bar (i.e., preferred versus null color and/or orientation), because many previous studies have shown that surround stimuli are widely suppressive when the center stimulus is effective, and

Bar Type	Orientation	Color
A	Preferred	Preferred
B	Null	Preferred
C	Preferred	Null
D	Null	Null

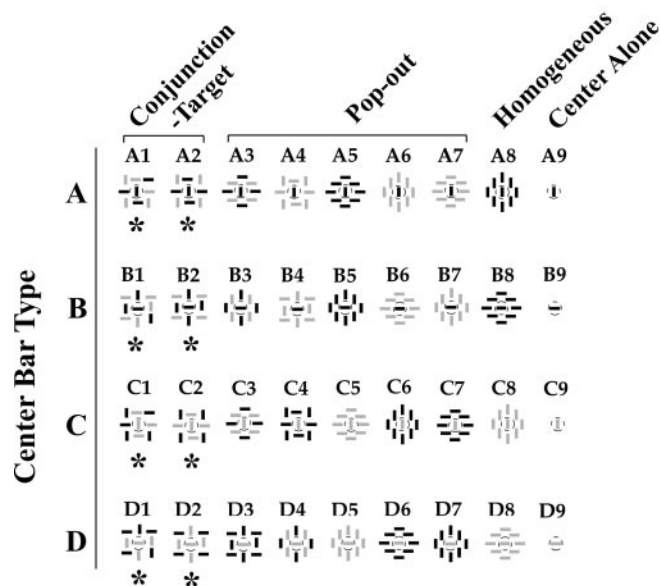


Figure 1. The stimulus set was constructed using four different bar types representing the four possible combinations of the most preferred or the least preferred (null) color and orientation (inset at top) of the cell. Of the 36 stimuli in the stimulus set (icons at bottom), four (“Center Alone” stimuli) consisted of a single bar presented within the CRF (dashed circle). The others were center-surround stimuli with the center bar surrounded by many additional bars, or distractors, in the nonclassical surround. Nine stimuli were constructed for each of the four types of center bar (rows A–D), including one Center Alone stimulus and eight center-surround stimuli. The center-surround stimuli consisted of three subclasses: the homogeneous stimuli, which did not contain any feature discontinuities between the center and the surround, pop-out stimuli, in which the center bar was distinguishable from the surround bars in terms of color and/or orientation, and the conjunction-target stimuli (asterisks), in which the target was defined by a unique conjunction of color and orientation. Note that both the pop-out and the conjunction-target stimuli contained feature discontinuities between the center and surround. The center bar in the neurophysiological experiments was the search target in psychophysical experiments. The actual stimuli used in either set of experiments differed from the icons shown in this figure in many respects. See Materials and Methods for additional details.

often facilitative when the center stimulus is weak (Sillito et al., 1995; Polat et al., 1998; Somers et al., 1998).

Visual stimulation. The display screen subtended a $19 \times 15^\circ$ of the visual field (screen resolution, 1280×1024 pixels; refresh rate, 75 Hz). Each stimulus was presented with the designated center bar centered on the CRF, with the surround bars (when present) in the nonclassical surround, filling the remainder of the display. The center-surround stimuli consisted of 59–110 bars, depending on the cell (see below), but for any given cell, all center-surround stimuli contained the same total number of bars. The surround bars, when present, were distributed in the surround in randomly jittered rows and columns so that no higher order structure among bars (e.g., rows or columns) was apparent by visual inspection. Surround stimuli with more than one bar type in the surround had equal (or nearly equal) number of each bar type in the surround, and the different bar types were distributed uniformly in the

surround. For any given stimulus, the location of the surround bars was randomly shuffled from one repetition to the next. This, together with the random spatial jittering of the bar locations and fixation jitter, minimized the likelihood that a given subregion of the surround was consistently stimulated by the same bar across repetitions. All stimuli, including those used during the initial receptive field mapping, were presented against a uniform gray background. All bars and the background had a luminance of 30 cd/m², as measured by Tektronix J17 photometer ("equiluminant condition"), except where noted otherwise.

Recording procedures. The experiments were controlled and the data collected using the CORTEX software package (courtesy of Dr. Robert Desimone, National Institute of Mental Health). After a cell was isolated for study, the CRF of the cell was mapped, and its receptive field preferences were determined using a mouse-driven bar on the display of the computer. The preferred and null orientations were determined independently of each other at a resolution of 5° each. The preferred and null colors were selected from a repertoire of six equiluminant colors. The cells in our sample in general had crisply delineated CRF boundaries, so that the center-surround distinction was clear and robust. The CRF diameters (range, 1–1.8°; mean, 1.2°; median, 1.3°) varied with the eccentricities (range, 1–6.2°; mean, 1.8°; median, 2.2°) as expected from previous studies (Van Essen and Zeki, 1978; Snodderly and Gur, 1995). To ensure that no surround bar stimulated the CRF during fixation, stimuli were constructed so that the closest points of any two bars were >1.2 CRF diameters or >1.2° apart, whichever was greater. The eye position of the animal was monitored throughout the trial using a scleral search coil, and the trial was aborted if the eye deviated by >0.5° from the fixation point at any time during the trial. Single-unit recording was performed using standard procedures. Recording coordinates were randomly chosen from within the craniotomy. Stimuli were presented one per trial for 1 sec each while the animal fixated for a juice reward at the end of the trial. We recorded from a total of 106 cells from three hemispheres of two monkeys during this experiment.

We also performed a "contrast effect experiment", which was identical to the main neurophysiological experiment described above except that only stimuli A1-A9 and B1-B9 were used in this experiment. Each of these 18 stimuli was presented at two different stimulus-background contrasts (Weber contrast, $\Delta I/I$) of 33 or 0%. The contrasts were achieved by setting the luminance of the background to either 30 or 20.1 cd/m² ("equiluminant" and "higher contrast" conditions, respectively) while maintaining the luminance of all bars at 30 cd/m² (the same as in the main experiment). We recorded from a total of 21 cells from one monkey during this experiment.

Analyses of neurophysiological data. Data analyses were performed using custom written S-Plus (Seattle, WA) or Matlab (Natick, MA) programs. For each cell, the response to a given stimulus was averaged from the net firing rate of the cell across 10 correct repetitions. The net firing rate was determined by subtracting the background firing rate, calculated using a 200 msec time window immediately preceding the stimulus onset, from the evoked response of the cell. We systematically tested many different time windows for calculating the evoked response (window widths of 50–1000 msec, starting at 0–950 msec after the stimulus onset). The results were qualitatively similar across a wide range of window parameters (data not shown), although the differences among the various stimulus types were progressively less prominent with larger time windows, because of the fact that the effects of surround modulation were most prominent during the first 150 msec after the stimulus onset, and the overall response usually decayed rapidly after the initial response transient. For the data presented in this report, evoked responses were calculated using a 200 msec window starting 30 msec after the stimulus onset. We tested each of the 127 cells recorded during both experiments to determine whether the stimulus-evoked responses significantly differed from the background firing rate (two-tailed *t* test; $p < 0.05$). Eighty-five cells from the main experiment and 21 cells from the contrast effect experiment met this criterion and were used in this study.

Given the fact that stimuli were not directly comparable across different center bars (because both the center and surround bar types varied across rows, see Fig. 1), we generally analyzed the responses to stimuli

with each of the four center bars separately, except where it was appropriate to do otherwise, as noted.

Tests of significance. All tests of significance were performed using randomization. A test of significance using randomization consists of determining whether the value of a user-defined test statistic calculated from the actual data differs significantly from the distribution of the same test statistic calculated from randomized data (for review, see Manly, 1991). Briefly, an appropriate test statistic was first calculated using the actual data. The data were then randomized in a manner appropriate for the given test, and the test statistic was recalculated using the randomized data. The randomization process was repeated 10⁶ times (10³ times in case of dendrogram analyses), and the proportion of times the randomized test statistic exceeded the actual test statistic constituted the one-tailed probability *p* that the value of actual test statistic was indistinguishable from chance.

To perform a conventional test of significance using randomization, the corresponding test statistic [e.g., *t* statistic for *t* test, *F* ratio for ANOVA, and the *q* statistic for Tukey's honestly significant difference (HSD) test, etc.] calculated from the original data were compared with those calculated from the randomized data, and the *p* value was determined as described above. This procedure effectively corrects for deviations of data set from normality (Manly, 1991). To determine the statistical significance of a user-defined index (see next paragraph), the randomization procedure was repeated using the given index as the test statistic. To correct for multiple comparison artifacts, we used Tukey's HSD test (S-Plus function *multcomp*).

Neurophysiological indices. The pop-out stimulus selectivity index (PSI) for a given center bar was calculated as $(R_{\text{homogeneous}} - R_{\text{pop-out}})/R_{\text{homogeneous}}$, where $R_{\text{homogeneous}}$ was the response of the cell to the homogeneous stimulus, and $R_{\text{pop-out}}$ was the response to a given pop-out stimulus, the response to which differed most from (i.e., most suppressed or enhanced relative to) $R_{\text{homogeneous}}$. We used this (unsigned) magnitude of difference rather than either suppression or enhancement alone as the criterion, because the responses to the pop-out stimuli were suppressed in some cases and enhanced in others relative to the homogeneous stimulus with the same center. Note that either relative enhancement or suppression can potentially help distinguish between the two types of stimuli. Similarly, we calculated the corresponding conjunction-target stimulus selectivity index (CSI) as $(R_{\text{homogeneous}} - R_{\text{conjunction-target}})/R_{\text{homogeneous}}$, where $R_{\text{conjunction-target}}$ was the response of the cell to a given conjunction-target stimulus (with the same center bar) the response to which deviated most (i.e., most suppressed or enhanced) relative to $R_{\text{homogeneous}}$. For each cell, we calculated the PSI and CSI values for each of the four center bars, indicated in the subscript of the index by the appropriate center bar designation (e.g., PSI_{pref} and PSI_{null} denote PSI values calculated for the preferred and null center bars, respectively).

The pop-out preference index (PPI) for a given center directly compared the response of a cell to its most effective pop-out stimulus with its response to its most effective conjunction-target stimulus with the same center bar. The PPI for the preferred bar in the center was defined by $\text{PPI}_{\text{pref}} = (\text{PO}_{\text{pref}} - \text{CT}_{\text{pref}})/(\text{PO}_{\text{pref}} + \text{CT}_{\text{pref}})$, where PO_{pref} and CT_{pref} are the responses of the cell to its most effective pop-out and conjunction-target stimuli, respectively, with the preferred bar in the center. The PPI values with the other three center bar types were also calculated in a similar manner.

The response variation index (RVI) measured the modulation (i.e., variation) of given responses of a cell across a given subset of pop-out stimuli or conjunction-target stimuli with a given center bar. $\text{RVI}_{\text{pref,po}}$ measured the modulation of the responses of the cell across the five pop-out stimuli with the preferred center bar (stimuli A3-A7). To calculate $\text{RVI}_{\text{pref,po}}$, we first calculated the conventional *F* ratio (Snedecor and Cochran, 1989) of the responses of the cell across the five stimuli, defined as $F = \text{MS}_{\text{between}}/\text{MS}_{\text{within}}$, where $\text{MS}_{\text{between}}$ was the stimulus-to-stimulus variance, and $\text{MS}_{\text{within}}$ was the average trial-to-trial variance. $\text{RVI}_{\text{pref,po}}$ was defined as the *F* ratio calculated from the actual data divided by the average *F* ratio calculated from 10⁶ rounds of randomization. $\text{RVI}_{\text{pref,ct}}$ similarly measured the response modulation across the two conjunction-target stimuli with the preferred center bar (i.e., stimuli A1 and A2). RVI_{po} and RVI_{ct} values for the other three center bars were

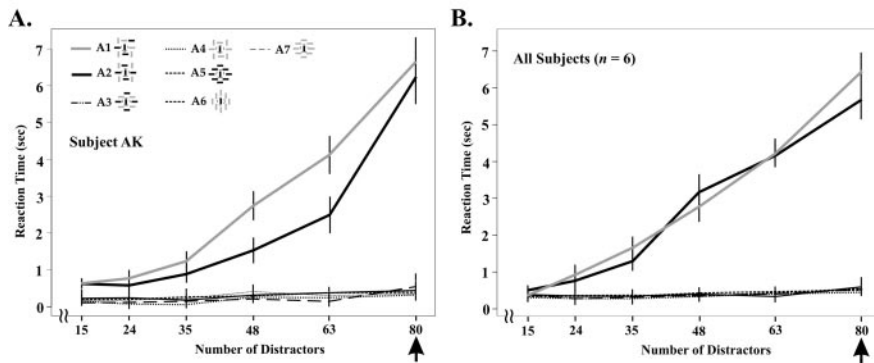


Figure 2. Reaction times for selected stimuli. Each of the 28 stimuli with non-homogeneous surrounds was tested in a visual search experiment in which human subjects searched for a single target among varying numbers of distractors (see Materials and Methods for details). In this figure, average reaction times (\pm within-group SEM) for eight selected stimuli (A1–A8) are shown as a function of the number of distractors for each stimulus for an individual subject (A) or for all subjects (B). For the data shown, the color and the orientation values used were the same as those for the exemplar cell shown in Figure 4A; the number of distractors used for that cell is denoted by the arrow. Similar results were obtained for each of the many other combinations of color and orientation values tested (data not shown).

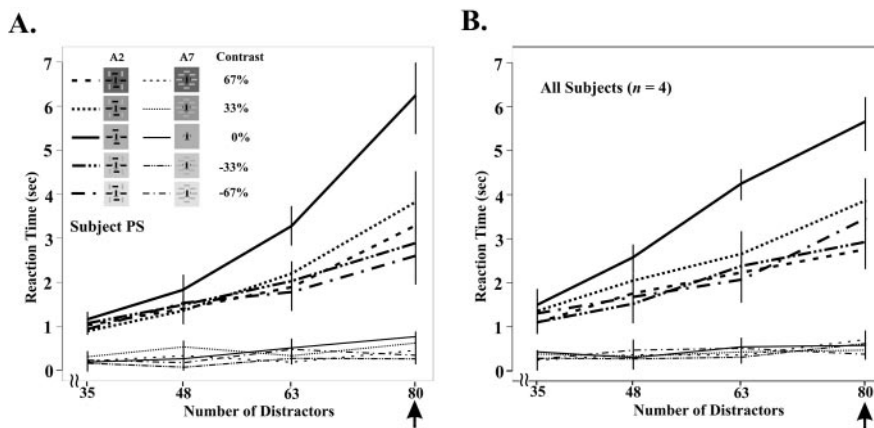


Figure 3. Effect of stimulus-background contrast on visual search. Representative pop-out and conjunction-target stimuli were retested at systematically varying stimulus-background contrasts (icons; see Materials and Methods for details). The resulting average reaction times (\pm within-group SEM) are shown for each contrast level as a function of the number of distractors in the given stimulus for an individual subject (A) or for all subjects (B). For the data shown, the color and the orientation values used were the same as those in the exemplar cell shown in Figure 11A; the arrow denotes the number of distractors in the stimuli used for that cell. Similar results were obtained for each of the many other combinations of color and orientation values tested (data not shown).

also calculated similarly. Note that a given RVI_{po} had much greater statistical power than the corresponding RVI_{ct} ($df = 4$ for RVI_{po} vs $df = 1$ for RVI_{ct}). The response modulation comparison test was essentially a randomized version of a two-way ANOVA with stimulus type (i.e., pop-out vs conjunction-target) and stimuli (i.e., of either type) as the two factors, used for determining whether the patterns of response significantly varied between pop-out and conjunction-target stimuli with a given center bar (Manly, 1991; p. 69). To pass this test for a given center stimulus, a cell had to have $p < 0.05$ for the stimulus type factor, the interaction factor, or both.

The average surround modulation index (ASM) was calculated for each cell as

$$\sum_{i=1}^n [(S_i - C_i)/C_i]/n,$$

where S_i and C_i are, respectively, the responses of the cell to the i th center-surround stimulus and the corresponding center alone stimulus. We calculated the absolute, and not the signed, difference between S_i and C_i because most cells in our sample were suppressed by some center-surround stimuli and enhanced by others.

The response modulation index (RMI) was calculated in the same manner as the RVIs above, except that the response modulation was measured across all 32 center-surround stimuli, instead of just the pop-out or conjunction-target stimuli with a given center bar.

Population analyses. To analyze patterns of response correlation across the population, we used hierarchical cluster analysis (HCA) (for review, see Kachigan, 1991). We used HCA (S-Plus routine *agnes*) to derive a graphical binary tree, or “dendrogram”, of the stimulus set, so that stimuli that elicited similar responses across the V1 cells population were clustered closer together on nearby branches, and those which elicited disparate responses were separated on distant branches.

We also analyzed the population response data using metric multidimensional scaling (S-Plus routine *cmdscale*), which plots the data so that the distances between the data points, in our case the stimuli, represents the similarity of the responses of V1 cells to the stimuli (Kruskal and Wish, 1978).

To construct the population average peristimulus time histograms (PSTHs), 10 msec bins spanning the 1400 msec interval around the stimulus presentation (with the stimulus onset and offset at 0 and 1000 msec, respectively) were used. For each cell, a PSTH was constructed separately for each of the 36 stimuli using these bins. The 36 PSTHs for each cell were then normalized, so that the bin during which the firing rate of the cell was maximal (for any stimulus/stimuli) had a value of 1.0. The normalized PSTHs were averaged for each stimulus individually across all cells.

Human psychophysics

Human psychophysical experiments were identical to the neurophysiological experiments except as noted. All stimuli were displayed on a 19 inch Sony Multiscan E500 monitor (but using the same screen settings). The luminance of the stimuli were adjusted to be the same as that used in the physiological experiments using a PhotoResearch PR650 Photometer. The color and orientation values used for constructing the psychophysical stimuli were either the same as those actually used for selected V1 cells or a random combination of values from the repertoire used for all cells. Not all parameter values were tested for all subjects, and only a representative subset of stimuli were tested for some subjects.

Trials were performed as previously described (Hegd  and Felleman, 1999) with minor modifications. Briefly, the subjects were instructed to search the stimulus for a single unique bar (“the odd man out”) and press a designated key when the search target was found. Stimuli were presented one per trial. Forty percent of the trials contained a pop-out stimulus, another 40% contained a conjunction-target stimulus, and the remaining 20% of the trials contained no target (catch trials). Subjects indicated the lack of target using a different key. No feedback was provided in any of the trials, and the “incorrect” trials were not repeated. Each subject performed at >95% across trials (data not shown).

Reaction time was measured for each stimulus type over four, five, or seven different bar array sizes, depending on the experiment. Each subject performed several practice trials, the data from which were discarded.

The “stimulus-background contrast experiment” was identical to main psychophysical experiment, except that stimulus-background contrast varied systematically from one condition to the next, whereas the

bars remained equiluminant with each other. In the “non-equiluminant” conditions, the luminance of the background was 33, 67, 133, or 167% of that of the bars, corresponding to a Weber contrast of +33, +67, –33, or –67%, respectively, between the stimuli and the background. In the “equiluminant” condition, the stimulus-background had a Weber contrast of 0%, as in the main experiment.

Analyses of reaction times. The reaction time for each stimulus was averaged across 40 repetitions presented over four sessions. We measured the dependence of reaction time t on the number of distractors n using either the Spearman’s correlation coefficient or the monotonic increment index (MII). To calculate MII, we first calculated MII_{raw} from the actual data, defined as $MII_{raw} = \sum \Delta t / \Delta n$, where Δt was increment in reaction time in milliseconds when the number of distractors increased by Δn . We then recalculated this index using data points from individual repetitions of given stimulus with a given number of distractors and calculated the root mean squared deviation between MII_{raw} from the actual data and the randomized data. We calculated the average of these differences, MII_{RMS} , from 10^6 rounds of randomization. MII was defined as MII_{raw} / MII_{RMS} , and its statistical significance was determined as described for the neurophysiological indices above. Note that MII explicitly takes trial-to-trial variations, including those caused by chance and/or practice effects, into account.

Results

Psychophysical testing of center-surround stimuli

The stimulus set included a representative collection of putative pop-out and conjunction-target stimuli (Fig. 1). To verify that these stimuli were able to produce the expected perceptual effects under our experimental conditions, we tested them in a visual search experiment using human subjects under experimental conditions similar to those used in the neurophysiological studies (see Materials and Methods). Figure 2A illustrates the reaction times of an individual subject to eight pop-out and conjunction-target stimuli with the same center bar, all presented at equiluminance. As expected from previous studies (Treisman, 1980, 1988; Wolfe, 1994), the reaction times for the conjunction-target stimuli (A1 and A2) increased monotonically with the number of distractors (Spearman’s correlation coefficient $r_s = 0.92$; $p < 0.05$). The MII (see Materials and Methods) value for these stimuli collectively was 56.1, indicating that the reaction time increased at an average rate of ~ 56 msec per distractor, when random variations in reaction time were accounted for. The MII value was statistically significant ($p < 0.05$) in each case. By contrast, the reaction time for pop-out stimuli (stimuli A3–A7) remained short and statistically unchanged as the number of distractors increased ($r_s = 0.06$, $MII = 0.004$; $p > 0.05$ in both cases). The results were similar for each of the other five subjects individually, and for the average response of all six subjects (Fig. 2B) (three-way ANOVA, subjects \times search type \times number of distractors, $p > 0.05$ for subjects). Similar results were also obtained with the other three types of center bar (bar types B, C, and D), and many different combinations of bar color and orientation (data not shown). These results confirm that our

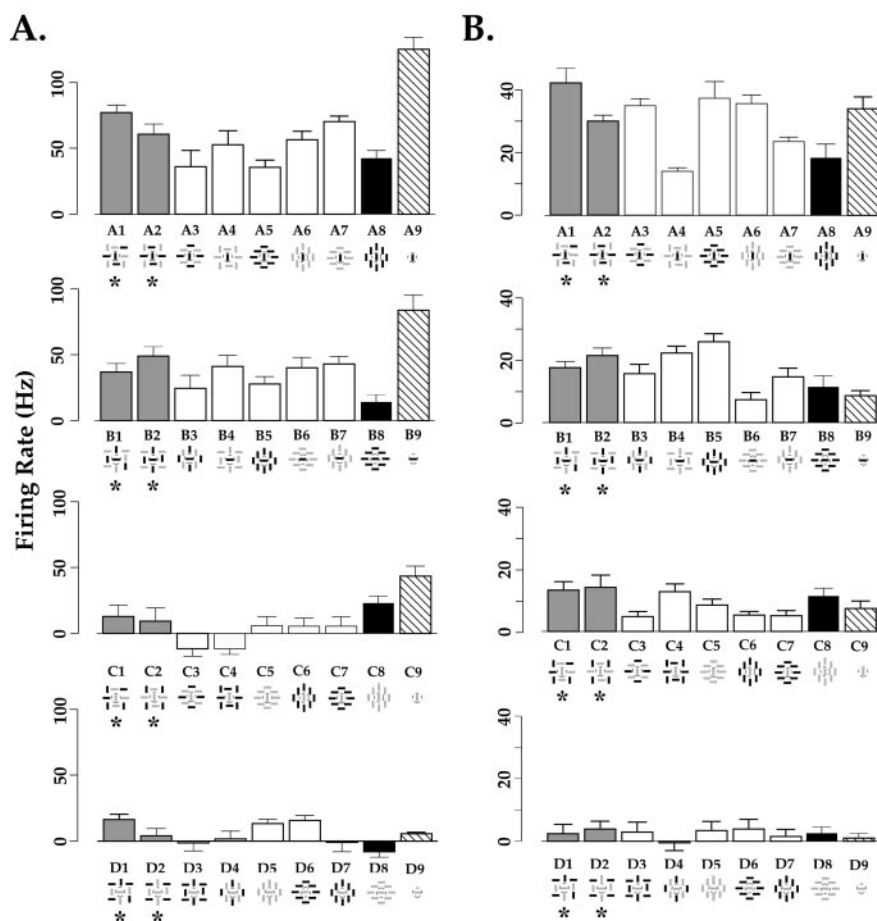


Figure 4. Responses of exemplar V1 cells to the stimulus set. *A* and *B* show the net responses of two individual V1 cells to each of the stimuli averaged across repetitions. Error bars indicate SEM. Negative firing rates represent suppression of responses below background levels. For clarity, the bar plots are filled differently according to the stimulus subclasses defined in Figure 1.

pop-out and conjunction-target stimuli were able to elicit the expected perceptual effects. They also demonstrate that the pop-out and conjunction-target effects persist for color and/or orientation at equiluminance (Luschow and Nothdurft, 1993).

Effects of stimulus-background contrast on visual search

As noted earlier, the neurophysiological experiments were performed under both the equiluminant and non-equiluminant conditions. To compare the reaction times elicited by pop-out and conjunction-target stimuli under the two types of conditions, we conducted a stimulus-background contrast experiment, in which we systematically varied the stimulus-background contrast while maintaining all stimulus elements at equiluminance (see Materials and Methods for details). Figure 3A shows the reaction times of a naive subject to a representative pop-out and conjunction-target stimulus each at five different stimulus-background contrasts. The reaction times to the pop-out stimulus (thin lines) were statistically indistinguishable across various contrasts (two-way ANOVA, luminance \times number of distractors, $p > 0.05$ for both factors and interaction), indicating that pop-out was not significantly affected by variations in stimulus-background contrast. On the other hand, for the conjunction-target stimuli (thick lines), the reaction times increased significantly faster for the equiluminant condition than for any of the non-equiluminant conditions (Tukey’s HSD test; $p < 0.05$), although the non-equiluminant conditions were indistinguishable among themselves (two-way ANOVA, conditions \times number of

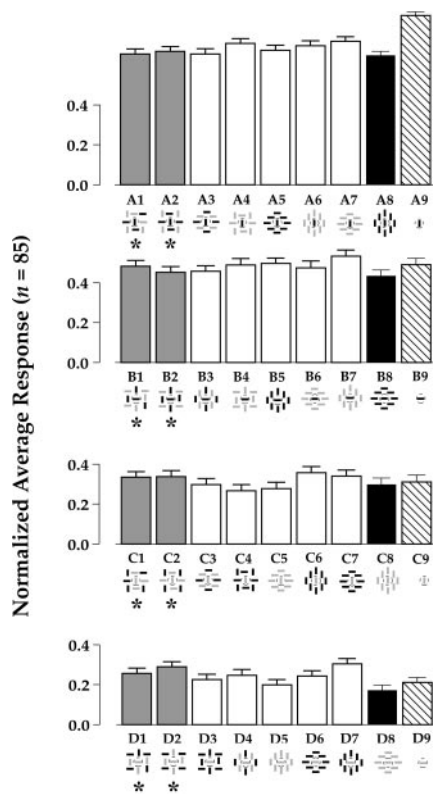


Figure 5. Population average response. The response of each neuron was normalized to a maximum of 1.0 and averaged across all cells, so that each bar represents the mean population response (\pm cell-to-cell SEM) to the corresponding stimulus. For clarity, the bar plots are filled differently according to the stimulus subclasses.

distractors, $p > 0.05$ for condition). We obtained similar results for the each of two other naive subjects (data not shown), and for all four subjects in the study together (Fig. 3B), but not for the one non-naive subject in the study (J.H., one of the authors), for whom the reaction times for the equiluminant condition were indistinguishable from those for the non-equiluminant conditions (Tukey's HSD test; $p > 0.05$; data not shown). Together, these results indicate that the psychophysical distinction between the pop-out and conjunction-target stimuli were at least as prominent, often more prominent, at equiluminance as at non-equiluminance. Thus, neurophysiological differences between the two sets of stimuli, if any, can be expected to be most evident under equiluminant conditions.

Neuronal responses to center-surround stimuli

We studied the responses of 85 cells from area V1 of two alert, fixating macaques to each of the 36 stimuli presented under equiluminant conditions (see Materials and Methods). Figure 4 shows the responses of two V1 cells to the stimuli. The cell shown in Figure 4A was selective for both color and orientation of the center bar (compare A9 vs B9 vs C9 vs D9; Tukey's HSD test; $p < 0.05$). For each of the four center bars (rows), the responses of the cell to the center-surround stimuli were mostly (but not always) suppressed relative to the corresponding center-alone stimulus, regardless of whether the center-surround stimuli were conjunction-target (asterisks) or pop-out type. The modulation of responses among the center-surround stimuli was significant for stimuli with each of the four center bars (one-way ANOVAs; $p < 0.05$ in each case).

There are many computationally meaningful ways of assessing the selectivity this cell for pop-out stimuli. As indicated earlier, many previous studies have compared the responses to a given pop-out stimulus with the responses to the corresponding homogeneous stimulus (Knierim and Van Essen, 1992; Lamme, 1995; Zipser et al., 1996; Nothdurft et al., 1999, 2000). We performed similar analyses, with corrections for multiple comparisons where appropriate, which took into account the fact that our stimulus set contained many more pop-out stimuli than homogeneous stimuli (see Materials and Methods). For each of the four center bars, the response of the cell to the homogeneous stimulus was statistically distinguishable from at least one pop-out stimulus with the same center bar (e.g., A8 vs A7; D8 vs D6, etc.; Tukey's HSD test; $p < 0.05$). Furthermore, the most effective stimulus of the cell with the null center bar was a pop-out stimulus (D6), as was the stimulus that elicited the largest surround suppression (C3). Thus, when the comparison is limited to pop-out and homogeneous stimuli, and conjunction target stimuli were excluded from the analysis, this cell was selective for one or more pop-out stimuli by many potentially meaningful criteria. Note, however, that none of these analyses distinguishes between selectivity for the pop-out per se and a mere selectivity for the existence of center-surround feature discontinuities in pop-out stimuli, absent from homogeneous stimuli.

A very different picture of the selectivity of the cell emerges when the responses to the conjunction-target stimuli are taken into account. In general, the cell responded similarly to many pop-out and conjunction-target stimuli, so that if the responses to the pop-out stimuli were excluded from the analysis, the cell could be judged to be selective for conjunction-target stimuli by many of the same criteria as those used to infer pop-out selectivity above. For instance, for every pop-out stimulus that elicited a response distinct from the response to the corresponding homogeneous stimulus (e.g., stimuli A7 vs A8, D6 vs D8), there was a conjunction-target stimulus that elicited a similar response (A1 and D1, respectively; Tukey's HSD test; $p > 0.05$). Nonetheless, the response of the cell to each of these conjunction-target stimuli was distinguishable from the corresponding homogeneous stimulus (Tukey's HSD test; $p < 0.05$). The fact that the cell responded similarly to two stimuli that both contained center-surround discontinuities but differed in terms of the nature of the discontinuities indicates that the responses of the cell likely reflected only a selectivity for the existence of the center-surround discontinuities.

The cell in Figure 4B illustrates some additional complexities of center-surround responses in V1 and the methodological challenges they pose for the analyses of pop-out selectivity. Like many cells in our sample, the responses of this cell were enhanced by some center-surround stimuli and suppressed by others, so that the cell was selective for either pop-out stimuli or conjunction-target stimuli depending on whether surround enhancement or surround suppression was the criterion used. For instance, a conjunction-target stimulus (A1) elicited the largest surround enhancement, whereas a pop-out stimulus (A4) elicited the largest surround suppression with the same center (A9). When only the magnitude of surround modulation was considered and the sign of the modulation was disregarded, the pop-out stimulus A4 elicited the largest surround modulation. However, the response modulation by the conjunction-target stimulus A1 had the largest magnitude when measured against the homogeneous stimulus, A8 (instead of the center-alone stimulus A9). Note that the same set of comparisons result in different conclusions for stimuli with the other three center bars (centers B-D). Thus, whether

a particular cell could be classified as pop-out-selective or conjunction-target-selective depended on the criterion used.

Figure 5 shows the average response of all 85 V1 cells to each stimulus. Note that V1 cells on average were surround-suppressed with the preferred bar in the center (top row), although with weaker bars in the center (rows 2–4), the surround modulation was generally, albeit modestly, facilitative, consistent with earlier results (Polat et al., 1998; Somers et al., 1998). Nevertheless, the population response to pop-out versus conjunction-target stimuli was indistinguishable for each of the four center bars (pairwise *t* tests; $p > 0.05$ in all cases). Although this analysis does not rule out the possibility that V1 contains a subpopulation of pop-out-selective cells, this scenario is unlikely, since the cell-to-cell variations (Fig. 5, error bars) were no larger for pop-out stimuli than for other stimuli (one way ANOVA across center-surround stimuli; $p > 0.05$).

Distinguishing pop-out selectivity from selectivity for center-surround feature discontinuities

As indicated above, pop-out-selective cells must, at a minimum, be able to distinguish pop-out stimuli from both homogeneous stimuli and conjunction-target stimuli. To determine the extent to which V1 cells can do this, we compared the response of each cell to a given homogeneous stimulus both with its response to a selected pop-out stimulus relative to homogeneous stimuli, we calculated a PSI for each of the four center bars for each cell (see Materials and Methods for details). The distribution of PSI values with the preferred bar in the center, PSI_{pref} is shown in the *x*-axis histogram of Figure 6. For about six-tenths of the cells (50 of 85, 59%) the PSI values were negative, indicating that for these cells a pop-out stimulus elicited a larger response than the corresponding homogeneous stimulus did. For the remaining four-tenths of the cells (41%), the response to the pop-out stimulus was lower than the response to the homogeneous stimulus (i.e., positive PSI values). The difference between the responses to pop-out versus homogeneous stimuli was statistically significant for approximately one-third of the cells (28 of 85, 33%; Tukey's HSD test; $p < 0.05$; filled bars), indicating that these cells were selective for pop-out or, alternatively, for the existence of the center-surround discontinuities. About two-thirds of the cells (59 of 85, 69%) were selective for a pop-out stimulus by this criterion for at least one of the four center bars (data not shown).

We similarly calculated a CSI (see Materials and Methods) for each center bar, which measured the selectivity of the cell for a conjunction-target stimulus relative to the homogeneous stimulus with the same center bar. The CSI values were statistically signifi-

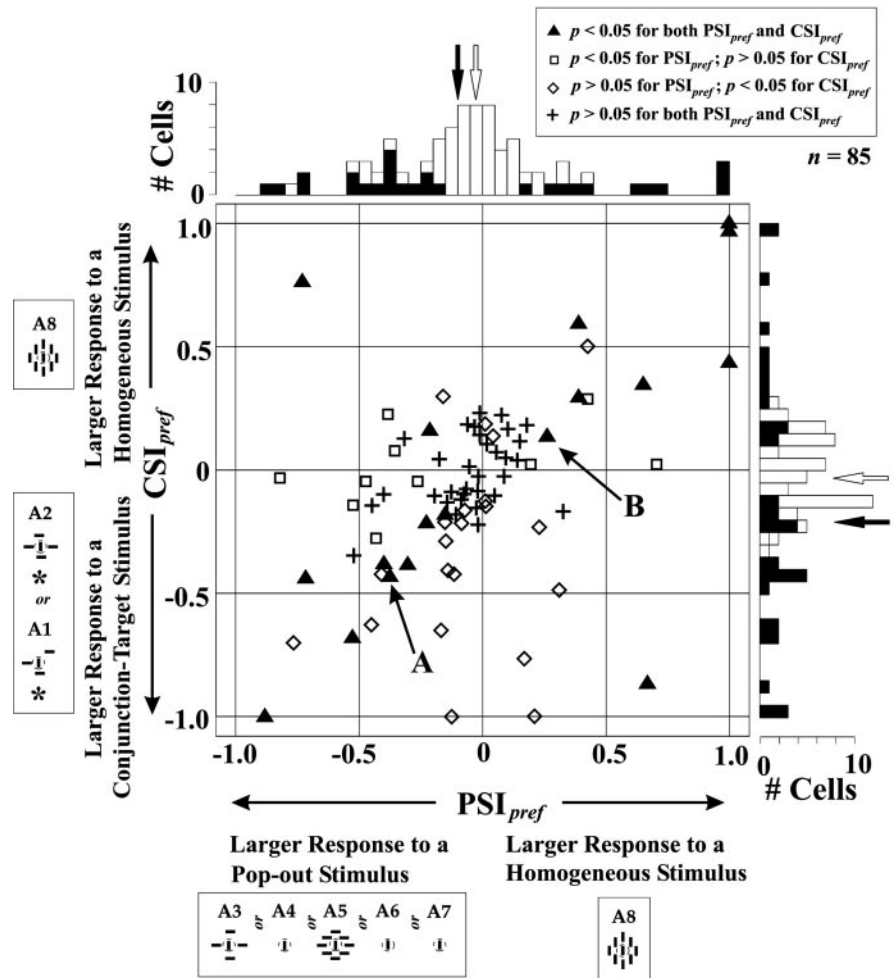


Figure 6. Comparison of responses to pop-out versus conjunction-target stimuli. The selectivity of each cell for pop-out or conjunction-target stimuli with the preferred bar in the center was measured using the selectivity indices PSI_{pref} and CSI_{pref} , respectively, as described in Materials and Methods. In the scatterplot, the PSI_{pref} value (*x*-axis) of each cell is plotted against its CSI_{pref} value (*y*-axis) according to whether the value of either index was statistically significant (see inset). Outliers with index values > 1 were normalized to 1. The histogram on either axis shows the distribution of the corresponding index values. The filled and open bars denote the cells for which the value of the corresponding index was statistically significant ($p < 0.05$) or insignificant, respectively. The filled and open arrows denote the corresponding sample means, calculated before the outliers were normalized. Note that negative values of PSI_{pref} and CSI_{pref} represent a preference for, and not response suppression by, pop-out and conjunction-search stimuli respectively, relative to the homogeneous stimuli. In this and subsequent figures, the cells shown in Figure 4, *A* and *B*, are denoted by the corresponding letters.

cant for ~44% of the cells (37 of 85; Tukey's HSD test, $p < 0.05$; filled bars in the *y*-axis histogram), indicating that these cells were able to distinguish between conjunction-target versus homogeneous stimuli.

The PSI_{pref} and CSI_{pref} values were moderately, but significantly, correlated with each other ($r = 0.48$; $df = 84$; $p < 0.05$), indicating that V1 cells tended to respond similarly to the given pop-out and conjunction-target stimuli. Only about two-fifths of the cells (18 of 85; 21%) were able to distinguish the homogeneous stimulus from both the corresponding pop-out and conjunction-target stimuli (Fig. 6, triangles in the scatterplot), but only two of these 18 cells were also able to distinguish between the corresponding conjunction-target and pop-out stimuli (Tukey's HSD test; $p < 0.05$), indicating that these cells were genuinely selective for the pop-out stimuli (color and orientation pop-out in each case; data not shown) with the preferred bar at center. Note that if pop-out selectivity were defined based solely on the basis of pop-out versus homogeneous stimulus compari-

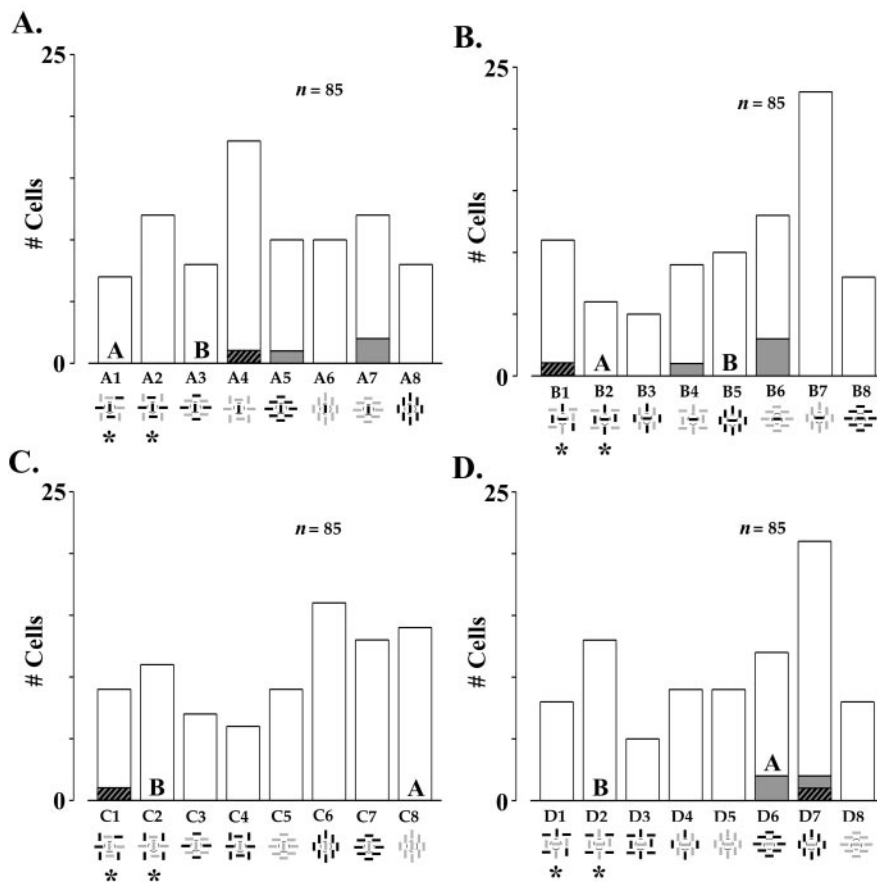


Figure 7. Preferred center-surround stimuli of V1 cells. V1 cells were classified according to their most effective center-surround stimulus with the given center bar. *A–D* show the results for each of the four center bars. The cells for which the response to the most effective stimulus with the given center bar was statistically distinguishable from the responses to the most effective stimuli from the other two stimulus subclasses with correction for multiple comparison artifacts (two-tailed Tukey's HSD test; $p < 0.05$; hatched bars) and without (two-tailed *t* test; $p < 0.05$; gray bars) are indicated. No cells had $p < 0.05$ using the Bonferroni correction for multiple comparison. See Materials and Methods for details.

son, a substantially larger proportion of cells (denoted collectively by squares and triangles in the scatter plot, corresponding to the filled bars in the *x*-axis histogram) would be designated as pop-out-selective.

We obtained qualitatively similar results when we repeated the above analyses using only the responses that were either suppressed or enhanced relative to the response to the homogeneous stimulus, although the selectivity for the various stimulus types, as expected, was generally more modest (data not shown). The results were also similar when we directly compared the responses of the cell to its most effective pop-out stimulus with its response to its most effective conjunction-target stimulus with the same center using the PPI (see Materials and Methods). The distribution of the PPI values for the 85 cells in this experiment will be presented later in a different context (see Fig. 12*A*).

Preferred center-surround stimuli of V1 cells

To determine the relative preponderance of selectivities for various types of stimuli in V1, we classified the cells according to their most effective center-surround stimulus with a given center bar. The distribution of the 85 cells that preferred each of the eight center-surround stimuli with the preferred bar in the center (stimuli A1–A8) are shown in Figure 7*A*. The proportions of cells which preferred different stimuli was indistinguishable from random (Kolmogorov–Smirnov test for Goodness of Fit; $p > 0.05$),

indicating that none of the eight stimuli was disproportionately effective for V1 cells. Although pop-out stimuli were the most effective subclass of stimuli, eliciting preferred responses from about two-thirds of the cells (58 of 85; 68%), this was statistically indistinguishable from that expected from the fact that five of the eight stimuli (62.5%) were pop-out stimuli (binomial proportions test; $p > 0.05$). The proportion of cells that preferred a stimulus from the other two subclasses of stimuli, the conjunction-target or homogeneous stimuli, were also similarly indistinguishable from random (20 and 9% cells, respectively; binomial proportions test, $p > 0.05$). Across all eight stimuli, the response to the given preferred stimulus was significantly larger than the responses to the most effective stimuli from the other two subclasses for only a small minority of cells, with correction for multiple comparison effects (Tukey's HSD test; $p < 0.05$; hatched bars) or without (*t* test; $p < 0.05$; gray bars). Importantly, the proportion of cells with $p < 0.05$ by either method was indistinguishable from that expected from chance at 5% significance level (i.e., probability of type I error $\alpha = 0.05$). Similar results were obtained for stimuli with the other three types of center bars (Fig. 7*B–D*). Together, these results indicate that V1 cells show no pronounced preference for any particular type of center-surround stimuli, including pop-out stimuli. Furthermore, the fact that few cells unambiguously prefer any given type of center-surround stimulus over others is consistent with a distributed, rather than local (i.e., explicit) coding of center-surround discontinuities.

Response modulation across pop-out versus conjunction-target stimuli

Given that individual V1 cells are unlikely to explicitly represent pop-out, we studied the extent to which the response profiles of V1 cells convey information about center-surround stimuli in general. To do this, we compared the response variation across pop-out stimuli versus across conjunction-search stimuli using the RVIs based on the conventional *F* ratio (see Materials and Methods). The *x*-axis histogram in Figure 8 shows the distribution of $RVI_{\text{pref,po}}$ values, which measured the response modulation of V1 cells across the five pop-out stimuli with the preferred center (stimuli A3–A7). The average $RVI_{\text{pref,po}}$ value was 1.78, indicating that the average response modulation across the five pop-out stimuli was ~ 1.78 -fold larger than that expected from chance. However, the response modulation was statistically significant for only a small number of cells (10 of 85, 12%; filled bars). The response modulation across the conjunction-search stimuli with the same center bar (stimuli A1 and A2), as measured by $RVI_{\text{pref,ct}}$ was also significant for only a few cells (9 of 85, 11%; filled bars in the *y*-axis histogram). The response modulation was significant for both sets of stimuli for five cells (6%; denoted by the overlapping filled triangles in the top righthand

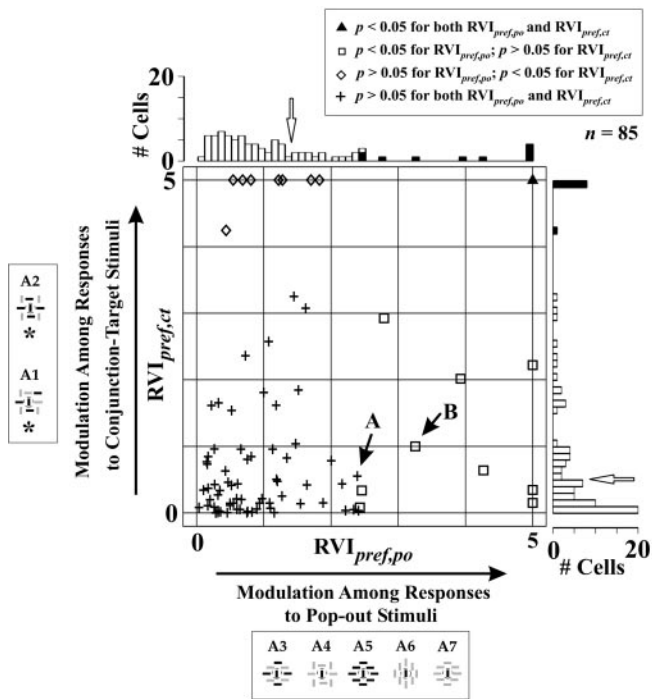


Figure 8. Response modulation across pop-out stimuli versus conjunction-target stimuli. For each cell, the response variation across all five pop-out stimuli with the preferred center bar was calculated using a response variation index ($RVI_{pref,po}$) as described in Materials and Methods. The corresponding index for the conjunction-target stimuli, $RVI_{pref,ct}$, was similarly calculated. The two indices are plotted against each other here using the same conventions as in Figure 6. The averages for the filled bars in the x - and the y -axis histograms were out of the histogram range at 8.56 and 13.84, respectively (data not shown).

corner of the scatterplot). The response modulation between the two sets of stimuli significantly differed for about one-quarter of the cells (22 of 85, 26%) of the cells, as measured by the response modulation comparison test, equivalent to a two-way ANOVA with stimulus type \times stimuli as the two factors (see Materials and Methods). Collectively, these results indicate that V1 cells convey a modest amount of information about center-surround stimuli.

Hierarchical cluster analysis of the population response

To determine whether V1 cells distinguish pop-out stimuli from conjunction-target stimuli at the population level, we used hierarchical cluster analysis, which groups the stimuli which elicit similar responses from the V1 cell population closer together than those that elicit dissimilar responses (see Materials and Methods for details). Figure 9 shows the results of this analysis in a dendrogram format, where the vertical distance between any two stimuli is a measure of how closely the population responses they elicited were correlated. Each of the four primary clusters contained all (and only) the stimuli with the same center bar type, indicating that the center bar was the most important determinant of the population response and the surround stimuli only played a modulatory role. Within the individual clusters, the segregation of the pop-out and the conjunction-target stimuli from each other was clearest for the stimuli with the preferred bar at center (far left branch), in that the vertical distance between the two conjunction-target stimuli (stimuli A1 and A2) was less than half the vertical distance to the nearest pop-out stimulus (stimulus A3; distances of 0.09 vs 0.23 U; data not shown). However, this separation was statistically insignificant ($p > 0.05$) as measured by D ratio, which measured the ratio of between-subcluster dis-

tances to within-subcluster distances in a manner similar to ANOVA (Hegd  and Van Essen, 2003). Moreover, the homogeneous stimulus (A8) elicited responses closest to the orientation pop-out (A5) stimulus. The separation of pop-out vs conjunction-target stimuli was statistically insignificant ($p > 0.05$) for the other three clusters as well. Similar results (data not shown) were obtained when the population response was analyzed using multidimensional scaling (MDS) (see Materials and Methods). Together, these results indicate that V1 cells as a population did not distinguish between the pop-out versus conjunction-target stimuli.

These results could, in principle, result from a scenario where different subpopulations of V1 cells are selective for pop-out versus conjunction-target stimuli, and these distinctions between the subpopulations average out across the overall population. However, both HCA and MDS failed to reveal any significant clustering among V1 cells based on center surround responses in general and differential responses to pop-out versus conjunction-target stimuli in particular (data not shown). As noted earlier, no such clustering was evident in the distributions of any of the index values either (Figs. 6, 8). Together, these observations indicate that the lack of pronounced pop-out selectivity in the overall V1 population was not because such selectivity was confined to a small but distinct subpopulation of pop-out-selective cells.

Time course of responses to pop-out versus conjunction-target stimuli

Because a defining psychophysical distinction between pop-out and conjunction-target stimuli is the reaction time, we studied whether the temporal dynamics of the corresponding neuronal responses also have distinguishing temporal characteristics. We performed two different analyses, one at the level of individual cells and the other at the population level.

Figure 10, *A* and *B*, shows the time course of the responses of the exemplar cell illustrated in Figure 4*A* to the nine stimuli with the preferred bar at center. The response time courses of the most effective conjunction-target versus pop-out stimuli (stimuli A1 and A7 respectively) (Fig. 4*A*) were statistically distinguishable (two-tailed Kolmogorov–Smirnov test for goodness of fit; $p < 0.05$) for this cell. The same was true, for at least one of the four center bars, for 23 (27%) of the 85 cells (data not shown). However, for all but two (2%) of the cells, the surround modulation (i.e., suppression or enhancement relative to the corresponding center bar alone) was evident in earlier bins for the conjunction-target stimulus than for the pop-out stimulus, contrary to what would be expected if the shorter reaction times for pop-out were correlated with shorter latencies for surround modulation.

The results were qualitatively similar when we repeated these analyses at the population level. Figure 10*C* shows the average time course of the V1 cell responses to the nine stimuli with the preferred bar in the center during -20 to 200 msec, spanning the interval during which the responses were most distinctive from each other. Whereas the responses to the center surround stimuli as a group were distinguishable from the responses to the center alone during each bin within the 40 – 150 interval (binwise t tests; $p < 0.05$ in all cases), the responses to the center-surround stimuli among themselves were not (binwise one-way ANOVAs; $p > 0.05$ in all cases). Furthermore, the time courses of responses to the most effective conjunction-target and pop-out stimuli for the V1 cell population (stimuli A2 and A7, respectively) (Fig. 5) were statistically indistinguishable from each other during any bin within

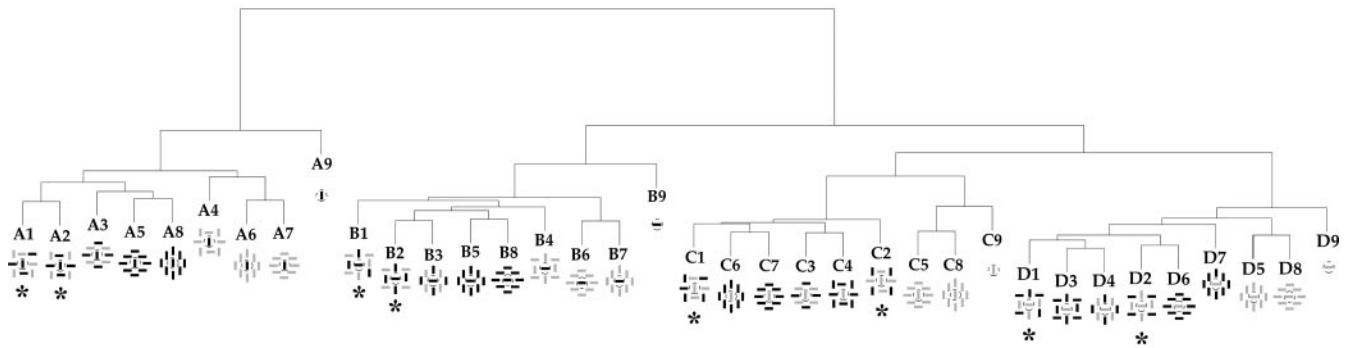


Figure 9. Analysis of population response patterns. Hierarchical cluster analysis was used to arrange the stimuli in a dendrogram so that the vertical distance between any two (sets of) stimuli is a measure of the similarity of responses they elicit from the V1 cell population. See Materials and Methods for details.

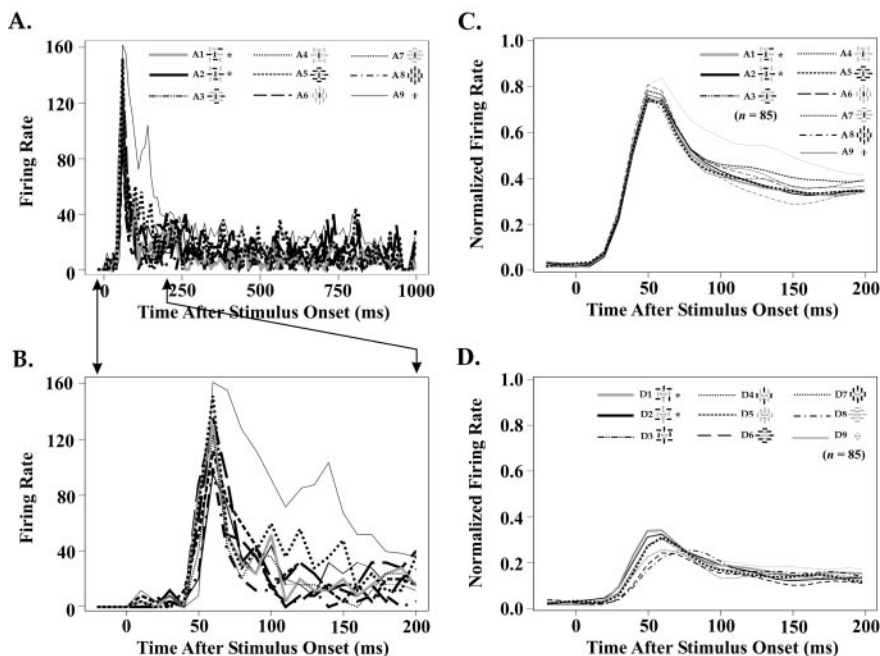


Figure 10. The time course of surround modulation. *A, B*, Response PSTHs of the exemplar cell illustrated in Figure 4*A* for selected stimuli (inset). *C, D*, The normalized population average PSTHs. In all cases, the stimulus onset and offset were at 0 and 1000 msec, respectively. *B–D* show 0–200 msec interval of this period (as illustrated for the exemplar cell by the double arrows between *A* and *B*), during which the distinction among the PSTHs was most prominent for both the exemplar cell and for the population. Using larger time windows did not qualitatively alter the results. See Materials and Methods for details.

the -200 to 1200 msec interval (two-tailed Kolmogorov–Smirnov test for goodness of fit; $p > 0.05$; data not shown). With the null bar in the center (Fig. 10*D*), the results were similar, except that both the response onset and the modulatory effects were less pronounced in magnitude and had a slightly longer latency, although the time course of responses to the various stimuli were statistically indistinguishable from each other (binwise one-way ANOVAs, $p > 0.05$ in all cases). Similar results were obtained for stimuli with the other two remaining center bar types (data not shown). Together, the above results indicate that the differences in the psychophysical reaction times for pop-out versus conjunction-target stimuli are not reflected in the temporal dynamics of the V1 center-surround responses.

Controls

The degree of surround modulation

In principle, the lack of pronounced selectivity for pop-out stimuli could arise from a lack of surround modulation, such that the surround stimuli were too ineffectual to elicit distinct responses

to pop-out stimuli. To address this issue, we measured the average surround modulation of all V1 cells using the ASM. The average ASM value of V1 cells was 0.24 (excluding outliers with $ASM > 1.0$ [$n = 82$]; average ASM for all 85 cells, 1.73; data not shown), indicating that the response of V1 cells to center-surround stimuli was on average about one-quarter larger or smaller (depending on the stimulus) than the corresponding center stimulus alone. Importantly, the ASM values of the cells were not correlated with their pop-out selectivity as measured by the PPI ($r = 0.04$; data not shown), indicating that the lack of pop-out selectivity was not a result of lack of surround modulation.

Response modulation across all center-surround stimuli

To ascertain that the responses of V1 cells in our sample were modulated across the various center-surround stimuli, we calculated the RMI for each cell, which measured the non-random variation in the responses of the cell across all 32 center-surround stimuli (see Materials and Methods). The average response modulation was 1.34, indicating that the response of V1 cells was modulated on

average 1.34-fold above chance levels. The response modulation was statistically significant ($p < 0.05$) for approximately one-half of the cells (46 of 85, 54%; data not shown).

We also tested whether the lack of observed selectivity for the pop-out stimuli was attributable to high levels of noise (i.e., trial-to-trial variation) in the data. We found that the noise levels in our data were comparable to, and frequently lower than, those reported for V1 cells by many previous studies (Dean, 1981; Vogels and Orban, 1991; Gur et al., 1997; data not shown). Moreover, the noise levels were indistinguishable between the responses to pop-out versus conjunction-target stimuli (two-tailed t test; $p > 0.05$; data not shown), indicating that the lack of pronounced selectivity for pop-out stimuli was not attributable to larger noise levels in the pop-out responses. More importantly, this means that the lack of pop-out selectivity is unlikely to have been a consequence of lack of statistical power in our dataset, because increasing the power, i.e., sampling the neuronal re-

sponses over a larger number of repetitions, is likely to improve the observed overall selectivity of V1 cells to both pop-out and conjunction-target stimuli.

Effects of stimulus-background contrast on surround modulation

Thus far, we have described neuronal responses to center-surround stimuli presented against an equiluminant background. As described earlier (Fig. 3), the psychophysical distinction between the pop-out and the conjunction-target stimuli is usually more prominent when the stimuli are presented against an equiluminant background than when presented against a higher-contrast background. Thus, if the activity of a given V1 cell reflected a genuine selectivity for pop-out stimuli, this selectivity may be expected to diminish under conditions that diminish the perceptual distinction between the pop-out and the conjunction-target stimuli.

To explore this possibility, we performed a second experiment, the contrast effect experiment (see Materials and Methods), in which we recorded the responses of an additional 21 V1 cells to 18 selected stimuli (Fig. 1, rows A and B, stimuli A1–A9 and B1–B9) presented at two stimulus-background contrasts: 0% (or equiluminant condition, as in the main experiment) or 33% contrast (higher contrast condition). The 21 cells we studied during this experiment were indistinguishable from the 85 cells in the main experiment in terms their average firing rate during the equiluminant conditions and average surround-modulation as measured by the ASM index (t tests, $p > 0.05$ in each case; data not shown), indicating that two sets of cells sampled the same parent population.

Figure 11 A–C shows the responses of three individual V1 cells to each of the 18 stimuli in the equiluminant and high contrast conditions (open and filled bars, respectively). For each cell, the responses were significantly modulated across the center-surround stimuli with either center bar under either contrast condition (one-way ANOVAs; $p < 0.05$ in all cases). Relative to the responses in the equiluminant condition, the responses in the higher contrast condition were generally enhanced for the cell in A, generally suppressed for the cell in B, and enhanced for some stimuli and suppressed for other stimuli for the cell in C. However, none of the three cells was selective for pop-out stimuli by any of the criteria outlined above for the main experiment. Across all 21 cells (D), the responses to stimuli under equiluminant versus higher contrast conditions were statistically indistinguishable (two-way ANOVA, conditions \times stimuli, using non-normalized data, not shown; $p > 0.05$ for condition and interaction factors), although the responses were modulated significantly across the stimuli under both conditions ($p < 0.05$ for the stimulus factor), indicating that the neuronal responses in general did not reflect the psychophysical distinction between the two conditions. These results also indicate that the lack of unambiguous pop-out selectivity in the main experiment was not an artifact of low stimulus-background contrast.

To compare the selectivity of a given cell for pop-out stimuli

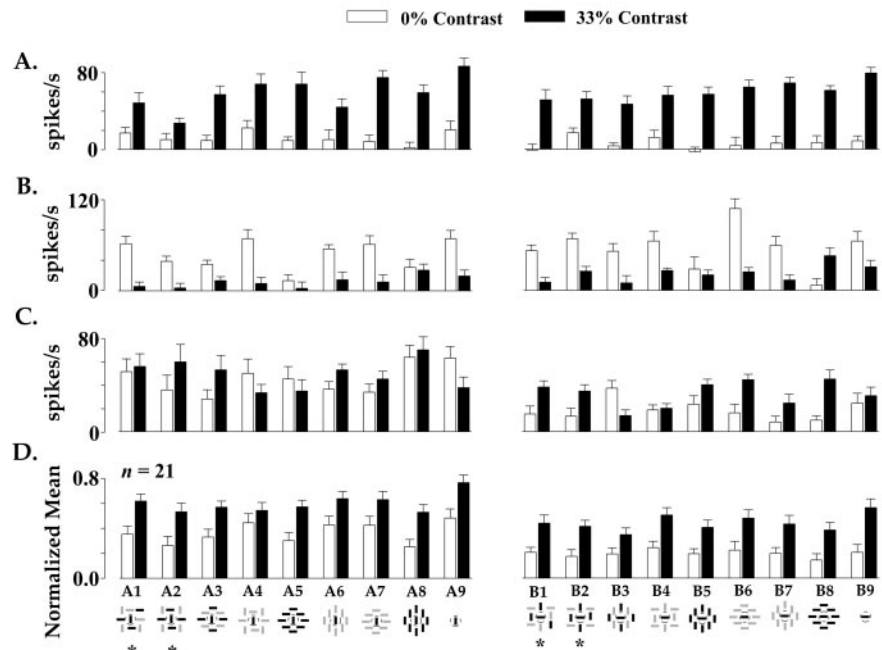


Figure 11. Effects of stimulus-background contrast on center-surround responses in V1. Stimuli A1–A9 and B1–B9 were presented against equiluminant or higher-contrast backgrounds (0 or 33% contrast, respectively). A–C show the responses of three individual cells. D shows the normalized population average of all 21 cells in this experiment. See Materials and Methods for details.

under the equiluminant (eq) versus higher contrast (hc) conditions, we calculated a PPI for either condition with the preferred bar at center ($PPI_{pref,eq}$ and $PPI_{pref,hc}$, respectively). The $PPI_{pref,eq}$ values of these 21 cells were indistinguishable from the corresponding $PPI_{pref,eq}$ values for the 85 cells in the main experiment calculated in an identical manner (t test; $p > 0.05$) (Fig. 12A), indicating that observed pop-out selectivity in the two experiments was similar under equiluminant conditions. For the 21 cells in the contrast effect experiment, $PPI_{pref,eq}$ and $PPI_{pref,hc}$ values were poorly correlated with each other (Fig. 12B) (correlation coefficient $r = -0.12$; $p > 0.05$), and were statistically indistinguishable from each other (paired t test; $p > 0.05$), indicating that the selectivity of individual neurons for pop-out stimuli did not vary with the stimulus-background contrast, contrary to what would be expected if the responses of these cells reflected a genuine selectivity for the pop-out stimuli or the pop-out percept.

Discussion

Selectivity of V1 cells to pop-out and conjunction-target stimuli

We have found that V1 cells typically respond similarly to pop-out and conjunction-target stimuli. The responses to the two types of stimuli were widely indistinguishable from each other by any of the many response measures. The inability of V1 cells to robustly distinguish between the two types of stimuli is not attributable to a lack of responsiveness to center-surround stimuli or to a lack of surround modulation itself. Our results indicate that center-surround responses in V1 do not provide a strong, explicit representation of pop-out in particular, or of figure-ground segregation in general.

On the other hand, the responses of V1 cells were substantially modulated among center-surround stimuli, indicating that information about many types of center-surround is represented collectively across the V1 cell population. Furthermore, the pro-

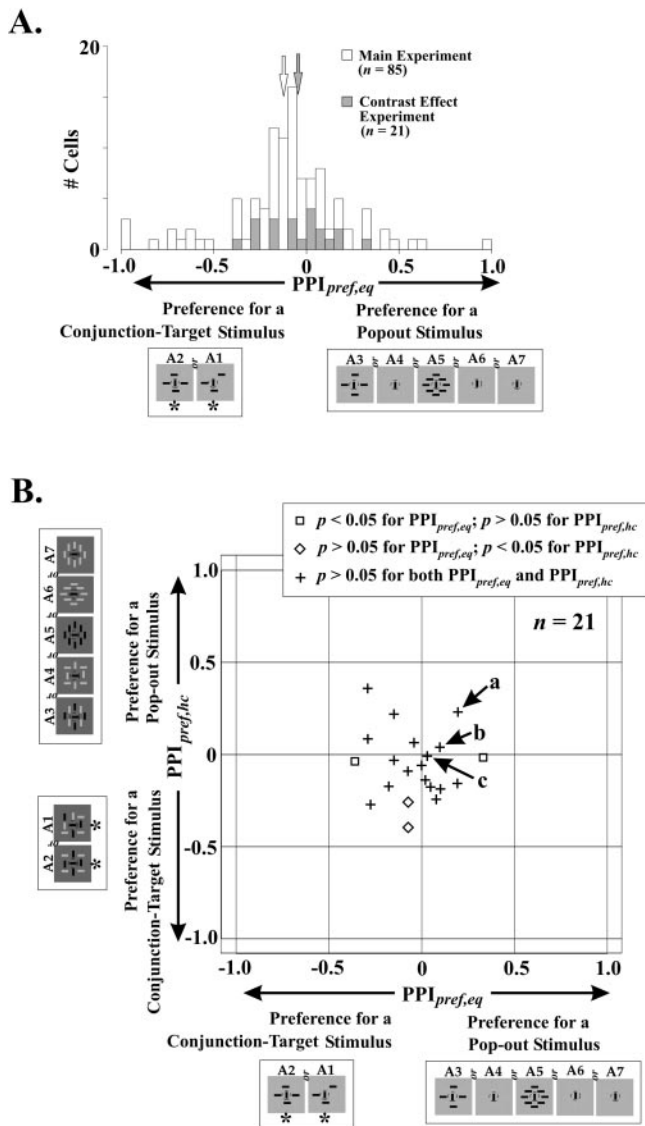


Figure 12. Preference for pop-out stimuli with equiluminant versus high contrast backgrounds. *A*, The responses to pop-out versus conjunction-target stimuli with the preferred bar at center at equiluminance was compared for each cell using a pop-out preference index (PPI_{pref,eq}) as described in Materials and Methods. The joint distribution of PPI_{pref,eq} values from the cells in the main experiment (open bars) and the contrast effect experiment (filled bars) is shown. The open and the filled arrows denote the corresponding sample means. *B*, Comparison of pop-out selectivity under equiluminant versus high stimulus-background contrasts. For each of the 21 cells in the contrast effect experiment, a PPI value was calculated under equiluminant versus high stimulus-background contrasts (PPI_{pref,eq} and PPI_{pref,hc}, respectively) and plotted against each other in this plot using different symbols according to whether one or both PPI values were statistically significant ($p < 0.05$) for a given cell. The exemplar cells shown in Figure 11 *A–C* are denoted by the corresponding lowercase letters *a*, *b*, and *c*. Note that in either panel, negative PPI values represent a preference for pop-out stimuli over conjunction-target stimuli, and not response suppression by pop-out stimuli.

portions of cells that preferred different types of center-surround stimuli were comparable, consistent with a distributed representation of many different types of feature discontinuities. Further processing of the feature discontinuity information, presumably in higher visual areas, is likely needed before an explicit representation of pop-out emerges.

Previous studies of pop-out selectivity in V1

Our results differ from previous studies of pop-out selectivity in area V1 in many important ways. Some previous studies have

suggested that pop-out-selective cells in V1 represent neural correlates of perceptual pop-out and, more generally, of figure-ground segregation (Lamme, 1995; Zipser et al., 1996; Kastner et al., 1997; Nothdurft et al., 1999, 2000). For the reasons outlined above, our results suggest this is not likely to be the case (Hochstein and Ahissar, 2002).

On a more sensory level, a large number studies have reported that V1 cells are selective for pop-out stimuli (Knierim and Van Essen, 1992; Lamme, 1995; Zipser et al., 1996; Kastner et al., 1997; Nothdurft et al., 1999, 2000) (for review, see Albright and Stoner, 2002). Our results find little evidence for this. We believe the discrepancy may be in large part attributable to the fact that many of these studies assessed the selectivity for pop-out stimuli relative to the homogeneous stimuli. We show that this comparison, taken by itself as evidence of pop-out selectivity, tends to greatly overestimate the preponderance of pop-out selectivity, because it does not account for the proportion of cells which are selective for the existence of feature discontinuities in general, without being specifically selective for pop-out. We found that most V1 cells fail to distinguish pop-out stimuli from conjunction-target stimuli, even when they distinguish both from homogeneous stimuli.

The above observations underscore the importance of differentiating nonspecific selectivity for center-surround feature discontinuities from selectivity for pop-out per se. The two issues cannot be clearly distinguished using discontinuities in single feature (e.g., orientation only), as most of the previous studies of pop-out selectivity have done, because in this case, feature discontinuity by itself results in pop-out. Two or more visual features are needed (e.g., orientation and color), so that the figure shares one of the features with some background elements and other feature(s) with other background elements. This strategy of separating pop-out from feature discontinuity is novel to our study.

Another possible explanation for the discrepancy between our results and those from many earlier studies is the fact that our animals passively viewed the stimuli, as opposed to performing a visual search task. However, we believe this is unlikely to have been a contributing factor, for three main reasons. First, previous studies have shown that behavioral tasks do not alter the ability of V1 neurons to differentiate between pop-out versus non-pop-out stimuli (Rossi et al., 2001; Marcus and Van Essen, 2002). Second, many studies have reported pop-out selectivity in area V1 of passively fixating or even anesthetized animals (Lamme et al., 1998; Nothdurft et al., 1999, 2000). Third, as we indicate throughout this report, pop-out stimuli are distinguishable from conjunction-target stimuli at a purely sensory level in terms of the nature of the feature discontinuities between the center and the surround. Thus, a cell selective for pop-out stimuli without reflecting the pop-out percept can still be expected to distinguish between pop-out and conjunction-target stimuli at a purely sensory level. Collectively, these observations indicate that center-surround responses in V1 do not explicitly represent pop-out stimuli at a purely sensory level either.

Recently, Rossi et al. (2001) re-examined the selectivity of V1 cells for figure-ground discontinuities of the type reported by Lamme (1995) using similar stimulus conditions and found little evidence of selectivity when the figure-ground border was $>1^\circ$ away from the CRF. When the figural border was closer to or within the CRF, however, many V1 cells distinguished orientation-defined figure-ground borders from uniform textures. Similar selectivity for feature discontinuities located near or within the CRF has been reported many others as well (Lee et al.,

1998; Das and Gilbert, 1999). But as noted earlier, although these results clearly represent selectivity for feature discontinuity between the center and the surround, it is unclear whether they necessarily reflect a selectivity for pop-out. In light of our results, unambiguous selectivity for pop-out, near or away from the CRF, still remains to be demonstrated in V1 or elsewhere in the visual cortex.

Psychophysical studies of pop-out

Mechanisms of pop-out and conjunction target search have also been addressed by many psychophysical models of visual search (Treisman, 1985, 1988, 1999; Treisman and Gormican, 1988; Duncan and Humphreys, 1989; Treisman and Sato, 1990; Cohen and Ivry, 1991; Cohen and Rafal, 1991; Wolfe, 1994, 1999). Unfortunately, it is difficult to determine specific implications of our results for any of these models (or vice versa), because none of the models propose (or for that matter claim to propose) specific, falsifiable hypotheses about how pop-out is represented neurophysiologically in V1 or elsewhere in the visual system. However, it is worth noting that some of the models, most notably those of Treisman (1985, 1988, 1999) and Wolfe (1994, 1999), appear to rely critically on local (as opposed to distributed) representations of feature values called “feature maps”. Feature discontinuities, such as those in a pop-out stimulus, are explicit from this map, because “pop-out for a target defined by a single distinctive feature is mediated by the unique activity it generates in the relevant feature map” (Treisman, 1988, p. 205; Wolfe, 1994, p. 204). No neural correlates of feature maps have been found (Treisman, 1999). More importantly for the present context, this mechanism of pop-out appears incompatible with a distributed representation of pop-out in V1 or elsewhere. Given the computational feasibility and neurophysiological plausibility of distributed representation of feature-discontinuities (our results; Rossi et al., 2001; Marcus and Van Essen, 2002), it is worth re-examining whether explicit representations of feature discontinuities are a prerequisite for pop-out to occur.

In conclusion, the notion that area V1 contains correlates of pop-out or figure–ground segregation is likely too simplistic. A more nuanced perspective may be that V1 plays an important role in the early processing of the feature discontinuities as a part of distributed network.

References

- Albright TD, Stoner GR (2002) Contextual influences on visual processing. *Annu Rev Neurosci* 25:339–379.
- Cohen A, Ivry RB (1991) Density effects in conjunction search: evidence for a coarse location mechanism of feature integration. *J Exp Psychol Hum Percept Perform* 17:891–901.
- Das A, Gilbert CD (1999) Topography of contextual modulations mediated by short range interactions in primary visual cortex. *Nature* 399:655–661.
- Dean AF (1981) The variability of discharge of simple cells in the cat striate cortex. *Exp Brain Res* 44:437–440.
- Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. *Psychol Rev* 96:433–458.
- Gur M, Beylin A, Snodderly DM (1997) Response variability of neurons in primary visual cortex (V1) of alert monkeys. *J Neurosci* 17:2914–2920.
- Hegd  J, Felleman DJ (1999) The pop-out in some conjunction searches is due to perceptual grouping. *NeuroReport* 10:143–148.
- Hegd  J, Van Essen DC (2003) Strategies of shape representation in primate visual area V2. *Vis Neurosci* 20:313–328.
- Hochstein S, Ahissar M (2002) View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36:791–804.
- Julesz B (1984) A brief outline of the texton theory of human vision. *Trends Neurosci* 7:41–48.
- Kachigan SK (1991) Multivariate statistical analysis. New York: Radius.
- Kastner S, Nothdurft H-C, Pigarev N (1997) Neuronal correlates of pop-out in cat striate cortex. *Vision Res* 37:371–376.
- Knierim JJ, Van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67:961–980.
- Kruskal JB, Wish M (1978) Multidimensional scaling. Newbury Park, CA: Sage Publications.
- Lamme VAF (1995) The neurophysiology of figure–ground segregation in primary visual cortex. *J Neurosci* 15:1605–1615.
- Lamme VAF, Zipser K, Spekreijse H (1998) Figure–ground activity in primary visual cortex is suppressed by anesthesia. *Proc Natl Acad Sci USA* 95:3263–3268.
- Lee TS, Mumford D, Romero R, Lamme VA (1998) The role of the primary visual cortex in higher level vision. *Vision Res* 38:2429–2454.
- Lee TS, Yang CF, Romero RD, Mumford D (2002) Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nat Neurosci* 5:589–597.
- Levitt JB, Lund JS (1997) Contrast dependence of contextual effects in primate visual cortex. *Nature* 387:73–76.
- Li W, Thier P, Wehrhahn C (2000) Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *J Neurophysiol* 83:941–954.
- Luschow A, Nothdurft HC (1993) Pop-out of orientation but no pop-out of motion at isoluminance. *Vision Res* 33:91–104.
- Manly BFJ (1991) Randomization and Monte Carlo methods in biology. New York: Chapman and Hall.
- Marcus DS, Van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88:2648–2658.
- Nothdurft H-C, Gallant JL, Van Essen DC (1999) Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia. *Vis Neurosci* 16:15–34.
- Nothdurft HC, Gallant JL, Van Essen DC (2000) Response profiles to texture border patterns in area V1. *Vis Neurosci* 17:421–436.
- Palmer J, Verghese P, Pavel M (2000) The psychophysics of visual search. *Vision Res* 40:1227–1268.
- Polat U, Mizboe K, Pettet MW, Kasamatsu T, Norcia AM (1998) Collinear stimuli regulate visual responses depending on cell’s contrast threshold. *Nature* 391:580–584.
- Rossi AF, Desimone R, Ungerleider LG (2001) Contextual modulation in primary visual cortex of macaques. *J Neurosci* 21:1698–1709.
- Sillito AM, Cudeiro J, Murphy PC (1995) Visual cortical mechanisms detecting focal discontinuities. *Nature* 378:492–496.
- Snedecor GW, Cochran WG (1989) Statistical methods. Ames, IA: Iowa State UP.
- Snodderly DM, Gur M (1995) Organization of striate cortex of alert, trained monkeys (*Macaca fascicularis*): ongoing activity, stimulus selectivity, and widths of receptive field activating regions. *J Neurophysiol* 74:2100–2125.
- Somers DC, Todorov EV, Siapas AG, Toth LJ, Kim DS, Sur M (1998) A local circuit approach to understanding integration of long-range inputs in primary visual cortex. *Cereb Cortex* 8:204–217.
- Treisman A (1980) A feature-integration theory of attention. *Cognit Psychol* 12:97–136.
- Treisman A (1985) Preattentive processing in vision. *Comput Vision Graphics Image Processing* 31:156–177.
- Treisman A (1988) Features and objects: the fourteenth Bartlett memorial lecture. *Q J Exp Psychol* 40A:201–237.
- Treisman A (1999) Solutions to the binding problem: progress through controversy and convergence. *Neuron* 24:105–110:111–125.
- Treisman A, Gormican S (1988) Feature analysis in early vision: Evidence from search asymmetries. *Psychol Rev* 95:15–48.
- Treisman A, Sato S (1990) Conjunction search revisited. *J Exp Psychol* 16:459–478.
- Van Essen DC, Zeki S (1978) The topographic organisation of rhesus monkey prestriate cortex. *J Physiol (Lond)* 277:193–226.
- Vogels R, Orban GA (1991) Quantitative study of striate single unit responses in monkeys performing an orientation discrimination task. *Exp Brain Res* 84:1–11.
- Wolfe JM (1994) Guided search 2.0: a revised model of visual search. *Psychon Bull Rev* 1:202–238.
- Wolfe JM (1999) The psychophysical evidence for a binding problem in human vision. *Neuron* 24:11–17:111–125.
- Zipser J, Lamme VAF, Schiller P (1996) Contextual modulation in primary visual cortex. *J Neurosci* 16:7376–7389.