Spatial Resolution for Feature Binding Is Impaired in Peripheral and Amblyopic Vision

Peter Neri and Dennis M. Levi
School of Optometry and Helen Wills Neuroscience Institute, University of California, Berkeley, California

Submitted 1 December 2005; accepted in final form 17 January 2006

Neri, Peter and Dennis M. Levi. Spatial resolution for feature binding is impaired in peripheral and amblyopic vision. J Neurophysiol 96: 142–153, 2006. First published January 18, 2006; doi:10.1152/jn.01261.2005. We measured spatial resolution for discriminating targets that differed from nearby distractors in either color or orientation or their conjunction. In the fovea of normal human observers, whenever both attributes are big enough to be individually visible, their conjunction is also visible. In the periphery, the two attributes may be visible, but their conjunction may be invisible. We found a similar impairment in resolving conjunctions for the fovea of deprived eyes of humans with abnormal visual development (amblyopia). These results are quantitatively explained by a model of primary visual cortex (V1) in which orientation and color maps are imperfectly co-registered topographically. Our results in persons with amblyopia indicate that the ability of the fovea to compensate for this poor co-registration is consolidated by visual experience during postnatal development.

INTRODUCTION

We perceive individual objects as possessing multiple attributes: a red ball falls behind a brown fence. This perceptual representation combines four different attributes (color, shape, motion, and depth) and provides an example of feature binding (Treisman 1998).

Feature binding can fail (Treisman 1996). When this occurs, normal observers experience illusory conjunctions of physically disjunct features (Treisman and Schmidt 1982; Wolfe and Cave 1999). The failure is chronic in patients with subcortical (Ward et al. 2002) but particularly cortical (Cohen and Rafa1 1991) damage like visual neglect (Eglin et al. 1989; Estermann et al. 2000) and Balint’s syndrome (Friedman-Hill et al. 1995; Robertson et al. 1997), and there is some evidence for impaired binding in schizophrenia (Alain et al. 2002). Our goal was to identify and measure aspects of the visual stimulus that lead to selective failure of feature binding in both normal and pathological vision.

We focused on the ability to resolve small visual items. We chose this measure for two reasons. First, the assessment of resolution has a well-established history in visual psychophysics and can be easily related to the physiological concept of neuronal receptive fields (RFs) (Levi and Klein 1985). Second, it pertains directly to a major issue and currently active debate in contemporary neuroscience: how the representation of space (and its resolution) “glues” together separate representations of object properties (Treisman 1998; Yu et al. 2005). We opted for a design that allowed us to factor out potential effects of crowding (Parkes et al. 2001), working memory (Luck and Vogel 1997), and coarse attentional deployment (Treisman 1998; Yu et al. 2005). We achieved this by normalizing the spatial resolution for binding features (orientation and color) by the resolution for discriminating features individually (orientation or color), in almost identical stimuli.

We found that the resolution for binding features is much coarser than for discriminating them in peripheral vision but not in the fovea where they are equal. We also found that abnormal visual development (amblyopia) produces a similar impairment of binding in foveal vision. To understand the potential neural basis for these effects, we used established knowledge about the physiology of V1 to construct a plausible model of this cortical area that could solve our tasks. Our model captured all aspects of the data, suggesting that threshold elevation for the conjunction task derives from imperfect spatial registration between orientation and color maps in cortex. Our results in normal observers show that foveal circuitry has the ability to compensate for this mis-registration. However, our results in amblyopic observers demonstrate that the acquisition of this ability is disrupted by abnormal visual deprivation, providing evidence that its establishment is regulated and consolidated during postnatal development.

METHODS

Observers

Four normal and five amblyopic [4 strabismic and 1 anisometropic (MR)] adult observers participated in this study. Of the amblyopic observers, only JS (strabismic) had residual stereopsis (70 arcs). Viewing was binocular for normal observers. All observers participated in all experiments, with the exception of the oblique orientation discrimination experiment for which we could only test three (AL, NL, and PN) of the four normal subjects. All observers except PN were naïve to the purpose and methodology of the study.

Stimuli

As seen in Fig. 1, segments were 1–2.5 arcmin thick and could vary in length across a range of 2–70 arcmin (the range had to be constrained differently in some experiments, to ensure that large peripheral patterns would not reach the fovea). They could be either bright (74 cd/m2) or dark (0 cd/m2) on a gray (37 cd/m2) background (i.e., they had the same contrast but opposite polarity). Orientation could either be vertical or horizontal with the exception of the oblique orientation-discrimination task for which it could either be horizontal or 30° clockwise away from horizontal.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
The average distance between segment centers was 1.5× segment size for all set sizes except in the spaced-out condition (gray symbols) for which it was 3×. This distance defined the size of an imaginary square within which segment position was jittered uniformly (without allowing any part of segment to be outside the square) to ensure that our stimuli did not consist of a rigid lattice structure. Viewing distance was 57 and 171 cm for peripheral and foveal conditions, respectively, in normal observers and 114 cm for amblyopic observers. We chose these values for the following reasons. For peripheral presentations, we needed subjects to be close to the monitor so that our segments could be big enough to be visible even when many of them had to be presented (as in the 64-set size condition). However, this distance was not appropriate for the fovea because the pixel resolution of our monitor did not allow us to make stimuli small enough for this condition. We therefore had to increase viewing distance by threefold for the foveal condition, allowing us to resolve the psychometric function with even better resolution than for the periphery. This change of viewing distance was necessary to ensure that the lack of difference between orientation and conjunction thresholds in the fovea was not simply due to our inability to resolve threshold differences in the fovea. For amblyopic subjects, neither distance would have worked for the reasons just described with the additional complication of large acuity variations in the amblyopic eye. We settled for a compromise between the two distances used in normal observers. Data for amblyopic subject JT with known target configuration (Fig. 6D, inset) was collected at 57 cm distance, and this manipulation had no effect on our results.

**Trial timeline**

Each trial consisted of two presentations. Each presentation lasted 150 ms with the exception of amblyopic observers JT and MR, who found this duration too difficult and were therefore tested with a longer duration (300 ms). Provided stimulus duration is short enough to preclude eye movements (which would allow subjects to foveate to peripheral targets), we believe it plays little role in the main effects we demonstrate [in fact, peripheral processing may be faster than foveal processing (Carrasco et al. 2003)]. The two presentations were separated by a 1-s gap. On each presentation, an array of 16, 36, or 64 segments was flashed either at fixation (foveal condition) or at 7° horizontal eccentricity (peripheral condition). For each eccentric presentation, the array could appear randomly either to the left or to the right of fixation to ensure that subjects would not be tempted to foveate to either side before stimulus appearance (such strategy would not be advantageous for this design). Only one of the two presentations displayed a target array (see following text), and subjects were asked to select which interval (1st or 2nd presentation) contained the target (2 temporal alternative forced choice). We also tested a subset of the observers (PN, AL, and NL) with both target and nontarget arrays presented simultaneously, one (randomly chosen) to the right and the other one to the left of fixation (2 spatial alternative forced choice, see RESULTS).

**Tasks and instructions to subjects**

For orientation discrimination, the nontarget array contained segments of the same orientation, whereas in the target array, the orientation of one (randomly selected) segment differed from all the others (Fig. 1A). In both intervals, exactly half the segments were black and half were white, randomly assigned. We used the same design for color discrimination, swapping color for orientation (B). For the conjunction task, there were exactly half black, half white, half vertical, and half horizontal segments in each interval (C). In other words, the two intervals were identical in overall color and orientation content. However, color and orientation were paired differently for individual segments in the two intervals. In the nontarget interval, all segments of one orientation also had the same color. In the target interval, we swapped the orientation of two segments, making them “odd.” This task can only be performed by truly conjoining the two features. Before testing, observers were shown example stimuli that remained on the screen for ease of inspection. We explained to them exactly how target and nontarget segment arrays were constructed and how they differed from each other. For the conjunction task, subjective reports indicated that most observers segregated the array by color and then performed orientation discrimination on the segregated subset. However, there is evidence that this subjective impression does not correspond to what subjects actually do, which is to simultaneously segregate both features (Friedman-Hill and Wolfe 1995). The exact strategy used by subjects to perform our task is not particularly relevant to our results. Our main concern was that the conjunction task (however it was performed) required the use of information from both features. This goal is indeed achieved by our balanced stimuli.

**Threshold estimation**

We measured percentage correct as a function of segment length, thus obtaining a psychometric curve. We estimated threshold size by fitting a Weibull function to the raw data \(W(size) = 1 - 5 \times \exp\left(\frac{-(size/threshold)^\alpha}{\beta}\right)\), combined across blocks. This is equivalent to taking the 82% correct point on the psychometric curve. We also repeated our threshold measurements using a Gaussian fit (and taking the 75% correct point on the curve) and confirmed all the effects we report here. However, the raw psychometric curves converged more closely to the Weibull fit than to the Gaussian fit, so we present results from the former analysis here. We also examined whether we could resolve differences in the value of \(\beta\) (steepness of the curve) across tasks and conditions. Although we found that average \(\beta\) values for the orientation discrimination task had a slight tendency to be higher than those for the conjunction task [orientation task: \(\beta = 3.9 \pm 1.5\) (linear axis), \(6.1 \pm 1.9\) (log axis); conjunction task: \(\beta = 2.5 \pm 1.9\) (linear axis), \(5.6 \pm 4.1\) (log axis)], this difference was not consistent for individual observers and could not be adequately resolved by the accu-
racy of our fits. We ensured that thresholds were stable before collecting data that was used for threshold estimation. Most of our tasks, because of their simplicity, required only limited pretraining. We observed some learning at the very beginning of individual testing, but performance saturated very quickly.

**Modeling**

We challenged our model with stimuli that were identical to those used in the psychophysics down to their pixel definition. Our model consists of two layers. Output from the orientation-(o)-selective layer is obtained by

\[
\text{output} = E_{\text{horizontal}} - E_{\text{vertical}}
\]

where orientation energy \(E\) for stimulus image \(I\) is obtained by two-dimensional (2-D) pixel convolution

\[
E_{\text{horizontal}} = (I \otimes RF_{\text{horizontal}})^2 + (I \otimes RF_{\text{vertical}})^2
\]

and similarly for \(E_{\text{vertical}}\) (illustrated in Fig. 2C). RF was a Gabor function obtained by multiplying a 2-D sinusoid (sin for odd, cos for even) of frequency 0.2 cycle/pixel [converted to 10.23 (fovea) and 4.97 (periphery, 7°) cycles/°], consistent with estimates of peak spatial frequency for detecting lines (Levi and Waugh 1994)] with a 2-D Gaussian of SD 2 pixels [converted to 2.34 (fovea) or 4.83 arcmin (periphery), consistent with estimates of RF size in monkey V1 (Dow et al. 1981)]. For fovea, we used an equivalent eccentricity value of 0.32° estimated from the average region occupied by our foveal stimuli at orientation discrimination threshold size, for all set sizes. \(E\) as defined in the preceding text is the currently accepted implementation of a complex-cell output (Pollen and Ronner 1983). Output from the color(c)-selective layer was

\[
\text{output} = I \otimes RF_{\text{color}}
\]

where RF\(_{\text{color}}\) was a 2-D Gaussian of same SD as detailed in the preceding text (illustrated in Fig. 2B). For peripheral simulations, both output\(_{\text{o}}\) and output\(_{\text{c}}\) were sub-sampled every other pixel to simulate the lower RF coverage of V1 at 7° eccentricity, which we estimated from 10 relevant publications (Beard et al. 1997; Cowey and Rolls 1974; Dougherty et al. 2003; Duncan and Boynton 2003; Engel et al. 1994; Fox et al. 1987; Grusser 1995; Sereno et al. 1995; Slotnick et al. 2001; Stevens 2002). For orientation detection, the interval with largest [max(output\(_{\text{o}}\)) − min(output\(_{\text{o}}\))] was selected as the target. Similarly for color detection. For the conjunction task, values in the output matrix were locally scrambled within square regions of 5 × 5 pixel\(^2\) [equivalent to a cortical area of 2 × 2 mm\(^2\)].

**FIG. 2.** How does the model identify targets? A: example target stimulus for the conjunction task. This is convolved with a color filter (B) and with an orientation extractor (C), constructed by taking the difference between horizontal and vertical orientation energy (energy is computed by summing square outputs of Gabor filters in quadrature pairs). D: corresponding output, from the color layer, \(E\) is output\(_{\text{c}}\), from the orientation layer. These are multiplied (pointwise) to obtain the conjunction output\(_{\text{con}}\) in \(F\), but the multiplication is noisy because of misregistration between \(D\) and \(E\) (see METHODS). The misregistration is simulated at the cortical level to match the extent of intrinsic cortico-cortical connections, as shown in \(H\) (black line shows probability distribution of intrinsic connections as estimated from Fig. 10 in Stettler et al. 2002), gray line shows the distribution obtained in this study by scrambling the registration between color and orientation maps). The final scalar response to stimulus \(A\) is obtained as shown in \(G\) by taking the difference between the maximum and the minimum values in output\(_{\text{con}}\). \(I\) shows responses to the other possible target configuration (target segments are black and horizontal and white and vertical, rather than black and vertical and white and horizontal as in \(A\)). \(J\) and \(K\) to the two configurations for nontarget stimuli. The max-min rule \((G)\) returns greater responses for \(A\) and \(J\) than for \(B\) and \(K\), leading to correct identification of target stimuli. Model performance in the other 2 tasks (orientation discrimination and color discrimination) is explained by similar logic.
RESULTS

Foveal versus peripheral thresholds in normal observers

Subjects viewed stimuli like those in Fig. 1 and were asked to select the segment array that contained an odd element (in Fig. 1, the odd element is always in the array to the right of the central fixation cross). In the orientation-discrimination task (A), the odd element differs only in its orientation. Similarly, in the color-discrimination task (B), it only differs in its color. In the conjunction task (C), the target array contains two odd elements that differ in both color and orientation but not in either alone. The target is invisible to a device that monitors either orientation alone or color alone because each segment array contains exactly half black, half white, half vertical, and half horizontal segments (the reader can verify this by counting segments of different colors and orientations in Fig. 1C). We then reduced the overall size of the stimulus (by shrinking it uniformly) until observers were unable to perform the task and determined a threshold segment size for all three tasks. We also studied the effect of overall number of segments (set size) in three quasi-logarithmic steps (16, 36, and 64).

In Fig. 3, size threshold for the conjunction task is plotted on the y axis against threshold for discriminating orientation on the x axis for all set sizes (symbol size scales with set size). We plot the orientation discrimination threshold because this was invariably larger than the color discrimination threshold, and we expect the conjunction threshold to be limited by the largest single-attribute threshold (i.e., the least visible feature should limit the visibility of combined features). Our prediction that conjunction and orientation discrimination thresholds should be equal is confirmed in all four normal observers when stimuli are presented at the fovea (open symbols fall on the solid unity line; paired t-test for difference between conjunction and orientation thresholds across subjects and set sizes is not significant, $P = 0.65$).

We expected that, due to the lower spatial resolution in the periphery, eccentric presentation of our stimuli would lead to an elevation in orientation discrimination thresholds, and our prediction was confirmed (solid symbols are displaced to the right of open symbols; see METHODS for details on eccentric presentation). We also expected that conjunction thresholds would be equally elevated. Surprisingly, what we found is that the elevation in conjunction thresholds is much higher than expected from the elevation in orientation discrimination thresholds (black solid data points fall above unity line; paired t-test as for fovea gives $P < 10^{-4}$; 3-way ANOVA gives $P < 10^{-5}$ for interaction between retinal location (fovea vs. periphery) and task type (discrimination vs. conjunction), whereas no other interaction (retinal location and set size, set size, and task type) is significant ($P > 0.3$), but there is an effect of each individual factor: retinal location $P < 10^{-5}$, task type $P < 10^{-5}$, set size $P < 0.02$). In other words, there is a qualitative difference between fovea and periphery that is not simply accounted for by the difference in spatial resolution between these two regions of visual space. The reader can verify this by comparing the difficulty of each task when viewing the stimuli of Fig. 1 directly, versus viewing them peripherally while fixating the cross (the efficacy of this simple demonstration may vary depending on individual differences).

We also tested a subset of the observers (PN, AL, and NL) with both target and nontarget arrays presented simultaneously, one (randomly chosen) to the right and the other one to the left of fixation (2 spatial alternative forced choice, AFC). The data are shown in Fig. 3E. For subject PN (circles), we obtained very similar results to those obtained using the two temporal AFC method (D). For the other two subjects [upright triangles (HW) and inverted triangles (NL)], performance in the conjunction task was so poor that thresholds could not be measured reliably (labeled as N/A) even though they could always be measured for the orientation task, with the exception of the smallest set-size in subject HW (upright triangle) for whom we could measure both thresholds. In summary, the conjunction deficit is more pronounced when measured using the spatial rather than the temporal version of our 2-AFC design. This is not particularly surprising because each presentation in the spatial 2-AFC design required subjects to process twice as much information in the same amount of time. However, this result cannot be directly compared with the foveal data because we used a temporal design for the fovea and the spatial version of the 2-AFC paradigm is not applicable to the fovea. In the rest of the paper, we therefore focus on data obtained using the temporal 2-AFC version of our peripheral task.

Modeling based on V1 circuitry

To understand what neural mechanisms may underlie the differences between fovea and periphery, we constructed a model of V1 with the following constraints: physiologically plausible and able to make quantitative predictions for our
measurements. Specifically, our model involves an orientation-selective layer of neurons (black) stacked on a color-selective layer of neurons (gray), Fig. 4A. Both layers are topographically organized, providing a cortical spatial map of the feature they are selective for (B and C). The orientation discrimination task is performed by the model based on the output of the orientation-selective layer only and similarly for color discrimination. The conjunction task is performed by combining the two feature maps to generate a joint color-orientation spatial map (see Fig. 2F). However, the spatial accuracy with which the two maps are combined depends on the precision of the local connectivity between the two neuronal layers (solid and dashed connecting lines in Fig. 4A).

In our model, this precision is limited by the inter-layer wiring scatter that we introduced to simulate cortical scatter and cortico-cortical connectivity. An equal degree of wiring scatter across the cortical sheet has substantially different effects in fovea and periphery because of the way RF size and magnification vary across cortex between these two regions of visual space (Dow et al. 1981; Van Essen et al. 1984). Figure 4 shows that if a neuron in the orientation-selective layer is wrongly connected to a noncorresponding neuron in the color-selective layer (solid line in A), the corresponding loss in RF overlap at the fovea (B) is minimal: both neurons still report on the same segment in the stimulus (D), thus leading to correct binding. In the periphery, however, the same cortical misassignment results in RFs being completely nonoverlapping (C): the two neurons now report on different segments (E), thus incorrectly binding the color of one with the orientation of another. This leads to threshold elevation for the conjunction task in the periphery, as shown in Fig. 5A. Model predictions (diamonds) match the experimental data (circles, subject average). Figure 5B shows ratios between conjunction and orientation-discrimination thresholds (y axis) at different set sizes (x axis) for different conditions. Again, our model [shown by light (fovea) and dark (periphery) gray shading] does an excellent job of explaining the data. It also simulates the very small threshold sizes for color discrimination (data not shown).

An interesting aspect of the simulation is that not only does our model correctly predict the threshold elevation for conjunction in the periphery, but it also captures the set size effect (increasing trend with increasing set size) for both individual thresholds (Fig. 5A) and for their ratios (Fig. 5B). This may not be intuitively obvious in that we did not incorporate any special mechanisms (e.g., a serial searcher for conjunction) that would be trivially expected to produce this pattern. The effect derives from the fact that additional segments increase the opportunity for generating false targets (illusory conjunctions).

The model prediction for foveal ratio between conjunction and orientation at set size 16 is slightly smaller than one (left portion of light gray shading in Fig. 5B), meaning that our model predicts that size threshold in the conjunction task should be smaller than size threshold in the orientation discrimination task for this condition. This effect stems from the fact that the stimuli we used for the conjunction task contained two target segments, whereas there was only one target segment in the orientation discrimination task [although the 2 tasks were equivalent in terms of the number of configurations (i.e., 2) that targets could randomly take: (vertical and black) + (horizontal and white) or (vertical and white) + (horizontal and black) in the conjunction task, and vertical or horizontal in the orientation discrimination task]. Indeed, we observed this effect in some human observers (smallest open symbol in Fig. 3, A and D) and in the subject average (left-most open symbol in Fig. 5B), but it was rather small.

Our model allows us to make a number of additional predictions. For example, it predicts that if segments are spread further apart from each other, there will be less opportunity for wiring scatter to mislabel color and orientation at each segment location. This should result in a reduction of the conjunction deficit in the periphery, as shown in gray diamonds in Fig. 5A and in gray shading in B. We were able to confirm this prediction experimentally within error (gray circles in Fig. 5 show the mean data; data for individual subjects shown in Fig. 3).

Another important prediction is that thresholds for the orientation discrimination task should be elevated if the discrimination is made between segments that differ only by 30° rather
than 90° in orientation. Crucially, our model predicts that a similar elevation should be observed both at fovea and periphery [thin gray (fova) and black (periphery) lines in Fig. 5B] because there is no conjunction involved in this comparison (only the orientation map is used) and therefore no role for registration noise between color and orientation maps. Again we confirmed this prediction experimentally (small circles in Fig. 5B, open for fovea and solid for periphery).

**Observers with abnormal early visual development (amblyopia)**

We were interested in exploring the developmental basis of the differences we measured between fovea and periphery. In other words, we wished to establish whether these differences are already present at birth or whether they are developmentally regulated. Because RF overlap and cortical scatter have been previously implicated in amblyopic vision (Kiorpes and McKee 1999; Levi and Klein 1985), our success in using these factors to explain normal peripheral vision led us to suspect that the amblyopic fovea might also show threshold elevation for our conjunction task. We therefore performed the same experiments on observers with amblyopia. Figure 6, A–E, uses the same format as in Fig. 3 except open symbols now refer to the fovea of the dominant eye and solid symbols to the fovea of the amblyopic eye. The dominant eye, like normal eyes, shows no selective impairment of conjunction thresholds (unity line, paired t-test P > 0.25) with the exception of observer JS. In contrast, the amblyopic eye, in line with the proposal outlined at the beginning of this paragraph, shows a marked deficit that parallels the deficit we observed in the periphery of normal subjects [compare solid symbols in Figs. 3 and 6; 3-way (eye, task, set-size) ANOVA between deprived eyes of amblyopic observers and eyes of normal observers gives P < 0.01 for interaction eye task with no other significant interac-

---

**FIG. 5.** Subject averages match model predictions. A: conjunction vs. orientation thresholds for all subjects (circles) and for model simulations (diamonds) at fovea (open) and periphery (solid). Gray is for the condition already described in Fig. 3. Symbol size refers to set size (small circles, 16; medium circles, 36; large circles, 64). Solid line shows unity. B: conjunction/orientation threshold ratios against set size. Circles are experimental data (subject average ±1 SE). Gray shading and thin lines are model predictions (±1 SD): light gray for fovea, mid-gray for the gray condition in Fig. 3 (spaced-out segments), dark gray for periphery. Thin lines and small circles refer to the oblique orientation discrimination task (for these points, the ordinate plots oblique/orthogonal threshold ratios for orientation discrimination, black line refers to periphery and gray line to fovea (the gray lines are hardly visible because they superimpose with the black lines)). Dotted line shows ratio = 1.

**FIG. 6.** The amblyopic fovea (solid) behaves like a normal periphery. A–E: same plotting conventions as in Fig. 3 except here solid refers to fovea of the amblyopic eye, open to the fellow dominant eye. F: conjunction/orientation threshold ratios for amblyopic eye and set size 64 segments (large solid symbols in A–E) are plotted against amblyopic/dominant threshold ratios for orientation discrimination (same set size) in all 5 subjects. D, inset: data obtained for subject JT when the exact configuration of the target was known (see METHODS).
tions ($P > 0.05$), and each factor significant at $P < 0.02$; ANOVA between amblyopic eyes and dominant eyes in amblyopic observers gives $P = 0.02$ for interaction eye task with no other significant interactions ($P > 0.05$), and each factor significant at $P < 0.03$). Interestingly, JS was our only amblyopic observer with residual stereovision, and previous studies have reported that amblyopes with residual stereo-competence often show transfer of the deficit in their amblyopic eye to the dominant eye (Ho et al. 2005; McKee et al. 2003). Our results are consistent with these previous reports in that JS is the only subject that shows conjunction impairment in both eyes.

Our conjunction task is unconventional, in that “target” segments may take any combination of color and orientation (see Fig. 1), whereas in most studies the exact target configuration is known. We therefore tested $PN$, $AL$, and $JT$ (amblyopic) with a version of our tasks in which the target was always of a specific orientation and color, and subjects were informed of this. In other words, they were looking for, say, a black vertical target segment in all stimuli and tasks. The resulting thresholds are shown by the squares in Fig. 3 for the periphery of subjects $PN$ and $AL$ (at set size 36) and by the inset in Fig. 6D for the amblyopic subject $JT$. They are very similar to those obtained when the exact configuration is not known.

We worried that the effect we observed in the amblyopic eye could simply be a consequence of its lower spatial resolution (the definition of amblyopia). Lower spatial resolution leads to higher orientation discrimination thresholds, meaning that to resolve single attributes, our stimuli needed to be larger in the amblyopic eye than in the dominant eye. Although always centered at the fovea, it is possible that they extended to the periphery in the amblyopic eye (because they were larger at threshold) but not in the dominant eye. If this was the case, then the effect we observed in our amblyopic observers may simply be a replication of our earlier result in the normal periphery and only related to amblyopia in interesting ways (i.e., as a consequence of reduced spatial resolution). To test for this possibility, we reasoned that if the elevation in conjunction thresholds is related to the spatial resolution deficit in the amblyopic eye, then these two quantities should be positively correlated across our sample, i.e., amblyopes with larger deficits in spatial resolution between dominant and amblyopic eye should also show larger elevations in conjunction thresholds. Instead, we found a negative correlation ($R = -0.87$) between the ratio conjunction threshold to orientation threshold on the one hand (plotted on the y axis in Fig. 6F), and the ratio orientation threshold in amblyopic eye to orientation threshold in dominant eye on the other hand (x axis), for the largest set size we used (64, the set size for which we observed the largest effect in conjunction deficit). In other words we found some evidence for the opposite trend: observers with good resolution in their amblyopic eye showed large elevations in conjunction thresholds (e.g., $AP$), whereas those with poor resolution showed relatively smaller elevations (e.g., $MR$).

We repeated our measurements with segments further apart (gray symbols) for set size 64 in our amblyopic sample. With the exception of one observer ($AP$), for whom this manipulation completely eliminated the elevation in conjunction thresholds (gray symbol in Fig. 6A), all remaining observers were unable to perform the task when segments were spaced further away from each other (no data are plotted because we were unable to measure conjunction thresholds), contrary to a straight-forward prediction based on the results obtained in the periphery of normal observers.

We used our model to fit the amblyopic data, using a peripheral simulation of 3° “equivalent eccentricity” (see Discussion for a detailed explanation of this concept). As shown in Fig. 7 our model [light (dominant) and dark (amblyopic) gray shading] can account for the main results within experimental error, indicating that the amblyopic fovea behaves like the parafovea of normal observers in our tasks. The model does not predict the extremely poor performance of most amblyopes when 64 segments are spaced further apart. However, the twofold increase in overall stimulus size introduced by this manipulation causes many segments to impinge on the periphery of the amblyopic eye where performance in our tasks is unmeasurably poor. We do not know how to specify the equivalent eccentricity in our model under these conditions.

**DISCUSSION**

**Relations to previous literature**

The foveal superiority for feature binding we demonstrate in this study may be related to an interesting illusion by Wu et al. (2004), where dot-by-dot pairing of color and motion in the periphery is misperceived in favor of a different pairing presented at the fovea. In the absence of foveal stimulation, peripheral perception reverts to veridical. Wu et al. hypothesize that this effect may result from higher reliability of foveal signals compared with peripheral ones, biasing the visual system to rely on the former when available. Indeed, we have demonstrated that for small visual items (like those used by Wu et al.), the periphery is unreliable in local binding of visual features, thus providing support for their proposed explanation of the illusion they report. We considered making the link stronger by repeating our measurements using motion instead of orientation. However, we realized that this experiment would be ill-designed because changes in stimulus size would co-vary with changes in stimulus velocity, thus generating ambiguous results (alternatively, if the duration of the stimulus

![FIG. 7. Amblyopic subject averages match model predictions. Same plotting conventions as in Fig. 5B. Main panel: averages of 4 subjects with no stereo-vision; inset: data for the remaining subject (JS), who had residual stereo-vision. Model predictions for amblyopic eye (dark gray shading) correspond to a peripheral simulation of 3° equivalent eccentricity. JS is plotted separately here to emphasize the potential effect of residual stereo-vision [i.e., interocular transfer of the conjunction deficit (Ho et al. 2005)], but this subject was included in all statistical tests reported in the paper.](image)
Our results are also likely related to the phenomenon of illusory conjunctions (Treisman and Schmidt 1982): in very brief displays of, say, a green letter O and a red X, normal observers sometimes perceive the O to be red and the X to be green. Although some of our amblyopic observers (JT in particular) did report that it sometimes looked as though there were more than the physical number of odd elements (2 in the case of our stimulus design for the conjunction task), it is very hard to gauge one’s perceptual experience at threshold. Therefore we cannot say with certainty whether the threshold elevation we observe for the conjunction task is due to perceiving too many targets or to perceiving none (both conditions can generate an erroneous response).

Carrasco et al. (1995) have described an “eccentricity effect” whereby error percentage and reaction time increase for conjunction searches as stimuli are presented at larger eccentricities. In subsequent work (Carrasco and Frieder 1997), they showed that this effect is accounted for by cortical magnification. By using target size as our threshold measure, we can equate the accuracy of detecting conjunctions in fovea and periphery (consistent with Carrasco and Frieder 1997)—however, we show here that the scaling required is much greater for conjunction than for detection. More importantly, in our experiments, we factor out cortical magnification by taking the ratio between conjunction and discrimination thresholds (Figs. 5B and 7), so the effects we report cannot be simply accounted for by changes in cortical magnification.

Hess and Field reported increased spatial uncertainty in peripheral (Hess and Field 1993) and ambiopic (Hess and Field 1994) vision, which they interpreted as deriving from disarray noise in the spatial map of the visual stimulus. This model fails to account for our findings in that it predicts no difference between discrimination and conjunction thresholds because it only perturbs the spatial location of objects within a single map. To explain our results, it is necessary that color and orientation are unbound in different maps (Ashby et al. 1996; Treisman 1998) before their spatial binding is perturbed by mis-registration of the two maps. This is the model we used in our simulations (see next section).

More generally, it is worth emphasizing that spatial localization alone does not in any way explain our results. The issue here is whether features are bound or unbound in the early representation of stimulus elements. If they are bound, any amount of localization impairment will not affect the conjunction task more than the orientation task. In other words, if segments were coded as having a specific orientation and contrast polarity in their early representation, then the system should be able to perform both single- and double-feature tasks equally well, no matter how poor the ability to localize target elements. Our results imply that any perturbation must be applied to unbound feature representations before binding. Indeed, at peripheral threshold size for the conjunction task segments were well above both orientation and color threshold sizes: segments were about three times larger than required for their orientation to be visible, and about seven times larger than required for their color to be visible. Clearly, the limiting factor was not the ability to process individual attributes but rather the ability to combine them. Another way of thinking about this issue is in terms of task demands. The orientation and the color tasks do not require subjects to localize target segments within the array. The conjunction task implicitly requires spatial localization before the two features are bound but does not require it after they are bound. After binding, task demands are the same for all three tasks with respect to spatial localization, so spatial localization alone cannot explain the differences we observed. It is only before binding that an argument can be made for a role of spatial localization in determining the differences we observed. In our model, noisy registration between feature maps and poor spatial localization within each feature map are virtually indistinguishable, so our model may be formulated in terms of poor spatial localization in the periphery but only provided such poor spatial localization is applied to each feature map before binding. Notice, however, that for this explanation to capture our results, the spatial localization would have to be several times poorer than the corresponding resolution for that individual feature.

At first sight, our results may appear to contradict previous studies on feature conjunction. More specifically, our results in the fovea may appear to contradict the common belief that conjunction tasks are harder than simple feature search even in the fovea (Wolfe and Cave 1999), and our results in the periphery may appear to contradict previous findings that performance in a conjunction task was entirely predicted by performance in the constituent feature discrimination tasks (Eckstein et al. 2000). This apparent discrepancy disappears when one considers that in all previous studies spatial factors were unlikely to be limiting performance. These studies assessed performance using metrics (like reaction times or percent correct) that did not expose the limitations in spatial resolution exposed by our study. For this reason, our results cannot be directly compared with those obtained in previous studies, which focused on other factors than spatial resolution.

Crowding is a well-known phenomenon in which arrays of nearby objects can render elements that would, in isolation, be perfectly visible, unidentifiable. It is clear that crowding was operative in our experiments. We observed clear crowding for orientation discrimination both in fovea and periphery. In foveal presentations, discriminating the orientation of target segments embedded in distractor arrays was only possible up to a resolution (between 5 and 10 arcmin) ~1.8 times worse than when target segments were presented in isolation (averaged across observers), consistent with previous estimates of foveal crowding (Toet and Levi 1992; Wolford and Chambers 1984). We observed a similar degree of crowding in the periphery (~2.8-fold increase in size threshold for the embedded condition). We can therefore confirm that crowding was active in our orientation discrimination task for all the main conditions we tested. Clearly, we also observed crowding for the conjunction task. There were similar differences between embedded and isolated targets for orientation discrimination in both eyes of amblyopes (1.99 and 2.15 in dominant and amblyopic eyes, respectively).

There is no accepted knowledge that would allow us to construct a crowding model capable of providing quantitative...
predictions for all our experiments. The specifics of crowding are still debated, and the term itself is likely to include very different mechanisms that appear to behave similarly in some experimental conditions. Moreover, current descriptions of crowding fail to explain various aspects of our data—for example, it is not clear that crowding would account for the set size effects we observe (in multi-segment arrays the crowding power of each segment is reduced by mutual crowding from nearby segments, so the addition of more segments causes crowding as well as disinhibition of crowding), and oblique orientation discrimination is treated like a conjunction in current descriptions of crowding (Pelli et al. 2004). Finally, if anything we would expect more crowding in the stimuli we used for the discrimination tasks (where segments were all the same with respect to ≥1 dimension) than those for the conjunction task because crowding is most effective between more similar elements (Kooi et al. 1994). For all these reasons, in the next section we interpret the data using our V1 model, every aspect of which is clearly specified. This model shares some of its features with established theoretical frameworks, such as feature integration theory (Treisman and Gelade 1980) and its features with established theoretical frameworks, such as feature integration theory (Treisman and Gelade 1980) and guided search (Wolfe 1996). We suspect that this model may indeed account for a number of important aspects of crowding (Levi et al. 1985).

Relations to cortical physiology

Because orientation selectivity is first encountered in V1 along the visual pathway (Hubel 1982), this is the earliest stage that can support orientation discrimination in our experiments. Moreover, in performing this task observers could resolve segments of ~0.1° in the fovea and ~0.25° at 7° of eccentricity, figures that are compatible with RF size in monkey (Van Essen et al. 1984) and human (Kastner et al. 2001) V1. For these reasons, it seems safe to conclude that orientation discrimination relied on signals carried by neurons in either V1 or at some later stage in the visual hierarchy. Similarly, our conjunction task could not rely on signals at earlier stages. There is extensive evidence for the involvement of extrastriate cortex in feature binding (Robertson 2003). However, these downstream mechanisms rely on signals from V1. Therefore we consider the possibility that our tasks (including conjunction) were limited by signals in V1. We chose this conservative position because, by restricting ourselves to V1, we can make detailed quantitative predictions that would not be possible with a model based on areas whose physiology has not been characterized as thoroughly as V1.

There are two critical features to our model. Most other aspects of the model (e.g., how orientation energy is computed, or the size of RFs) are either widely accepted or only marginally relevant to the success of our simulations. The two critical features are that RF overlap is larger in the fovea than it is in the periphery and that V1 cortex contains separate, topographically registered representations of features (e.g., maps) but registration is noisy. If the first feature is removed from the model, we cannot simulate any difference between fovea and periphery. If the second feature is removed, we cannot show any elevation for conjunction thresholds compared with orientation discrimination thresholds. Because of their critical nature, we need to assess the plausibility of these two aspects of our model.

RF overlap is proportional to RF size × magnification. The reduced RF overlap for the periphery (compare B and C in Fig. 4) is a well-established fact of V1 physiology (Dow et al. 1981; Van Essen et al. 1984) (compare F and H in the same figure). We used an RF-overlap ratio of 2 between fovea and periphery in our simulations (roughly as shown in Fig. 4). This was rounded from a median of 2.4 that we estimated from 10 published papers attempting to measure cortical magnification in human V1 using fMRI (Dougherty et al. 2003; Duncan and Boynton 2003; Engel et al. 1994; Sereno et al. 1995), PET (Fox et al. 1987), VEP (Slotnick et al. 2001), psychophysics (Beard et al. 1997), migraines (Grusser 1995), cortically induced phosphenes (Cowey and Rolls 1974), and evolutionary scaling laws (Stevens 2002). For RF size, we used monkey estimates from Dow et al. (1981) (the only study providing information relevant to the fovea). This aspect of our model is therefore not only plausible but necessary for a veridical implementation of known V1 architecture. Although an early paper by Hubel and Wiesel (1974) claimed that RF overlap did not vary with eccentricity, these authors did not record from neurons with representations that were close enough to the fovea to observe the increased cortical point image size subsequently demonstrated by Dow et al. (1981) and Van Essen et al. (1984). The now widely accepted knowledge that RF overlap does vary with eccentricity is captured by Stevens’ scaling laws (Stevens 2002).

The most compelling evidence for the existence of feature representations in V1 comes from the optical imaging literature: eye-, orientation-, color- (Xiao et al. 2003), direction-, and disparity-selective maps have all been measured in V1/V2 of either cats or monkeys (Swindale 2001). These maps are co-registered at a macroscopic scale (Weliky et al. 1996; Yu et al. 2005), and it is believed that co-registration retains retinotopic structure at least coarsely (Swindale et al. 2000). However, it has not been possible so far to determine whether and how these maps are co-registered at the level of neuronal connectivity. We hypothesize that the extent of mis-registration is in units of cortical distance, is uniform across cortex (whether fovea or periphery), and matches the extent of horizontal cortico-cortical connections in monkey V1 (we used an inter-map wiring scatter of 0–2 mm, consistent with Angelucci et al. 2002 and Stettler et al. 2002) (see Fig. 2H). Given our present knowledge of V1 physiology, intrinsic connections are the best candidates for mediating inter-map connectivity. We realize that this link is tenuous, in that it is not evident that intrinsic connections should serve the purpose of registering feature maps, but this is the most reasonable guess that our current knowledge of cortex allows. There is no existing evidence on how eccentricity relates to the cortical extent of intrinsic connections (C.D. Gilbert, personal communication), so we opted for the most parsimonious assumption (equal size everywhere).

Given the structural plan just described, all parameters in our model are constrained by known facts of V1 physiology (the extent of such constraint is exemplified in Fig. 2H, which shows how closely we simulated the empirically determined spatial spread of cortico-cortical connections). In this sense, the model has no free parameters. It is also robust in that its behavior depends on very few critical features. We find it compelling that a model fully constrained and characterized by independent knowledge and data that was acquired by anatom-
Impaired binding in periphery and amblyopia

The fovea of deprived eyes in amblyopic subjects showed an elevation in conjunction thresholds that parallels a normal periphery at an equivalent eccentricity of ~3°. In normal eyes, this figure marks the point at which the cortical point image size drops from its high foveal value and levels off to a much lower (~1/3), more-or-less constant value for the periphery (Van Essen et al. 1984). Consistent with previous evidence (Levi and Carkeet 1993; Levi and Klein 1985), our data show that visual deprivation prevents foveal vision from crossing the 3° threshold that marks the sharp transition from parafovea to fovea.

A word of caution is necessary to interpret the 3° figure correctly. This figure was used as an equivalent eccentricity in our model to provide a fit to the average amblyopic data for conjunction/orientation ratios (Fig. 7) not to the absolute threshold values. Although we averaged absolute values for normal observers (Fig. 5), we did not do the same for amblyopic observers because this would make no sense given the wide range of amblyopic deficits represented in our sample (compare the loss in acuity for JS as opposed to MR, shown by the values on the abscissa of Fig. 6F). Averaging in the case of amblyopic observers only makes sense for ratios (as in Fig. 7), where the absolute amblyopic deficit is factored out. Ratios reflect the change in RF overlap between fovea and periphery in our model but do not depend very much on the absolute eccentricity used for the periphery. In other words, once RF overlap drops to parafoveal values, it stays more or less constant across periphery (Van Essen et al. 1984). Because our model was tailored to average amblyopic ratios, and not to absolute threshold values, the 3° figure for equivalent eccentricity indicates that RF overlap in the amlyopic fovea is similar to a normal parafovea, but it does not say anything about how absolute RF size in the amlyopic fovea compares to absolute RF size in the parafovea. The absolute equivalent eccentricity for the amlyopic fovea depends on the extent of the amlyopic deficit in each subject.

Our results demonstrate that the inability to develop normal foveal vision can have unexpected repercussions on spatial vision, repercussions that go well beyond a simple loss in acuity, resolution or localization accuracy as shown by the studies cited in the preceding text. Because accurate co-registration of feature maps requires highly developed foveal circuitry, amlyopic foveas are poor at resolving the conjunction of multiple features even when the feature size has been scaled to compensate for their reduced resolution (Fig. 7). In other words, there may be situations in which amlyopic eyes can resolve individual attributes (like orientation or color) but cannot resolve the way in which they are combined at each spatial location. This result could not be directly inferred from previous studies.

The earliest physiological effects of amlyopia are seen in V1 (Kiorpes and McKee 1999; McKee et al. 2003), and they generally occur before foveal specialization in the retina is complete at ~4 yr of age (Youdelis and Hendrickson 1986). Interestingly, the peripheral retina appears to develop much more rapidly than the fovea. From birth to beyond 4 yr of
age, cone density increases in the central region, due to both migration of receptors and reduction in their dimensions. It seems likely that these alterations in the retina are also reflected in V1. Indeed, the massive migration of retinal cells, and the alterations in the size of retina and eyeball (along with changes in interpupillary distance), may necessitate the plasticity of cortical connections early in life. We speculate that interruption of normal cortical development due to amblyopia results in reduced V1 RF overlap and increased scatter noise in neuronal topographic wiring, similar to the normal periphery.

**Summary**

Our data show that peripheral vision is incapable of resolving combinations of multiple features at spatial scales at which it can resolve them individually. We further show that foveal vision does not suffer from this limitation, provided it undergoes normal visual experience during postnatal development.

Our model of noisy co-registration between cortical maps for different visual attributes offers a solid framework for interpreting these results in terms of known physiology in V1 and generates quantitative predictions for the sizeable limitations that can arise when attempting to integrate multiple features at small spatial scales. Our results demonstrate that multi-attribute maps possess less resolving power than single-feature maps (up to a $3 \times$ difference). These constraints must be taken into account in situations where peripheral or deprived vision play relevant roles (for example in designing devices for visual aids).

On the assumption that our model is at least partly correct, future research will need to explain how foveal development accomplishes the formidable task of fully compensating for imperfect registration between feature maps in cortex. More specifically, it will be necessary to determine whether development reduces noise in inter-map connectivity so that it does not exceed the tolerances of foveal sampling or whether it increases foveal sampling so that it can tolerate connectivity noise or both.

**Acknowledgments**

We thank S. Klein, R. Li, M. J. Morgan, A. Popple, and J. Wolfe for useful comments.

**Grants**

This research was supported by National Eye Institute Grant ROI EY-01728.

**References**


