



# Early binding of feature pairs for visual perception

Alex O. Holcombe and Patrick Cavanagh

Vision Sciences Laboratory, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138, USA

Correspondence should be addressed to A.O.H. ([aholcombe@psy.ucsd.edu](mailto:aholcombe@psy.ucsd.edu))

If features such as color and orientation are processed separately by the brain at early stages<sup>1,2</sup>, how does the brain subsequently match the correct color and orientation? We found that spatially superposed pairings of orientation with either color or luminance could be reported even for extremely high rates of presentation, which suggests that these features are coded in combination explicitly by early stages, thus eliminating the need for any subsequent binding of information. In contrast, reporting the pairing of spatially separated features required rates an order of magnitude slower, suggesting that perceiving these pairs requires binding at a slow, attentional stage.

To determine the temporal resolution of the perception of feature pairs, we combined color and orientation features, either spatially superimposed or spatially separated (Fig. 1). In the superimposed condition, half the trials consisted of a display alternating between a red patch tilted leftward and a green patch tilted rightward (semicircular windowed gratings with a sinusoidal 45° tilt). In the other trials, the pairing of color and orientation was reversed. Observers reported whether the red was paired with rightward or with leftward tilt.

Observers fixated 0.2° away from the straight edge of the semicircle windowed grating 8.5° in diameter. The red and green stimuli were set to equiluminance separately for each observer at each temporal frequency with a minimum motion method<sup>3</sup>. Following this setting, the red and green 0.59 cycles/degree gratings had peaks of equal luminance and troughs of equal luminance, and the luminance of the peak and trough of each cycle differed by about 38 cd/m<sup>2</sup>. The trough of the mean red grating was 16.4 cd/m<sup>2</sup>, CIE ( $x = 0.59$ ,  $y = 0.35$ ), and its peak was 54.2 cd/m<sup>2</sup> ( $x = 0.41$ ,  $y = 0.34$ ). The mean green grating had trough of 17.4 cd/m<sup>2</sup> ( $x = 0.30$ ,

$y = 0.55$ ) and peak of 56 cd/m<sup>2</sup> ( $x = 0.32$ ,  $y = 0.39$ ). These values ensured that the sum of the bright red and dark green gave the same yellow as the sum of the dark red and bright green, thus camouflaging the orientations of the green and red gratings in the sum. Perfect camouflage in the sum was verified with each observer and, for one observer (AH), was tested over a range of contrasts bracketing the critical values to determine the sensitivity to possible unintended display deviations. These control data showed that a deviation of greater than 5% contrast would be required to bring performance to the 75% threshold, whereas the computed value fell within 2% of the optimal value (producing chance performance). The first and last pairing in the sequence was masked; the trial began with extremely rapid presentation of the stimuli, and gradually slowed to the intended temporal frequency (a few hundred milliseconds), after which the presentation rate gradually increased, ending the trial with a postmask.

In the spatially separated condition, the same pairings were used, but the color was presented as a uniform patch of saturated red or green, and the orientation was presented as an adjacent achromatic patch of tilted bars. Another block of the same experiment tested judgments of brightness and orientation pairings. A dark (30 cd/m<sup>2</sup>) semicircle windowed grating alternated with a bright (86 cd/m<sup>2</sup>) grating on a noise background. The difference between the peak and trough of both gratings was always 55 cd/m<sup>2</sup>, which ensured that the sum of the two gratings was the same regardless of the brightness–orientation pairing.

The critical rates for threshold accuracy (75%) in reporting the pairings were slower than 3 Hz for each observer in the spatially separated condition (Fig. 1). In the spatially superimposed condition, thresholds were nearly ten times better. The average for four observers in the color condition was 18.8 Hz, and the average in the brightness condition was a remarkable 35.5 Hz. This latter value corresponded to ~14 ms for each feature pair.

Several studies suggest that features are more likely to be processed together if they form part of a single object or group<sup>4</sup>. The critical rates for the spatially separated conditions may have been slower than rates for the superimposed conditions because the separated features appeared to be part of separate objects or groups. To test this possibility, we devised several variations that grouped the features together, based on displays shown to have effects in earlier reaction time studies<sup>4</sup>. However, even when the separated features were linked into a common object or surface, critical rates remained low (Fig. 2). These results do not rule out the possibility that an object context could provide a threshold advantage in the region of a dozen milliseconds, but any possible advantage (statistically insignificant in this study) was small compared to that afforded by superposition (300 ms).

When the superimposed features were alter-

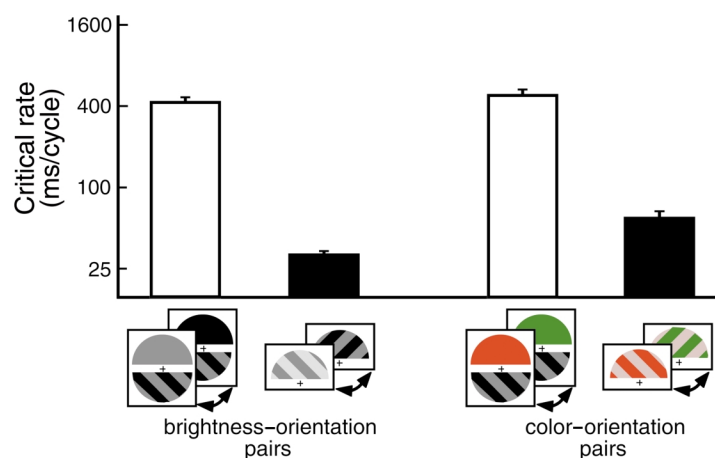
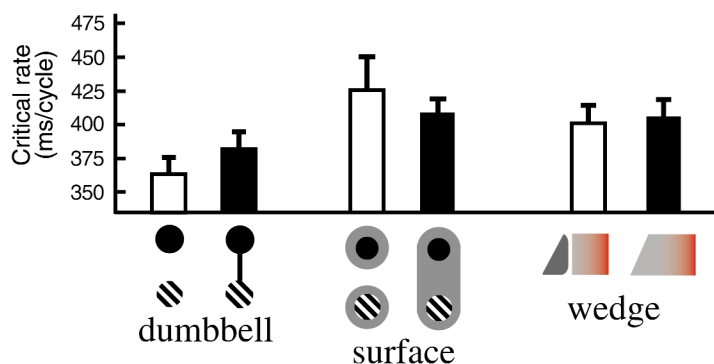


Fig. 1. The critical rates for 75% accuracy in pairing spatially superimposed features (depicted by the second and fourth icons along the horizontal axis) are nearly ten times faster than rates for pairing spatially separated features (first and third icons). Each bar represents the mean of the same four observers. Small vertical bars, 1 s.e.m.



**Fig. 2.** The critical rates for 75% accuracy do not differ significantly when spatially separated features are linked to form a group or object (white bars, unlinked versions as depicted by the left icon of each pair; black bars, linked versions as depicted by the right icon of each pair). Small vertical bars, 1 s.e.m. Averages of data from 25 observers are shown, except middle conditions, which shows average of data from 7 observers. The feature pairings of the dumbbell and lozenge-shaped stimuli alternated between combinations of light or dark on one end and left- or right-tilt on the other. The feature patches of the dumbbell were sharp-edged as depicted, whereas those of the lozenge stimulus were smoothly graded into their surround. The left edge of the wedge stimulus alternated between right- and left-tilted, whereas the right region graded smoothly from the gray to, alternately, either red or green.

nated at rates above approximately 10 Hz, observers reported that they did not experience the brief individual stimulus presentations separately. Instead, some observers reported that the two feature pairings slowly alternated in their awareness, or that one dominated. Others said that both seemed available simultaneously, as if they had been presented transparently. Still, underlying this percept is a high-speed process able to read out the combined values of color (or brightness) and orientation within the brief presentation of a single pairing. If this process were not able to resolve each brief presentation, it would be faced with the patterns summed across two or more intervals. Because of the way the stimuli were constructed, the pairings were camouflaged when the intervals were summed (red right plus green left is indistinguishable from red left plus green right). Although the alternate pairings had to be read out separately in the consecutive intervals, the experience of each pairing extended over time, as reflected in the reports of transparency and slow rivalry. It seems that subsequent processes before awareness integrate the paired representations over a relatively long interval.

The high pairing rates for spatially superimposed feature pairs suggest, among other possibilities, that some features may be assessed in combination from early levels. In this case, there is no binding problem. For example, nonlinear ON- and OFF-channels as early as retinal ganglion cells separate stimuli brighter and darker than the background, even at high flicker rates<sup>5,6</sup>. Once the bright and dark stimuli are represented in

separate channels so that they do not cancel, orientation analysis can proceed independently of the flicker. In the case of our color-orientation pairings, color-opponent cells may have a similar involvement by separating the red and green stimuli, even at high presentation rates<sup>7</sup>. In contrast, in the case of the spatially separated condition, it is unlikely that any one cell is selective for the orientation at one location and the brightness or color at another. Judging the pairings of spatially separate features would therefore require a combination of the responses of disparate neurons selected by later, slower and attentive stages<sup>8,9</sup>.

#### ACKNOWLEDGEMENTS

This work was supported by an NEI NRSA graduate fellowship to A.O.H. and EY09258 to P.C.

RECEIVED 24 AUGUST; ACCEPTED 12 DECEMBER 2000

1. Zeki, S. M. *Nature* 274, 423–428 (1978).
2. Treisman, A. & Gelade, G. *Cognit. Psychol.* 12, 97–136 (1980).
3. Cavanagh, P., MacLeod, D. I. A. & Anstis, S. M. *J. Opt. Soc. Am.* 4, 1428–1438 (1987).
4. Scholl, B. *Cognition* (in press).
5. Kuffler, S. W. *J. Neurophysiol.* 16, 37–68 (1953).
6. Lee, B. B., Martin, P. R. & Valberg, A. *J. Physiol. (Lond.)* 414, 245–263 (1989).
7. Gur, M. & Snodderly, M. *Vision Res.* 37, 377–382 (1997).
8. Verstraten, F. A. J., Cavanagh, P. & Labianca, A. *Vision Res.* 40, 3651–3664 (2000).
9. Lu, Z. & Sperling, G. *Vision Res.* 35, 2697–2722 (1995).