Rapid global form binding with loss of associated colors

Colin W.G. Clifford
Color, Form & Motion Lab, Visual Perception Unit, School of Psychology, The University of Sydney, Australia

Alex O. Holcombe
School of Psychology, Cardiff University, Cardiff, United Kingdom

Joel Pearson
Color, Form & Motion Lab, Visual Perception Unit, School of Psychology, The University of Sydney, Australia

Distributed neural processing creates a binding problem: the representations of the various features of an object are distributed across brain areas but must be associated with, or bound to, the same object. Here we determine the limits that binding imposes on the perception of global form in static flow fields defined by oriented dot pairs. The binding of local orientation signals into global form is shown to operate at rapid 20-Hz rates, implying that visual signals at the point of global form extraction retain precise temporal registration. Binding global form with color is limited to rates of 3-5Hz, showing that binding across attributes can impose a severe temporal limit on perception. Judgment of the temporal sequence of the global structures is also limited to slow rates. These results point to a substantial loss of temporal resolution in the visual system following the extraction of global form but preceding visual awareness.

Keywords: binding problem, global form, color, awareness, psychophysics

Introduction

Spatial form is processed hierarchically in the primate visual system, beginning with the extraction of local stimulus orientation in primary visual cortex (Hubel & Wiesel, 1968). Selectivity for more complex patterns of form is not observed at this stage (Smith, Bair, & Movshon, 2002) but is a common property of cells later in the form processing hierarchy (Gallant, Braun, & Van Essen, 1993; Gallant, Shoup, & Mazer, 2000; Hedégé & Van Essen, 2000, 2003; Tse et al., 2002) that may in turn provide the input to object and face recognition mechanisms (Rentschler, Treutwein, & Landis, 1994; Wilkinson et al., 2000).

Populations of neurons that process spatial form may be separate from those that process color and motion (Zeki, 1978). This separation yields a binding problem when, for example, two differently colored forms are presented. With the form pathway neurons signaling the presence of two forms and other neural populations signaling the presence of two colors, the issue is how the system determines which color belongs with which form. The need to resolve this ambiguity may be the cause of the perceptual asynchronies found when observers attempt to pair color and form (Moutoussis & Zeki, 1997a; Clifford, Arnold, & Pearson, 2003) or color and motion (Moutoussis & Zeki, 1997a; Moutoussis, & Zeki, 1997b; Arnold, Clifford, & Wenderoth, 2001; Arnold & Clifford, 2002; Nishida & Johnston, 2002; Bedell, Chung, Ogmen, & Patel, 2003; Clifford, Spehar, & Pearson, 2004).

In addition to binding different visual attributes, such as color and form, binding multiple instances of the same attribute is also sometimes necessary. For example, to perceive the spiral form of the Glass patterns of Figure 1A, the local orientations defined by dot pairs must be linked together into the spiral pattern. At the level of early visual cortex (V1 and V2), these local orientations are represented by distinct neurons, as the receptive fields of V1 and V2 neurons are too small to take in more than one dot pair. Furthermore, recent single-cell recordings have indicated that although in some instances stimuli outside the receptive field affect responses, in the case of Glass patterns, the global form does not significantly influence the response of V1 or V2 cells to a local dot pair (Smith et al., 2002; Movshon, Smith, & Kohn, 2003).

The need to bind together the spatially distributed orientations at a later stage could potentially impose severe temporal limits on perception. Indeed, pairing color and orientation of gratings when these attributes are spatially separated is limited to slow 3-Hz rates, whereas the pairing of these attributes when in the same location can be perceived at 20 Hz (Holcombe & Cavanagh, 2001). One possible explanation is that fine temporal precision may be sacrificed for extended integration as one ascends the visual hierarchy. The temporal precision of inputs to global form detectors may then be quite limited, but this is an open question. We used Glass patterns to investigate temporal limits on binding local forms into global form and on binding global form with color.

A further aim of these studies was to probe for a hypothesized temporal limit on subjective awareness. Holcombe (2001) and Holcombe and Cavanagh (2001) observed “temporal transparency”—gratings alternating faster than about 8 Hz seemed to be experienced together rather than as alternating. Interestingly, this phenomenon oc-
curred even in instances where the pairing of the orientation and color of the stimuli could be reported at alternation rates much faster than 8 Hz. The temporal transparency phenomenon suggests that by the time visual signals reach awareness, signals are integrated over about 120 ms. The dissociation between the temporal transparency limit and the color-form binding threshold indicates that some attributes are paired together before the stage of extended temporal integration.

Because it is unknown whether the long subjective integration phenomenon and dissociation with binding is a general property of visual awareness, we were interested to see whether they also occurred for global form and color. The reasons for using a stimulus alternation paradigm were twofold. One reason was our interest in the phenomenon of long subjective temporal integration--temporal transparency has no opportunity to occur with single presentation. Second was that with single masked presentation, temporally precise neural signals might not manifest in perception due to competition with the mask (di Lollo, Enns, & Rensink, 2000).

The Glass patterns (Glass, 1969) used here consisted of 2,000 pairs of dots. Together the local orientations defined by the dot pairs form global structure resembling a flow field (Figure 1A and 1B). To construct a Glass pattern, one dot in each pair is placed entirely randomly. The second dot is placed a short distance away, according to a vector that depends on the global form to be generated. To create translational global form, the vector separating the two dots of each pair is the same for all pairs. In the case of more complex form, the direction of the intra-pair separation vector depends on the pair’s position within the pattern. If, for all pairs, the intra-pair vector is perpendicular to the direction of the pair from the center of the pattern then the global form is concentric. If the intra-pair vector is oriented in the same direction as the center of the pattern, then a radial "sunburst" pattern is generated. If the intra-pair vector makes a ±45° angle with the direction of the center of the pattern, then a spiral pattern is created whose pitch, clockwise or anti-clockwise (Figure 1), depends on the sign of the angle of displacement.

Figure 1. Pairs of spiral clockwise and anti-clockwise Glass patterns were constructed from groups of four dots such that in their sums the two pairs were indistinguishable. B. The left and right panel both schematize two alternating spirals, using a very low density pattern, oversize dots, and a regular arrangement for illustration. The solid circles represent the dots presented on one of the two alternating frames, the outline circles the other. Corresponding pairs of dots form identical dot quartets for both pairs, which gives the pairs their identical sums. In the case of spiral patterns, the orientation of each quartet is 45° relative to a virtual line connecting it to the central fixation dot. C. Threshold rates of temporal alternation between stimulus pairs for 81.6% accuracy in identifying whether the alternating spirals were clockwise or anti-clockwise. Thresholds are shown for each of four subjects at each of two intra-pair dot separations.
Methods

Subjects in all experiments were two of the authors (CC and JP) and up to two experienced observers naïve to the purposes of the study (ER and WL). All had normal or corrected-to-normal vision.

Stimuli were generated using Matlab software to drive a VSG 2/5 graphics card (Cambridge Research Systems) and displayed on a gamma-corrected 21" Sony Trinitron GM 520 monitor (1024 x 768 resolution; 120-Hz refresh rate). Each stimulus frame, in addition to a central fixation spot, consisted of 4000 dots in a circular annulus with outer and inner diameters of 14.0° and 0.3°.

The basic unit in each Glass pattern was a pair, a quartet, or an octet of dots, depending on the stimulus condition, as described below and in the general text and figures. The position of each of the dot groups was randomly assigned according to a distribution uniform over area. Individual dots were then positioned within each group according to a global rule to produce patterns that could be spiral, translational, concentric, or radial, again depending on the stimulus condition. The overall duration of the train of stimuli was 667 ms. During this time, the stimulus was displayed at full (58.6%) contrast for the middle 333 ms and contrast was ramped on and off over the first and last 167 ms, according to a raised cosine envelope.

In Experiment 1, the basic pattern unit was a quartet of dots positioned at the corners of a square or diamond (Figure 1B). In separate conditions, the length of the side of the square was set to 0.18° or 0.50°. On alternate frames, alternate pairs of dots from within each quartet were displayed such that the overall pattern on all frames was either a clockwise or anti-clockwise spiral. Under no time pressure, subjects made a forced-choice judgment by pressing one of two buttons to indicate whether the stimulus contained clockwise or anti-clockwise global structure. Using the method of constant stimuli, the proportion of correct responses by each subject was recorded as a function of the duration of each stimulus frame. A Weibull function was fitted to the data for each subject, and the duration threshold was defined as the frame duration corresponding to 81.6% correct performance. From the duration threshold, the alternation threshold was calculated as the number of cycles (pairs of frames) in one second. For each alternation threshold an estimate of the associated SE was calculated using parametric bootstrapping (Efron & Tibshirani, 1998).

Experiment 2 investigated the spatial pooling of the mechanism underlying the global form judgment. The patterns were the same as in the first experiment, but in one condition the portion of the pattern exposed was reduced by 75% by showing 3 pie-shaped segments (Figure 2), as done by Wilson and Wilkinson (1998).

Experiment 3 was designed to address a concern that global form discrimination in the first experiment might be based on motion mechanisms. A basic pattern unit of eight dots was used, of which alternate sets of four were presented on alternate frames, positioned at the corners of a regular octagon of side length 0.50°. The orientation of each pattern unit was determined by its position relative to the center of the pattern. In one condition the alternating patterns were either (1) concentric and radial or (2) opposing spirals (Figure 3). Subjects made a forced-choice judgment between these two possibilities. In another condition translational patterns were used and observers judged whether the forms alternated between horizontal and vertical or between 45° and –45°. In other details the experiment was the same as the first.

In Experiment 4, observers again discriminated between clockwise and anti-clockwise spirals, just as in Experiment 1, but this time the stimulus was single brief exposure of one of the patterns rather than alternation. The brief exposure of the Glass pattern was followed by a noise mask presented for 500 ms. The mask consisted of randomly oriented dot pairs of the same size, intra-pair spacing, density and contrast as the Glass pattern (i.e., a pattern identical to the stimulus except at 0% coherence).

Experiment 5 investigated the binding of color with the global form. As in the first experiment, the basic pattern unit was a quartet of dots positioned at the corners of a square of side 0.18°. On alternating frames, the two dots from alternate diagonals of the square were presented. The color of the dots alternated along with the pattern, and subjects had to report the correct pairing of color and pattern (Figure 5) by pressing one of two buttons. In separate conditions, the patterns could either be left and right oblique translational Glass patterns or clockwise and anti-clockwise spirals. The colors were either red (CIE coordinates: .63, .34) and green (.28, .62) of equal luminance (26.0 cd/m²) or light (99.8 cd/m²) and dark (26.0 cd/m²) dots with the same chromaticity (.28, .30) as the 62.8 cd/m² background.

In Experiment 6, each subject was presented with a Glass pattern oscillating with period 600 ms between two pattern types (circular and radial) and two colors (red: .63, .34; green: .28, .62) of equal luminance (26.0 cd/m²). Intra-pair dot separation was 0.18°. The stimulus was presented until the subject made a response. The relative phase of the color and form changes was manipulated from trial to trial in steps of 12° (one-thirtieth of a temporal period), corresponding to 20-ms variations in the temporal alignment of color and form. During different runs of the experiment, subjects made judgments about color (what is the predominant color while the stimulus is red/green?) or form (what is the predominant pattern while the stimulus is radial/circular?). These reports, made by pressing one of two buttons, were recoded to represent the proportion of times that each color was paired with each type of form as a function of the relative phase of the oscillations. If the pairings were veridical, this distribution would be centered on physical synchrony (zero degrees of phase). The deviation of the centroid from physical synchrony was taken as a meas-
ure of the perceptual asynchrony of color and form processing.

In Experiment 7, we probed for a behavioral consequence of the subjective temporal transparency of the alternating global forms (Figure 7). All dots were dark and achromatic. There were two conditions, paired and unpaired. In the paired condition, the Glass pattern unit was again a quartet of dots positioned at the corners of a square of side 0.18°. Alternating pairs of diagonally opposing dots from each quartet were presented on alternate frames. On half of the trials, the dot pairs were chosen such that the global pattern alternated between clockwise and anticlockwise spirals. On the other half of the trials, on each frame 50% of the dot pairs were chosen from the clockwise spiral and 50% from the anticlockwise. Subjects were required to make a forced-choice judgment between these two possibilities. In the unpaired condition, the alternating pairs of dots were randomly positioned so did not sum to form a quartet. Subjects again had to identify whether the stimulus alternated between 100% coherent clockwise and anticlockwise spirals or between two spatially interleaved patterns.

In principle observers might “cheat” in each of these experiments by rapidly blinking or making large saccades so as to occasionally catch one of the stimuli without allowing the subsequent stimulus to mask it. This is a possible criticism of most masking experiments in the literature. However, in practice this strategy is difficult to use. In another investigation using the rapid alternation paradigm, extensive effort was made to monitor eye movements with all subjects but the results showed saccades and blinks were not a factor (Holcombe & Judson, 2004). In the present study, we are confident that our experienced observers made very few eye movements, and the results of the naïve observers did not systematically differ.

**Binding local form elements into global form is fast**

To determine the temporal limits of global form perception, we presented stimuli that alternated between two clockwise or anti-clockwise spiral Glass patterns. Each pair of patterns was composed of many groups of four dots (Figure 1B) arranged such that in the sum the two pairs of patterns were indistinguishable (Figure 1A). For a range of alternation rates, subjects attempted to determine whether the alternating spiral patterns were clockwise or anticlockwise. To make the discrimination correctly, local orientation information from across the visual field, represented in distributed form across the retinotopic map, must be bound together with a temporal precision corresponding to, or better than, the stimulus alternation rate.

The results of Experiment 1 revealed that the visual system can bind form elements together with remarkable temporal precision. Alternation rates for threshold performance (81.6% correct) averaged 22.0 ± 3.6 Hz when dots within a pair were close (0.18° apart) and 18.0 ± 1.9 Hz when dots within a pair were farther apart 0.50° (Figure 1C). To extract global form at a 20-Hz rate, the system must link disparate form elements together within the individual stimulus presentations of 25 ms. Furthermore, it does so via a very efficient spatial pooling process. Consider that, for the intra-pair dot separation of 0.18°, there was an average of 2.6 dots closer to any given dot than its partner dot within the pair. For the 0.50° separation, this number rose to 20. Hence, the relative position of nearest neighbor dots was not a reliable cue to the local orientation of the global pattern.

To better understand the spatial pooling underpinning the judgment of global form, we manipulated stimulus area in Experiment 2. In one condition, the full pattern was shown. In the other, only three pie-shaped segments of the pattern were shown, reducing total area by 75%. To quantify performance, we added randomly placed noise dots and measured the percentage of signal dots needed to perceive the global form. When the pair of Glass patterns alternated at 15 Hz, the 75% reduction of stimulus area increased signal percentage thresholds threefold (Figure 2). The size of this effect is in line with earlier work on static patterns, indicating that integration of information across the area of the pattern approaches the efficiency of the ideal observer (Wilson & Wilkinson, 1998; Morrone, Burr, & Vaina, 1995). Alternatively, the system might make preliminary decisions as to the global form on the basis of more local areas and then combine these decisions. The ideal integration result means that, instead, the system preserves the quantitative strength of the evidence for the global form to the stage where information is combined over a sizeable area. The finding depicted in Figure 2 implies that the system manages to accomplish this spatial integration just as efficiently at 15 Hz as in the static case. If signals were not maintained with fine temporal registration across the

![Figure 2](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933504/) Percentage signal thresholds for identifying static (0 Hz) and alternating (15 Hz) spiral patterns, as a function of stimulus area, for two subjects. That the slopes are equivalent in the two conditions indicates that the local forms are integrated over space with equal efficiency at low and high temporal frequencies.
pattern, efficiency would suffer as the evidence from the local regions would be degraded at the point of integration.

Under some circumstances, oscillations at high temporal frequency between static patterns are closely tied to the perception of motion (Victor & Conte, 2002). To ensure that apparent motion cues could not be responsible for the accurate performance at high temporal frequencies, for Experiment 3 we constructed sets of Glass patterns based on groups of eight dots rather than four so that any apparent motion between alternating frames would be inherently ambiguous (Figure 3). Depending on how the dots were paired between alternating frames, the resulting global structure alternated either between concentric and radial or between opposing clockwise and anti-clockwise spirals. As in the first experiment, in both cases the pairs were identical in the sum, ensuring that accurate discrimination would require temporal precision better than the alternation rate. Subjects were able to distinguish between concentric/radial structure and opposing spirals at temporal alternation rates of up to $20.9 \pm 2.2 \text{ Hz}$ (Figure 4). This temporal resolution is not significantly different to that obtained with the corresponding stimuli ($0.5^\circ$ dot separation) of the first experiment (paired $t = 1.23, p = 0.30$), indicating that form rather than motion is the cue for performance.

In the case of static Glass patterns, thresholds for different global forms reveal that humans are more sensitive to certain forms, such as concentric patterns, than to others, such as translational patterns (Wilson & Wilkinson, 1998). In Experiment 3 we also investigated whether the high temporal precision we found for perceiving global form was specific to particular patterns by comparing temporal thresholds with translational patterns to patterns with more complex form. Subjects distinguished alternation of horizontal and vertical patterns from alternation of oblique leftward-tilted and rightward-tilted patterns. Temporal frequency thresholds with the translational patterns were again remarkably high ($19.2 \pm 0.8 \text{ Hz}$) (Figure 4) and not significantly different from those for the more complex patterns (paired $t = 1.11, p = 0.34$). Hence, the scope of this high temporal resolution global binding process extends to a variety of Glass patterns, although complementary work on texture segmentation shows that it does not extend to arbitrary spatial arrangements (Motoryoshi & Nishida, 2001; Motoryoshi & Nishida, 2002; Forte, Hogben, & Ross, 1999).

Confirmation that the rapid extraction of global form is not restricted to conditions of stimulus alternation was provided by Experiment 4 in which subjects discriminated clockwise versus anti-clockwise spirals when the spiral pattern was presented in a single brief exposure and followed by a 500-ms mask of randomly oriented dot dipoles.

![Figure 3. A concern that global form discrimination in the first experiment might be based on motion mechanisms motivated the creation of alternating Glass patterns based on octets of dots. The left panel schematizes a radial Glass pattern (solid dots) alternating with a radial pattern (outline dots), and the right panel depicts alternation of opposing spirals. Apparent motion in either case should be ambiguous and rotational.](image)

![Figure 4. Alternation thresholds for discrimination of complex or linear patterns. As in the other experiments, the patterns were comprised of identical groups of dots (although presented at different times), in this case octets, such that temporal averaging over a cycle of alternation rendered them indistinguishable.](image)
Threshold performance was reached at durations of 17 ms and 21 ms for subjects CC and JP (data not shown), corresponding to the duration of a single stimulus presentation in the earlier experiments at alternation rates around 25-30 Hz. However, we should mention here that the results of more extensive experiments show that coherence thresholds are better with stimulus alternation, indicating that global form signals from successive intervals are pooled to some extent (Clifford, Holcombe, & Pearson, unpublished data, 2004).

**Binding color with global form is slow**

Pairings of local color and local orientation, when superposed, can be perceived at rates of nearly 20 Hz (Holcombe & Cavanagh, 2001). However, when the color and orientation features to be paired are spatially separated, the temporal limit for perceiving which color goes with which orientation drops to about 3 Hz. This dependence on spatial superposition has led to the suggestion that the high temporal resolution binding of color and orientation might be mediated by activity in an early visual area with very small receptive fields (Clifford et al., 2003), such as V1. Many cells in V1 carry information about multiple feature dimensions (McClurkin, Optican, Richmond, & Gawne, 1991; Lennie, Krauskopf, & Sclar, 1990; Johnson, Hawken, & Shapley, 2001). Consistent with the rapid perceptual alternation thresholds, these cells are also known to have high temporal resolution: Color-opponent cells in V1 respond to rates of chromatic modulation up to 30 Hz (Gur & Snodderly, 1997).

Global form is not extracted until brain regions situated later in the visual hierarchy, and the temporal limits on binding global form with color remain unexplored. Previous investigations show that while global form mechanisms exhibit limited chromatic tuning at equiluminance (Cardinal & Kiper, 2003), they are effectively color blind in the presence of luminance signals (Kovacs & Julesz, 1992). Similarly we find that, when summed, differently colored pairs of spiral Glass patterns appear to have the structure of the summed pattern rather than the colored components, so that the sum of red clockwise and green anti-clockwise spiral Glass patterns is perceptually indistinguishable from the opposite pairing unless extensive scrutiny is allowed (Figure 5). This constitutes a clear failure of color-form binding (see also Wilson, Wilkinson, & Assad, 1997), without the usual need to limit attention (Treisman, 1982) by employing lateral masking (He, Cavanagh, & Intriligator, 1996) or using brief exposures (Treisman & Schmidt, 1982). Glass patterns thus offer a chance to investigate the limitations on binding a high temporal resolution form signal that is extracted late in the visual system and hence is not coded in combination with color in early visual areas.

Experiment 5 investigated the binding of color with global form. While the spatial structure of Glass patterns can be extracted at rates of 18-22 Hz (Figure 1), and superposed color and orientation can be bound at nearly 20 Hz (Holcombe & Cavanagh, 2001), the temporal frequency threshold for binding color and global form is much lower (Figure 6). To ensure that the measured alternation threshold for binding color and form was not limited by the temporal resolution of color perception, we repeated the task using paired Glass patterns differing in luminance contrast rather than color. Although the maximum alternation rate for detecting temporal modulation of luminance is much higher than for chromatic modulation, we found that the alternation threshold for binding contrast polarity and the form of spiral Glass patterns was 4.9 ± 0.7 Hz, and for binding with color (red/green) was 4.5 ± 0.8 Hz. Thus, the

![Figure 5. Paired colored Glass patterns. In the sum, the two possible pairings of color and pattern (clockwise or anti-clockwise spiral) were indistinguishable without extended scrutiny. As depicted in the left panel, the stimulus was composed of dot quartets oriented just as in the previous experiments (see Figure 1B). The stimuli alternated between clockwise and anti-clockwise spirals of opposite colors, either red and green (shown) or light and dark. For a range of alternation rates, subjects attempted to report the color/pattern pairing. At the small scale of the figure, the sum may appear brown, but in the original stimulus the individual colors could be distinguished.](http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933504/ on 06/06/2017)
temporal frequency above which color and form could not be reliably bound was not significantly different for red-green and light-dark stimuli (paired $t_2 = 0.54, p = .64$). The close quantitative similarity of the two thresholds is not necessarily meaningful, because the light/dark and red/green differences were not equated in terms of cone contrast or multiples of detection threshold. Nevertheless, both stimuli were presented very far above detection threshold so the fact that both temporal limits are many times slower than the corresponding flicker fusion frequencies is meaningful. It suggests that binding of global form with color is limited by a later stage than that which limits simple color perception.

The similarity between the present threshold for binding color and global form and the threshold for binding spatially separated form and color elements (Holcombe & Cavanagh, 2001) suggests that both may reflect the slow stage which appears to limit arbitrary binding judgments—those not served by a specialized perceptual mechanism. Such judgments may require a time-consuming act of visual cognition or an effect mediated by cortical feedback.

Clifford et al. (2003) found evidence for a perceptual asynchrony between color and local form, with perception of color seeming to precede local form by about 50 ms in some circumstances. In Experiment 6, we varied the relative time of the color and global form changes to determine the perceptual asynchrony in two observers (details of this methodology are available in Clifford et al., 2003). The result was an asynchrony of about 40 ms in both observers, which is similar to that found with local form. Interestingly, Clifford et al. (2003) found that the perceptual asynchrony for color and local form gradually diminished as temporal frequency was increased, indicating that temporal frequency threshold is not limited by perceptual asynchrony.

Global form perception ultimately reflects integration across long intervals

While the global form of pairs of interleaved patterns can be identified at rapid 20-Hz rates, subjectively the temporal sequence of the stimuli appears to be lost. For example, rapid alternation between opposing spiral patterns yields the subjective experience of the two patterns coexisting rather than alternating. Previous work suggests that by the time information reaches awareness, visual signals have been averaged over long temporal intervals (Holcombe, 2001), but our present results demonstrate that certain complex forms are extracted first. This interpretation is further validated by a final experiment (Experiment 7) using spiral Glass patterns. Alternating pairs of diagonally opposing dots were presented on alternate frames (Figure 7). On half of the trials, the dot pairs were chosen such that the global pattern alternated between clockwise and anti-clockwise spirals. On the other half of the trials, both alternating frames contained 50% of the dot pairs from the clockwise spiral and 50% from the anti-clockwise. At high rates, subjects were unable to judge whether the two global forms were each presented at the same time or instead presented at entirely separate times. The threshold alternation rate for discriminating the two stimuli was 3.5 ± 0.3 Hz in the paired condition, where dot pairs from the two patterns were drawn from local groups of four dots (Figure 7), and 4.6 ± 1.0 Hz in the unpaired condition, where dot pairs from the two patterns were positioned independently from each other.

These two thresholds were not significantly different from one another (paired $t_2 = 1.81, p = .21$), showing that the rate at which the temporal sequence of the stimulus can be recovered is not critically dependent on its local spatial properties. Most importantly, these results confirm the subjective impression that the rapid 20-Hz rates at which global form can be extracted far exceed the rate at which judgments can be made about the temporal structure of the stimuli. At high rates, subjects are aware of the global forms presented, but do not have access to the individual temporal intervals (Figure 8).

Discussion

We demonstrated temporally precise binding of instances of local form into global form (Experiments 1-4). If, as is thought to be the case, local form signals are integrated into global form within a retinotopic high-level visual cortical area (Smith et al., 2002; Gallant et al., 1993; Hedgé & Van Essen, 2000, 2003; Wilkinson et al., 2000; Allison, Puce, Spencer, & McCarthy, 1999), then our result suggests that spatially disparate orientation signals are represented in precise temporal registration across the
Figure 7. Temporally mixed patterns were created by modifying the original clockwise spiral/anticlockwise spiral alternating stimulus; 50% of the dot pairs were temporally exchanged between the two frames, yielding alternation between two patterns that were 50% clockwise and 50% anticlockwise. The left panel shows the paired condition, in which dot pairs were arranged to form quartets. In the unpaired condition, dot pairs were positioned independently rather than arranged to form quartets. Temporal thresholds for discriminating between alternation of uniform spirals and mixed spirals were less than 6 Hz in both conditions, for each of three subjects (right panel). Figure 8 schematizes the percepts at various alternation frequencies.

Figure 8. Illustration of the percepts experienced at different rates for each class of alternating stimuli used in this work. The leftmost panel schematizes the global forms and indicates that they are experienced separately and alternating when presented at slow rates. At rates above approximately 4 Hz but slower than the global form identification threshold of approximately 20 Hz, observers continue to perceive both global forms, but they seem to be experienced at more or less the same time rather than in alternation. At still faster rates, the sum is perceived so one cannot distinguish between the pairs of corresponding global forms. At the intermediate rates (central column), in the case of colored Glass patterns (C), one still perceives green and red dots but can no longer determine which global form goes with each color. Also, observers cannot distinguish between alternation of intact spirals and alternation when dots of opposing spirals are temporally mixed, so that different parts of the spirals are presented at different times (D, bottom center).
QuickTime™ and a Photo - JPEG decompressor are needed to see this picture.

Movie 1. (view in loop mode) At alternation rates around 1 Hz, it is easy to pair color and form correctly.

Movie 2. (view in loop mode) At alternation rates around 5 Hz, the perceptual alternations between the two spirals give way to a perceptual superposition of clockwise and anti-clockwise spirals. The visual system no longer provides the correct pairing of the colors and forms.

retinotopic representation in extrastriate cortex. Specifically, we found that this computation was accomplished even at sustained 20-Hz rates, when each stimulus was present for only 25 ms (Experiment 1). Interestingly, however, this precision apparently only manifests behaviorally thanks to the elements’ arrangement into a coherent global form.

Previous work (Motoyoshi & Nishida, 2001; Motoyoshi & Nishida, 2002; Forte et al., 1999) has shown that simpler processes such as the extraction of texture edges also show high temporal precision. This may occur at an earlier visual area than that mediating Glass pattern perception (Smith et al., 2002; Hupe, James, Girard, & Bullier, 2001; Kastner, De Weerd, & Ungerleider, 2000). As for the temporal precision of processes more complex than that required for Glass patterns, Thorpe and colleagues have presented some intriguing results. They have shown that human brains can make the abstract determination of whether an animal is in a scene at latencies as short as 150 ms (Thorpe, Fize, & Marlot, 1996; Fabre-Thorpe, Richard, & Thorpe, 1998; Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001). However, follow-up work found that when low-level feature differences between categories are controlled for, the latency of the discriminating signal is highly variable, ranging from 150 to 300 ms (Johnson & Olshausen, 2003). Hence, the work of Thorpe and colleagues does not exclude the possibility that, despite impressive overall performance, the system might suffer from temporally imprecise signaling. Here we have shown that, at least in the case of the signals underlying the perception of global form, temporal precision is high.

For global form, the excellent temporal resolution in same attribute binding contrasts with the remarkably poor temporal resolution for binding this global form with a different attribute—color (Experiments 5 and 6). In viewing the alternating Glass patterns containing salient global structure, the subjects and experimenters were surprised by how much the alternation had to be slowed to perceive which color belonged to which pattern. The reader should be able to share this experience if the accompanying animated demonstration reproduces well on the reader’s display (Movies 1, 2, and Figure 9). This phenomenon shows that cross-attribute binding can impose a severe temporal limitation.

Figure 9. In the sum (or at alternation rates in excess of ~20 Hz), there is not time to extract global form before the next frame. Consequently, it is impossible to distinguish alternation between opposing spirals and alteration between concentric and radial patterns.
What is the neural basis of the present dissociation between cross-attribute and same-attribute binding? Recordings from neurons in V4 and other areas suggest that simple form and color are multiplexed on the same neural population (McClurkin & Optican, 1996). However, whether this is also true for the complex global form of Glass patterns is not yet clear. If color and complex global form are not multiplexed, then we may attribute the failure of binding to a lack of temporal precision across brain areas. But if the responses of individual neurons in V4 do code for both color and complex form, then the issue becomes why the temporal resolution of binding color and global form is not better.

Simple multiplexing of information on the same neurons may not be sufficient for perceptual binding. Consider that color information appears to develop at the same rate as pattern information in V1 neurons, but develops at different rates in the responses of V2 and V4 neurons (McClurkin & Optican, 1996). Hence, one possibility is that the failure of binding global form with color may be due to a lack of temporal correspondence between multiple feature codes within a brain area. However, this is not the theory that we favor.

From a psychological perspective, the inability to determine the color-form pairing is surprising because one might have thought one would be able to simply selectively attend to one of the colors and then identify the corresponding global form. However, the observer’s inability to do this is in accord with evidence from visual search that one cannot attend to color per se. Instead, the locations of the color must be identified first and attention then activates the corresponding spatial locations (Shih & Sperling, 1996; Moore & Egeth, 1998). The ineffectiveness of attention to color when brief displays are used (Shih & Sperling, 1996; Moore & Egeth, 1998) suggests that this process is time-consuming. Hence, although the presentation of the unicolored Glass pattern when in alternation may inform the observer of the locations of the dots belonging to a particular global form, this may take too long to be of use at greater than 4-Hz alternation rates. And at fast alternation rates, the detectors corresponding to the two global forms both point to the same region, so that the observer does not know which dots belong to a particular global form. Furthermore, the system appears to be modular in that the global form detectors have been reported not to signal color (Kovacs & Julesz, 1992; Wilson & Wilkinson, 1998), although Cardinal and Kiper (2003) did observe limited chromatic tuning in the detection of equiluminant Glass patterns.

An analogue of temporal transparency (Holcombe & Cavanagh, 2001) was experienced when two Glass patterns were alternated rapidly—observers reported that both patterns were experienced simultaneously. This result provides further support for the hypothesis that visual awareness reflects extended temporal averaging (Verstraten, Cavanagh, & Labianca, 2000; Holcombe & Cavanagh, 2001). In this case, global forms are extracted from small temporal windows. These representations of global forms are then combined over longer intervals, preventing conscious access to which form was presented when (Experiment 7). However, an important difference from the temporal transparency observed for gratings is that there the color remains bound to the orientation of the grating (Holcombe & Cavanagh, 2001), while with Glass patterns the associated colors are lost (Experiment 5).

We have seen that the orchestration of perceptual binding may be limited by the modular arrangement of visual processing subsystems. The binding of local form into global form is accurate at high rates of oscillation, while the binding of global form and color fails at those same rates. Our findings add to the many recent results that have added strong constraints for any theory of binding. Describing all these constraints is beyond the scope of this article, but in a pending book chapter we attempt to provide a more complete account (Clifford, Holcombe, & Pearson, 2004).

**Acknowledgments**

This work was supported by a Queen Elizabeth II Fellowship and Discovery Project Grant awarded to CC by the Australian Research Council. Support was also provided by a postdoctoral National Eye Institute National Research Service Award to AOH and by National Eye Institute Grant EY01711 awarded to D. MacLeod. We are grateful to Peter Neri, David Eagleman, Marcus Taft, Bart Krekelberg, and David Alais for helpful suggestions during the course of this work.

Commercial relationships: none.
Corresponding author: Colin W. G. Clifford.
Email: colinc@psych.usyd.edu.au.
Address: Color, Form & Motion Lab, Visual Perception Unit, School of Psychology, The University of Sydney, NSW 2006, Australia.

**References**


