Independent, synchronous access to color and motion features

Alex O. Holcombe[a,*] and **Patrick Cavanagh**[b,c]

[a] School of Psychology, University of Sydney, Sydney, NSW 2006, Australia

[b] Department of Psychology, Harvard University, Cambridge, MA, USA

[c] Département de Psychology, Université de Paris 5, Paris, France

Abstract

We investigated the role of attention in pairing superimposed visual features. When moving dots alternate in color and in motion direction, reports of the perceived color and motion reveal an asynchrony: the most accurate reports occur when the motion change precedes the associated color change by ~100 ms [Moutoussis, K., & Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London B, 264*, 393–399]. This feature binding asynchrony was probed by manipulating endogenous and exogenous attention. First, endogenous attention was manipulated by changing which feature dimension observers were instructed to attend to first. This yielded little effect on the asynchrony. Second, exogenous attention was manipulated by briefly presenting a ring around the target, cueing the report of the color and motion seen within the ring. This reduced or eliminated the apparent latency difference between color and motion. Accuracy was best predicted by timing of each feature relative to the cue rather than the timing of the two features relative to each other, suggesting independent attentional access to the two features with an exogenous attention cue. The timing of attentional cueing affected feature pairing reports as much as the timing of the features themselves.

Keywords

Binding; Object perception; Attention; Feature asynchrony; Neutral latency

1. Introduction

Different aspects of a visual object, such as its color and its motion, are to some extent processed by separate populations of neurons (Zeki, 1978). How these separate representations might be joined for unified perception is not understood. Adaptation studies show that at early stages, some features may be already combined or, perhaps, not yet separated (Blaser, Papathomas, & Vidnyanszky, 2005; Humphrey & Goodale, 1998; Vul & MacLeod, 2006). Attention seems to have no role in constructing these early, multiplex representations which may reflect conjoint tuning in single cells for multiple features such as color and motion (Croner & Albright, 1999; Dobkins & Albright, 1998). Whether such conjoint representations underlie the conscious perception of feature pairings is unknown.

At high-level stages of visual processing, features are apparently more separated, as lesions to parietal areas can selectively disrupt judgments of feature pairings (Friedman-Hill, Robertson, & Treisman, 1995). Recent evidence shows that the process pairing color and motion is capable

* Corresponding author. Tel.: +61 4 1685 5396; fax: +61 2 9036 5223. alexh@psych.usyd.edu.au (A.O. Holcombe)..

of accumulating information from disparate locations, consistent with the involvement of areas with large receptive fields like in parietal cortex (Cavanagh & Holcombe, 2005; Cavanagh, Holcombe, & Chou, submitted for publication). That the parietal areas apparently involved in binding may also mediate the allocation of attention (Serences & Yantis, 2007) supports Treisman's influential proposal that attention computes feature pairings by selecting a particular location and its associated features (Treisman & Gelade, 1980).

Perception of feature pairs can be limited to surprisingly coarse timescales. For example consider a patch of dots alternating in color between red and green and in motion between leftward and rightward. At very slow rates, it is easy to perceive the pairing of the color and motion. However, when the stimulus changes color and motion faster than about five times a second, it is difficult or impossible to determine the pairing (Arnold, 2005; Moradi & Shimojo, 2004) even though the features can easily be identified. The low temporal limit is consistent with the idea that a slow, high-level process like attention (Duncan, Ward, & Shapiro, 1994) pairs the features. In contrast, many perceptual computations that rely on specialized, preattentive detectors are very fast (Burr & Ross, 1982; Clifford, Holcombe, & Pearson, 2004a; Morgan & Castet, 1995).

In this paper, we consider the possibility that conscious reports of the features of an object reflect independent access by attention to each separate feature. We contrast this possibility to an alternative where the features of an object may be already joined together when accessed for report. We call this alternative a "zipper" mechanism. The zipper pairing of the features does not have to precede attention's access to support the predictions we will test. This "zippering" process may be preattentive and occur regardless of attentional state, or alternatively, it may be the act of attending that actually triggers the "zipping together" of the features. But with the zipper theory, crucially, access is to already paired features so any timing difference between the features that affects the accuracy of reporting the pairing of features should occur independently of any manipulation of attention. The feature pair reported is determined by relative timing of the two features. In contrast, for the theory of independent access to the separate features, it is the time of each feature relative to the attention cue that determines accuracy for that feature. Because of its link to the timing of an attention cue, independent access makes different predictions than does the zipper theory for the results of our Experiments 2 and 3.

We adapt a procedure developed by Moutoussis and Zeki (1997) and present color that alternates between red and green together with motion that alternates between inward and outward. We vary the relative timing of the color and motion reversals and evaluate the reports of which color is seen with which motion. In the first study using this method, Moutoussis and Zeki (1997) found that subjects' reports of the color and motion appeared most clearly paired (say, red moving only one direction followed by green moving only the opposite direction) when the motion reversals occurred about 80 ms before the color reversals: in other words, two features were best perceived as paired when they physically were asynchronous. From this result, Moutoussis and Zeki (1997) and others have inferred that color has a shorter sensory latency than motion. Such bottom-up latency theories of the asynchrony are compatible with a zipper that pairs features prior to access by attention in that both theories assume a baseline latency difference between the two features that is unaffected by attentional manipulations. However, we will report that the asynchrony does depend critically on attention, specifically, the timing of an exogenous attention cue. The timing of the attentional cue relative to each feature's onset is more important than the timing of the features relative to each other. This pattern of data fits with independent attentional access to features when reporting their pairing.

A previous study already examined the timing of access to the features as an explanation of the feature asynchrony found by Moutoussis and Zeki (1997). Enns and Oriet (2004)

investigated the possibility of sequential access determined by conscious strategy by varying the instructions to their subjects. They claimed that the asynchrony was reversed when subjects attempted to attend first to color and then reported the pairing compared to attending first to motion and then reporting the paired color. As will be explained below, however, we believe subjects did not use the attentional strategies that the experimenters intended.

Nishida and Johnston (2002) were able to dramatically alter the apparent asynchrony between color and motion, not by changing the instructions, but rather by manipulating the nature of the feature transients. The change in the transients may have had their effect via changing the attentional salience of the features, although they did not discuss the results in terms of attention.

In Experiment 1, we evaluate and discard the possibility that endogenous voluntary shifts of attention mediate the access to the features (as had been proposed by Enns & Oriet, 2004). Then, in Experiments 2 and 3, we manipulate the timing of attention directly with an exogenous cue, a ring, that suddenly highlights the stimulus. Observers were asked to report the feature pairing perceived while the cue was present. This manipulation allowed us for the first time to determine whether the critical variable for binding was the time of the two features relative to each other or the timing of each feature relative to the attention cue. In fact the latter was more important. When the cue onset was aligned to the motion change or the color change, the apparent color–motion asynchrony diminished to almost zero. This was a special case of the more general effect that each feature was best reported when it was synchronized with a cueing ring, regardless of the timing of the other feature. Each feature was thus reported as if it were accessed independently rather than being retrieved from a bound description. The asynchrony usually found may thus reflect an idiosyncrasy in how the two features are sampled by attention when in an uncued stream.

2. Experiment 1: Endogenous attention shifts do not explain the asynchrony

In a conference presentation, Enns and Oriet (2004) concluded that the asynchrony typically found can be reversed by changing the attentional strategy of the observers. In that study, half of participants were told to monitor for a particular motion feature and report the paired color. The other half were asked to monitor for a particular color feature and report the paired motion. A large difference between the groups was found. When participants monitored motion first and the color changed halfway through the motion period, they were more likely to report the second color than the first color. In contrast, for those participants who monitored for a particular color, when the stimulus changed motion direction halfway through then participants tended to report the second motion as paired with the color.

The results of Enns and Oriet (2004) are surprising because others had previously tried manipulating which feature was attended first and found little effect (Arnold, 2005; Clifford, Arnold, & Pearson, 2003). In one of these studies (Arnold, 2005), it appears that the same procedure was used as in Enns and Oriet (2004) – half of observers were told to monitor for a motion feature, and half to monitor for a color feature, with the task to report the pairing with the other feature. However, in that study only one or two naive observers were used in each group so the size of the effect, if any, remains uncertain.

Here, we use a display with 10 patches of dots in the periphery circularly arrayed around the fixation point. The dots in each of the patches alternate between moving towards the fixation point and away while their color alternates between red and green. Described in detail below, this display configuration is critical for later experiments in this paper. Despite the instructions to reverse the order of attention to the two features, the results show that the asynchrony persists, with observers most often reporting red inward when the green–red transition occurs 100–200 ms after the outward–inward transition.

As Enns and Oriet (2004) reported that a carryover of attentional strategy from other experimental conditions could dilute or eliminate the effect, we used only observers who had no experience with this feature binding task. Half were told to monitor for red and report the motion direction that dominated for the red dots, and the other half were asked to attend to inward motion and report the color that dominated during inward motion.

2.1. Methods

2.1.1. Stimulus—On a CRT controlled by a MacOS 9 computer with a custom C program that utilized the VisionShell C library (Comtois, 2003), ten random dot fields were presented arrayed around a circle, equally spaced (Fig. 1). The center of each 1.9° dot field was 5° from the white fixation point. Each dot field had a black background (<2 cd/m^2, but greater than scotopic levels) whereas the rest of the screen background was a dark grey (6 cd/m^2, CIE $x = .32$, $y = .40$). Each dot field contained approximately 30 square dots 0.15° wide dispersed randomly over the area, with sides of the dots oriented along the axis of motion. When the dots moved, new dots were revealed at the trailing edge, and dots at the leading edge disappeared, as if the dot field were moving behind an aperture. Hence, when the dot motion changed to the opposite direction, the dots that had disappeared became visible again. The sides of the virtual aperture were aligned with the axis of motion. The dots alternated between moving inwards directly toward the fixation point and moving directly away, at 7.2°/s over a period of 753 ms (1.33 Hz), yielding approximately 377 ms per color or motion. With the same period, they also alternated between red (40 cd/m^2, CIE $x = .63$, $y = .33$) and green (CIE $x = .29$, $y = .59$). The luminance for green was set for each participant individually to match their approximate equiluminance point – determined by flickering the dot fields between red and green at 8.5 Hz and adjusting the intensity of the green to minimize the subjective intensity of the flicker (Wagner & Boynton, 1972).

Adjacent dot fields were in counterphase – when a given field moved inward, its immediate neighbors moved outward, and when a given field was red, its immediate neighbors were green. To premask the stimulus train, the stationary dot fields appeared gradually by means of a roughly linear contrast ramp over the first three stimulus presentations (1.5 cycles of alternation) during which time the stimuli alternated between red and green at the same frequency as in the rest of the train. Subsequent to this premask period, the dots began moving and were presented at full contrast for 3 cycles. Finally, the dot fields disappeared in the same fashion that they had appeared – they became stationary and gradually diminished in contrast. The gradual ramping in and out avoided a sudden onset that might affect the asynchrony, as discussed in Experiment 2.

Fig. 1 shows a snapshot of the stimulus. Also provided is an interactive Flash (Adobe, Inc.) movie which allows one to manipulate the alternation rate, relative phase of the color and motion changes, and the ring marker used in Experiments 2 and 3 (http://www.psych.usyd.edu.au/staff/alexh/research/asynchronyVSS06/).

2.2. Design and procedure

The independent variable for the experiment was the in → red interval, which is the time between the onset of inward motion and the onset of red (Fig. 1). The in → red interval varied over the whole range of possible temporal offsets, taking a total of 16 different values (–377, –330, –283, –236, –189, –141, –94, –47, 0, 47, 94, 141, 189, 236, 283, 330 ms). Variation in the in → red interval yielded four physical pairings: in with red, in with green, out with red, and out with green. When the in → red interval was 0 or –377 ms, the features were exactly synchronized, so that the display consisted of only two stimuli. For example, for in → red interval = 0, the display alternated between in-red and out-green. These are considered the same pairing for our purposes and when this pairing was perceived, observers made a 'z'

keypress in this experiment. If the display instead was perceived to alternate between in-green and out-red, as it physically did for in → red interval = –377, they pressed '/'. For most values of the in → red interval, the stimulus physically presented comprised four feature pairings, in-red, out-red, out-green, out-red, in varying amounts. Observers were instructed to determine the predominant feature pairing in this situation. Two values for in → red interval should lead to the most unambiguous percepts, yielding the most consistent pairing reports. These are not expected to be the values where the stimulus was physically least ambiguous (0 and –377 ms), but instead to reflect a difference in when the color and motion are processed.

With 12 trials per each of 16 possible in → red intervals, there were 192 trials per subject. Two additional variables were which motion appeared first in the overall stimulus train and the phase in the cycle at the onset of the stimulus. These values were randomly chosen on each trial.

Participants fixated on the fixation point but could attend to any or all of the dot patches. At any point during the stimulus train, participants could make their response, pressing 'z' to indicate that they predominantly perceived red inward alternating with green outward, or '/' to indicate red outward alternating with green inward. These responses were described differently depending upon the attentional group the person was in. The attend red group was told to attend to the red phase and determine the paired motion, entering 'z' for inward and '/' for outward, whereas the attend motion group was told to attend to inward motion and report the paired color. The participant's keypress terminated the trial immediately (even if the stimulus was still being presented) and presentation of the next stimulus soon commenced.

Participants viewed the CRT screen from a distance of approximately 1 m. They participated in a variable number of practice trials until they were comfortable with the task. The plots of red inward responses against in → red interval were fitted by cosine functions as in previous literature (Bedell, Chung, Ogmen, & Patel, 2003; Clifford et al., 2003) and the functions were truncated or "capped" at 100% and 0% to accommodate the possible range of response proportions. In other words, the fitted cosine functions had a mean of 0.50 and when the best fit amplitude exceeded 0.5, predicted values above 1 were capped at 1 and predicted values less than 0 were set to 0. Standard errors were determined by bootstrap—the data were resampled without replacement 200 times and the standard deviation was taken of the phases of the fit curves for each (Efron & Tibshirani, 1993). In tests of the validity of these measures, these standard errors were found to not differ significantly from the more exact method of taking percentiles of the distributions of the bootstrap sample phase fits.

2.3. Results

Fig. 2 shows the data for each of the 10 participants of this first experiment where subjects either monitored for a particular color and reported the associated motion, or instead monitored for a particular motion, and reported the associated color.

Consider the top 10 plots of Fig. 2. Each represents a different naïve observer. The column of five plots on the left represents the observers who were told to pay attention when red appeared and report the predominantly paired motion direction. The independent variable was the relative timing of color and motion, represented on the horizontal axis as the "in → red" – the temporal interval between the moment when motion reverses from outward to inward and when the color reverses from green to red. In the case of observer CB, the data show that when the color and motion changes were perfectly in sync (in → red interval = 0) she reported "inward" – the correct answer – only about 50% of the time. The temporal phase relationship that led her to report that inward was paired with red most often was an in → red interval of 50–220 ms. When red appeared between 50 and 220 ms after the dots began moving inward, CB reported red inward at least 80% of the time. As the in → red interval increases, more and more

of the red period is during the outward phase of the motion. When the in → red interval is 377 ms, the onset of red coincides exactly with the onset of outward motion (red moves only outward, followed by green moving only inward. We have plotted the data for this interval at the –377 ms position on the graph as 377 ms and –377 represent the same point in the periodic cycle). For these trials of 100% alignment in the stimulus, observer CB reported red inward slightly less than 50% of the time, indicating that the perception of red moving outward was only slightly favored (rather than completely dominant as it is in the physical stimulus).

These data are summarized by fitting a cosine curve with the time of its peak as the best in → red interval – what previous authors have surmised to be the perceptual asynchrony (Moutoussis & Zeki, 1997). In the case of observer CB, the peak according to the capped cosine fit is at an in → red interval of 136 ms. Most researchers have explained this ~100 ms shift in perceptual alignment with the proposal that color has a shorter processing latency than motion, so that for optimal perceptual alignment, color should be presented later. However, Enns and Oriet (2004) found some evidence that the asynchrony could instead by caused by voluntary attention switching. On this account, the color lag might reflect a tendency to attend to the motion first before attending to the color.

The corresponding comparison is between the plots on the left versus on the right of Fig. 2. The column of plots on the right, of participants who monitored for inward motion and then reported the corresponding color, do indicate a slightly larger peak in → red interval. The peak values for the different observers are plotted together in the bottom summary plots of Fig. 2. Each observer's best-fit in → red interval is marked with their initials, and the mean is indicated by the right-hand circle for those observers who monitored for red (mean = 99 ms), and at right by the right-hand square for those monitoring for inward motion (mean = 134 ms). The mean red lag is hence greater by 35 ms for those who were told to monitor the stream for inward motion rather than color, although this difference is not statistically significant ($t(4) = 1.24$, $p = .25$). As is clear from the error bars on the means for the two groups, this trend for an effect of attentional strategy is small and certainly not enough to overcome and reverse the baseline effect of >100 ms lag for the color change.

Why were the results of Enns and Oriet (2004) different from ours and those of previous studies (Arnold, 2005; Clifford et al., 2003)? Enns and Oriet measured performance with only four relative timings of the motion and color transitions: two with the transitions in phase (red and up or green and up) and two with them out of phase (color switching in the middle of the motion and vice versa). The asynchrony measurement then depends on the two intervals where the motion and color were not synchronized, and for these the color changed halfway through a motion interval (relative phase of 90° and 270°). To speculate, participants may have noticed that half of trials were very difficult and possibly chosen to report the last seen value. For example, in the attend-red, report motion condition they would report the motion of the second half of the red interval. In the attend up, report color condition, they would report the color of the second half of the upward interval. The strategy would introduce a strong apparent asynchrony in the data that would reverse in the two attention conditions. We have no evidence that this is what Enns and Oriet's subjects were doing but in our experiments, there were many different relative timings and this strategy would be difficult to implement.

The results of the present experiment, in conjunction with the previous work of Clifford et al. (2003) and Arnold (2005) give confidence that voluntary switching between feature dimensions does not account for the better part of the asynchrony effect for this stimulus. Nonetheless, we do think that attention plays a critical role in mediating reports of feature pairing, but it is the stimulus transients that control attention (in the absence of conditions like those of Enns and Oriet) and its access to features. In the next two experiments a new transient is introduced, a cueing ring, that has a powerful effect on best feature timing. The cueing ring

is similar to classic exogenous cues and here it can be shown that it determines the timing of the feature access needed to pair features.

3. Experiment 2: Exogenous cueing

Nishida and Johnston (2002) have suggested that the color–motion asynchrony and the pairing mechanism that underlies it is controlled by the transients separating the features rather than processing of the features themselves. This in itself does not speak to whether the pairing process is a preattentive process or instead reflects independent attentional sampling of the features. However, if the pairing process is attentive, it should be affected not just by transients intrinsic to the color and motion but also by the external transients of an exogenous cue, classically the most effective stimulus for engaging attention (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989; Posner, 1980). In contrast to the power of an exogenous cue, the instructions in our first experiment here may have been ineffective in changing the time at which features were accessed because the feature transients might be too salient to ignore; they would drive access even when subjects were instructed to follow a different order of access.

If this were the case, then the asynchrony might be eliminated by keying the access to the color and motion features with an external transient rather than depending on the transients of the features themselves. For external transients, we use a ring (Fig. 3) that suddenly onsets around a moving colored dot field and ask participants to judge the feature pairing while the ring was present. The ring is an exogenous attentional cue, effectively providing transients that indicate the time and place to attend (Cheal & Lyon, 1991;Nakayama & Mackeben, 1989).

The relative timing of the ring, the color alternations, and the motion alternations were randomized over the entire range. This eliminated any possibility that bias might contaminate the color–motion asynchrony, or that the ring time could be used by the subject to infer something about the color or motion relationship. Observers did not know which of the circular array of dot fields was the target until the ring appeared around the target dot field, which it did at an unpredictable time, around an unpredictable dot field.

We parameterize the data with respect to two variables (Fig. 4). First, the time of the ring relative to the color changes (red $\rightarrow$ ring interval). Second, the time of the ring relative to the motion changes (in $\rightarrow$ ring interval). Each variable took one of 8 possible values on each trial, yielding 64 conditions.

If the feature reports are based simply on attention picking up the features during the ring's presentation, then the best relative timing of color and motion for consistent reports may depend on the timing of the ring relative to each feature. In contrast, feature reports may instead depend on the timing of the color relative to the motion, as would be expected if binding is based on the relative time of the features to each other rather than simply being those swept up by attention.

3.1. Methods

The red–green in–out stimulus was very similar to that of Experiment 1. Here, seven oscillating dot fields were presented in a circular array about fixation (Fig. 3, left). Each of the seven dot fields began at a phase in the alternation cycle that was chosen randomly on each trial. The relative timing of the color and motion alternation (in $\rightarrow$ red interval) was also assigned pseudorandomly to each dot field. This randomization prevented observers from judging the pairing before the critical target was revealed by the ring onset. The color–motion pairing randomization was more specifically that the six dot fields other than the target were each assigned a distinct in $\rightarrow$ red lag that together spanned the full range of possible in $\rightarrow$ red lags, with equal spacing. This and the random distribution of absolute phases of the dot patches'

motions yielded the impression of irregular contraction and expansion rather than the counterphasing of the previous experiments. The distribution of phases prevented any possibility that the observer might use the rhythm to help time the use of attention, although pilot experiments with two observers (data not shown) found no noticeable effect of this change. The target to be encircled by the cue was chosen randomly on each trial. However, in a pilot experiment in which two observers were told the target location in advance but instructed not to identify the feature pairing of the target until the cue appeared, the pattern of results was similar to those found here.

A slower rate was used here than in Experiment 1. In another experiment (Experiment 4), this change in rate was found to have no effect on the asynchrony. The stimulus color and motion changed every 470 ms, with the cuing ring presented for 470 ms around one of the 7 dot fields randomly selected on each trial.

The ring was an unfilled white circle (80 cd/m^2, diameter = 2.4°, thickness = .22°) and it appeared for one half-cycle, 470 ms. As mentioned in the introduction to this experiment, the time of appearance of this ring relative to both the motion and color changes (Fig. 4) varied through the whole cycle, on each trial taking one of eight values evenly spaced through the feature alternation cycle: –354, –236, –118, 0, 118, 236, 354, and 472 ms. With in → ring and red → ring intervals independently varied, 64 conditions were created. For consistent pairing reports the observer would have to perceive both the color and the motion, as there was no systematic mapping between ring, color, and motion. Each observer participated in eight trials for each of these 64 conditions, which were presented in pseudorandom order.

Participants reported the predominant feature pairing perceived in the dot field enclosed by the ring, during the interval the ring was presented. With no time pressure, the participant hit one of four keys: 'z' for red-in or 'x' for red-out, using the left-hand, or ';' for green-in or ''' for green-out, using the right-hand. The participant's keypress terminated the trial immediately (even if the stimulus was still being presented) and presentation of the next stimulus soon commenced.

The same style of on- and off-ramp of the presentation was used as in Experiment 1. In addition to the ramp, two cycles at full contrast were presented. Four observers participated, including the author AH. One participant was discarded because following the experiment he spontaneously reported that he did not follow the instructions. Pilot experiments for each observer without the ring revealed the usual asynchrony result of best inward motion onset before the red onset (data not shown). For comparison with asynchronies obtained when no ring was present, the data from the in-phase oscillating stimulus of the same rate described in Experiment 4 was used.

3.2. Results

3.2.1. Synchrony revealed—
Previous literature analyzed the color–motion timing that yielded the most consistent reports of the color and motion pairing. Here, however, we were also interested in any effect of the timing of the ring cue relative to each feature. To begin, to allow simple comparison to previous work we analyze the reports of color and motion as a function of color–motion timing in only those trials in which the ring was synchronized with inward motion. Fitting a capped cosine curve to the responses as a function of interval between the onset of the inward motion and the onset of red color, the peak of the curve (the highest rate of correct color–motion pairing reports) occurs close to when the timing of the motion change coincides with that the color change. This indicates that there is little or no color–motion asynchrony in this case (purple symbols in Fig. 5). This result contrasts strongly with the result when no ring was presented, where the peak occurs when the inward motion begins 100 ms or more before the dots turn red (black symbols). Now considering those trials where the ring was

synchronized with red, we again find little or no asynchrony (red symbols). For AH and CH, the peak in → red intervals were close to zero (–34 and 2 ms) and for HO, the 65 ms found amounts to a reduction of 133 ms from that when the ring was not presented (black symbols).

The differential latency hypothesis (e.g. Moutoussis & Zeki, 1997), which considers only the timing of the features relative to each other, cannot explain these results, as the ring should not differentially affect latencies for color and motion. To represent the full set of results and reveal the effect of ring timing, we plot the entire set of data as a function of the two ring timing variables, yielding a three-dimensional surface plot (for comparison, in Appendix A we also present the data in a format closer to that of previous studies).

In the surface plot showing the entire dataset for an observer, the dependent measure is the proportion of times the observer responded with the pairings of red-in or green-out (for our purposes, these are equivalent responses). The independent variables are in → ring interval and red → ring interval. For a particular oscillating stimulus, we are indifferent to whether the ring encircles red or encircles green. Instead, the critical variables are the relationship of the ring timing to the motion or color changes, regardless of which half-cycle the ring encloses. For example, the condition where the color and motion changes are synchronized and the ring is synchronized with green outward is the same for our purposes as that condition for which the ring is synchronized with the red inward phase. We can therefore combine together two halves of the data as shown in Fig. 6. In each case the relationship between color and motion, and the relationship of the ring to the color and motion transients, is the same. This collapses the 64 conditions into 32, representing the factorial combination of 4 red → ring intervals spanning –.125 and .25 and 8 in → ring intervals spanning –.375 and .5.

Plotting the collapsed data (Fig. 7) yields hill-shaped surfaces. The peak and lowest point of the surface are the two ring timings that yielded the most consistent responses, indicating where the feature pairing is perceptually least ambiguous. The position of the peak with respect to the red → ring axis is the best ring time relative to the color transitions, and its position with respect to the in → ring axis is the best ring time relative to the motion to elicit reports of in with red (or out with green, due to the data collapsing described previously). To determine these peak values, the data were fitted with a model (justified in the following section), and the model fits are plotted below each observer's data plot in Fig. 7. The fitted parameters are shown in Table 1.

The red → ring and in → ring parameters are the estimates of the position of the peak on the axes, in milliseconds. For participant AH, the ring time that yielded the most consistent pairing reports is 28 ms after the red onset and 0 ms after (synchronized with) the onset of the inward motion. If there were an actual difference in latency between the color and the motion, the best in → ring interval should be at least 100 ms larger than the best red → ring interval, to give the motion the head start it needed in previous experiments. Instead, this difference (last column of Table 1) is –28 ms, meaning that the peak of the surface corresponds to a stimulus with little difference in timing between the color and its corresponding motion. The results are similar for the other participants. In no case did the difference approach the 100–200 ms difference between color and motion time found when no ring was presented (grey symbols of Fig. 5). The delays between the red onset and ring onset that produced the most consistent responses are moderate (about 50 ms across the 3 subjects) as is the optimal delay between inward motion onset and ring onset (average about 30 ms). These results suggest that any latency between the response to the ring and access to the color or the motion is small.

Here, in the presence of conspicuous external transients (the onset and offset of the ring) cueing attention, little or no asynchrony is found. With the best timing of color and motion so clearly driven by the timing of exogenous attention, one must consider that perhaps the original

baseline asynchrony is itself a product of the particular mode of attentional access in those experiments. Perhaps in the absence of strong extrinsic transients from the ring, the asynchrony is produced by the order of access to the two features and this is set by the intrinsic salience of the color and motion transients. In this case, instead of attributing the asynchrony between color and motion to the voluntary switching of attention between the two features as Enns and Oriet (2004) did, we suggest it is set by involuntary asynchronous attentional access driven by the intrinsic transients of the stimulus. We argue that it is involuntary because our subjects were not able to overcome the order of access set by the intrinsic transients to read out the features in the instructed order (Experiment 1).

We suggest that the strong extrinsic transient of our cue ring is unrelated to either feature and so it drives access to both features equally well. Since access is not determined by either feature, the order of access may vary randomly from trial to trial and if there is asynchronous access with one beginning before the other, the delay would not consistently favor either. Alternatively, in this externally triggered condition, the access might be synchronous.

3.2.2. Features accessed independently rather than together—In the previous section, we discussed the implications of the position of the data peak for the relative latency of the color and motion and the binding mechanism. The shape of the data are also diagnostic of the nature of the feature binding mechanism. The introduction to this paper distinguished between two putative classes of binding mechanisms. Potentially, the features might be accessed independently so that only the timing of the attentional cue relative to each feature determines the accuracy of that feature's report. For any given timing of the cue, then, the most consistent response for one feature occurs when it is aligned with the cue, irrespective of the timing of the other feature. In contrast, if attention accesses features through a separate pairing mechanism (which we call zipper binding), then the relative timing of two features is important. For any given timing of the cue, the most consistent response for one feature should occur when it is aligned, not with the cue, but with the other feature. This alignment occurs when the stimulus timing is such that it offsets any sensory asynchrony between the two features prior to pairing. On our surface plot, such a constant color–motion asynchrony would yield an optimal timing for report that followed a diagonal, 45° ridge as pictured at top of Fig. 8. The prediction pictured is for a 150 ms asynchrony between motion and color; other values for the asynchrony would shift the ridge in one direction or another without changing its orientation.

For independent access, a difference in sensory latency between color and motion will also affect the optimal timing. However, unlike with the zipper mechanism, each feature is accessed independently so the feature reports will reflect simply the value of each feature that perceptually dominates during the cued interval, irrespective of the value and timing of the other feature. In other words, the probability of reporting a particular motion reflects only the proportion of the time that the motion was present along with the ring (after any difference in latency between ring and feature is taken into account). The same is true for color, with the probability of reporting a particular color reflecting only the time of the color relative to the ring. So unlike in zipper binding, the optimal temporal relationship depends on the independent relations between the color and the ring and the motion and the ring, and not on the relation between color and motion.

For many ring timings, sampling the stimulus during the ring will yield four features, two colors and two motions. If the pairing reported is the predominant color sampled along with the predominant motion sampled, the data will form a characteristic pattern. For simplicity, we have used a linear decline in performance as one moves away from the optimal coincident ring timing, which yields the square pyramid at bottom of Fig. 8. Comparison of the top and bottom of Fig. 8 reveal that the zipped and independent feature models make very different predictions for the shape of the data, because with independent access to the features, it does not matter

whether the features are temporally coincident in the stimulus. What matters is simply how long each feature is present along with the ring. Clearly, having only one color and one motion present within the ring (corrected for relative latency in processing or accessing the features) will be optimal for report with either binding mechanism. However, when the alignment is less than perfect, performance with independent access will reflect the proportion of time each feature is present irrespective of the timing of the other feature, whereas performance with zipper binding depends on the proportion of time that the two features are paired. For example, if red and inward motion appear simultaneously at the one quarter point of the ring interval and remain on together for 3/4 of the ring duration, performance in the case of direct binding should reflect the 3/4 proportion of the ring duration for each feature. This performance should then be the same when color and motion are not aligned but occupy the same proportion of the cued interval: when red is on at the time the ring appears and ends 3/4 of the way through while inward motion begins 1/4 of the way through the interval and remains on until the ring turns off. However, the two are present simultaneously now for only half the interval. In the independent access model, this temporal registration of the features with each other is not important. On the surface plot (Fig. 8), this predicts a single peak where the ring coincides with both features for the independent model, with responses dropping in each direction from this peak. In contrast, the zipper binding model is that feature pairings during the cued interval are accessed rather than individual features. The prediction is that a particular pair of features will be reported when they are the predominant pairing during the cued interval. The predominant pairing is the same regardless of the timing of the cue, as long as the relative color–motion timing is the same. Since a constant color–motion timing corresponds to a diagonal line on the plot in Fig. 8, the high proportion of a particular response there forms a diagonal ridge (Fig. 8).

The actual data are plotted in Fig. 7, top. The data clearly do not form a ridge, and much more closely resemble the prediction of the independent mechanism. To quantify this we fit both models to the data. First, the zipper model was fit with its simple differential latency cosine ridge (most consistent responses when the two features have a particular temporal alignment) positioned according to the differential latency parameter, with slope determined by the freely varying amplitude parameter. Second, the independent model was fit with freely varying color-ring asynchrony, motion-ring asynchrony, and amplitude parameters. Both models are capped as in previous experiments, with a minimum response proportion of 0 and maximum of 1. For subjects AH, HO, and CH the Pearson product-moment correlations were respectively $r^2 = .87$, .88, and .73 for the best-fitting independent model, compared to only $r^2 = .57$, .38, and .36 for the zipper model. The independent model thus fit much better. Comparing sum of squared errors for it (6494, 6285, and 8289 for the three subjects) to those of the zipper model (20790, 21485, and 20087) using a statistic appropriate for comparing the four-parameter independent model to the three-parameter zipper model (Dobson, 1990, p. 21) yields $F(1, 37)$ values ranging from 53 to 89 for the three subjects, which in each case corresponds to a $p < .000001$.

Does the ring have these effects by attentionally isolating the portion of the stimulus it cues, as if that were all that was presented? No, because unlike in our ring condition, with only a half-cycle stimulus snippet presented, a color–motion asynchrony is still found, as previously reported by Linares and López-Moliner (2006) and as found in a pilot experiment we conducted using the stimulus parameters of this experiment (data not shown). Our ring condition also differs from presentation of only a half-cycle in that a half-cycle presentation introduces temporal order cues including visual persistence and onset cues that do not occur in a longer stream (Beaudot, 2002; Dakin & Bex, 2002; Holcombe, Kanwisher, & Treisman, 2001) even when the longer stream is exogenously cued (Holcombe, Kanwisher, & Treisman, unpublished observation).

4. Experiment 3: Similar effects with a stepping exogenous cue

A stepping exogenous cue (depicted in Fig. 3) may have special effects that do not occur with a singly-presented exogenous cue. For example, Cavanagh et al. (submitted for publication) reported that stepping the cueing ring rather than just presenting it once could substantially raise the alternation rate at which the cued features can be reported. As in their work, in the present experiment the ring steps clockwise or counterclockwise about the stimulus array, visiting each dot patch for the duration of a single color and motion before stepping to the next, adjacent dot patch. We test the effect of the timing of the stepping ring on the pairing reports. As in the previous experiment, on each trial the time that the ring appears – or in this case, the time that it steps – takes one of eight different values relative to the color–motion alternation. As this experiment yields a pattern of results very similar to those of Experiment 2, it adds to the evidence favoring the independent access model.

4.1. Methods

Relative to Experiment 2, the color and motion alternation rate was slightly faster, 376 ms per motion and color interval, yielding a 752 ms cycle time (1.33 Hz). Ten dot fields were arrayed about fixation. In pilot experiments, this change in number of fields made no noticeable difference.

As in Experiment 1, adjacent fields alternated in counterphase (the phase of each dot field was one half-cycle different than its neighbors). The ring spent one half-cycle (376 ms) enclosing each dot field before visiting the next dot field, which due to the counterphasing was then at the same point as the previous dot field was when the ring arrived at it. The effect of this is that the portion of the oscillation highlighted by the ring in the equivalent condition in Experiment 2 is in this experiment cued over and over at each location of the ring.

The same eight values of in → ring interval and red → ring interval were used as in Experiment 2, again with eight trials at each combination. Whether the ring stepped clockwise or counterclockwise about the display was counterbalanced across trials. The ring began stepping one cycle (two steps) before the display began ramping on, to give participants time to begin tracking the ring, and the ring continued stepping through the stimulus train, yielding four steps at full contrast. Chosen randomly on each trial was the starting location of the guide.

Four experienced observers participated. As in the previous experiments, the luminance for the green dots was set for each participant individually to their approximate equiluminance point.

In a separate session, the asynchrony for this particular stimulus without the exogenous cuing by the circle was determined. To do this, the circle flashed very briefly at the location of the target but before the sequence got underway. The circle thus served only as a spatial cue rather than a temporal cue. The observer then judged the predominant feature pairing throughout the sequence of the target dot field, just as in the previous experiments in the literature and as in Experiment 1.

4.2. Results and discussion

When the ring was not presented, as usual the peak red inward responses occurred when the inward motion onset before the red. For the four observers of this experiment, the peak in → red interval and associated bootstrapped standard errors were 100 ± 35, 169 ± 15, 158 ± 8, and 171 ± 13 for AH, HO, CH, and GN, respectively.

For the data with the stepping ring, the independent feature model again gave a much better fit than did the zipper model. For subjects AH, HO, CH, and GN the Pearson product-moment

correlations were, respectively, $r^2 = .91$, .83, .84, and .86 for the best independent feature model, compared to only $r^2 = .47$, .49, .52, and .42 for the zipper model. The independent feature model fit significantly better, as comparing sum of squared errors for it (4983, 6386, 7401, and 5735 for the four participants) to those of the zipper model (27283, 17553, 21616, and 21928) using a statistic appropriate for comparing the four-parameter independent model to the three-parameter zipper model (Dobson, 1990, p. 21) yields $F(1, 37)$ values ranging from 64 to 166 for the four subjects, which in each case corresponds to a $p < .000001$.

As shown in Table 2, the best-fitting parameters of the independent feature model show that the feature reports with stepping ring yielded almost no asynchrony between color and motion.

The stepping ring thus has a very similar effect on the apparent color-motion asynchrony as does the singly-presented ring – in both cases the asynchrony is nearly abolished.

5. Experiment 4: Little effect of rate on best motion–color timing

Nishida and Johnston (2002) reported that the apparent perceptual asynchrony between color and motion diminished from about 100 ms at fast rates to close to 0 ms at 1 Hz for two out of three of their observers. They also found that a stimulus with just a single change, in a sense the slowest possible alternation rate, yielded zero asynchrony between color and motion. These results suggest that the asynchrony may vary with stimulus alternation rate.

However, Nishida and Johnston used a task very different from the feature reports used in the present paper and recent evidence suggests that this can give a very different result. Participants in Nishida and Johnston's study judged which feature changed first, motion or color, and the peak relative timing for synchronous judgments was taken as the motion–color perceptual asynchrony. Subsequent studies have shown that timing judgments in certain situations yield evidence of no asynchrony even when pairing judgments do. For example, Bedell, Chung, and Ogmen (2003) found that for a 1.4 Hz alternation rate, judging which change occurred first yielded no apparent motion–color asynchrony whereas reporting the dominant pairing of color and motion yielded an asynchrony of 150 ms or more. Clifford et al. (2003) documented a similar dissociation for color and orientation pairings. Both results can be explained by the consideration that the temporal order judgment may be made by identifying just the feature that onsets first, rather than identifying both features as is required for reporting the pair. So unlike the order report where asynchrony varies with alternation rate, the report of feature pairs may show a constant asynchrony at all rates. Some weak evidence for this was found by Bedell et al. (2003) and Moutoussis and Zeki (1997), as they tested at approximately 500 and 700 ms cycle durations and found no significant difference in best color–motion relative timing. Most recently, for presentation of only a single cycle, which might be considered an infinitely slow oscillation rate, we (in a pilot experiment mentioned in the discussion of Experiment 2) and Linares and López-Moliner (2006) also found a color–motion asynchrony.

In this experiment, we validate that the color–motion asynchrony seen without an exogenous attention cue is robust across the range of rates used in previous experiments of this paper.

5.1. Methods

Individual dot fields and their colors and motions were identical to those of Experiment 2. The pre- and post-masking and on and off ramping of contrast were also identical. Here, the seven dot fields were presented in phase, all with the target relative color–motion timing appropriate for that trial. Three rates were used: 0.71 Hz (1411 ms), 1.06 Hz (941 ms), and 1.33 Hz (753 ms).

Each of the observers AH, HO, GN, and CH participated in one block at each of the three rates, in counterbalanced order. For the two faster rates, eight equally spaced in → red lags through the possible range were used, with eight trials for each presented in pseudorandom order. For the slowest rate, twelve timings were used to accommodate the greater range of timings, again with eight trials for each.

5.2. Results and discussion

Little if any effect of alternation rate is evident in the data (Fig. 9). The top panel of Fig. 9 shows the data and fitted capped cosine functions, and the bottom panel plots the best color–motion timing as a function of cycle duration. Three out of four subjects showed little effect of alternation rate, and the fourth actually showed an increase of the interval by which motion best precedes the color change, contrary to the reduction in asynchrony at long cycle durations suggested by Nishida and Johnston (2002).

6. General discussion

The explanations proposed for the apparent asynchrony between color and motion span a wide theoretical gamut. The original report and some subsequent work have suggested that the illusion reflected a simple difference in latency between responses in visual cortices selective for color and those selective for motion (Moutoussis & Zeki, 1997; Zeki, 2003). Nishida and Johnston (2002) suggested that the illusion was caused not by simple latency differences but rather by a difference in the processing of the transients caused by motion changes and those caused by color changes. The evidence presented here supports the claim that the intrinsic transients in the original stimulus yield the asynchrony, since we are able to eliminate the asynchrony with the strong transients of the cueing ring. We assume that the ring, an exogenous attentional cue, overrides the unbalanced intrinsic transients and gives equal rapid access to both color and motion. Enns and Oriet (2004) had hypothesized that the asynchrony found between color and motion reflected voluntary switching of attention between dimensions. However, we found that endogenous attention as manipulated by instructions was not sufficient to change the asynchrony much, whereas the exogenous cueing provided by the ring had a large effect.

6.1. Ring transients allow access to motion and color without differential lag

Without the exogenous ring cue transients, the available transients are weak, especially the motion transients (Adams & Mamassian, 2004; Nishida & Johnston, 2002) which can only be detected at slow rates (Werkhoven, Snippe, & Toet, 1992). While the color transient is stronger and access to the new color happens rapidly, access to the new motion value may be delayed after the reversal until the transient elicits attention. The color reversal must be aligned with this delay for optimal response, leading to the usual asynchrony. In contrast, zipper binding gives optimal responses when the stimulus asynchrony compensates for any differential sensory delays for the features. Whatever asynchrony is seen without the ring must also be seen with the ring.

When pairing alternating motions and colors, the differential strength of the color and motion transients may drive observers to sample the color sooner after its onset than the motion. Experiments 2 and 3 show that sampling of the features based on an external transient (the ring) can eliminate the asynchrony by providing equal access to both feature. This would not be predicted by zipper binding theories, as they assume the best relative feature timing is set at an initial binding stage whose properties are independent of how attention later accesses the features.

Although the evidence here shows that external transients can reduce or eliminate the asynchrony between the color and motion, color-contingent motion aftereffects (Arnold, Clifford, & Wenderoth, 2001) demonstrate color motion asynchronies in the absence of any report or transient-triggered access in this task. The aftereffect may be proportional to the cross-correlation of time-varying response profiles to the two features (Clifford et al., 2003). If the motion response grows more slowly or decays more slowly than the color response, to maximize the correlation it would be best to begin the color after the motion. Note that the conscious response to the features may be suprathreshold through much of the interval and thus not be weighted as is the correlation underlying the aftereffect. Much more work is needed to uncover the relationship between the aftereffect and explicit reports.

6.2. Independent feature access by attention

In early visual stages, some features are coded together in cells that respond to multiple dimensions: for example, color, orientation, and direction (Leventhal, Thompson, Liu, Zhou, & Ault, 1995) and size and orientation (Blakemore & Campbell, 1969). These early multi-duty cells can have high temporal resolution and process combinations of features without perceptual awareness supporting a number of combined feature effects (Blaser et al., 2005; Houck & Hoffman, 1986; Humphrey & Goodale, 1998; Melcher, Papathomas, & Vidnyanszky, 2005; Rivest & Cavanagh, 1996; Vul & MacLeod, 2006; Wojciulik & Kanwisher, 1998). However, conscious perception may not have access to these representations any more than it can access the earliest multi-feature activity in retinal cells.

After the initial bound representations, eventually features seem to go their own way, perhaps to their own cortical areas (Zeki, 1978) which are likely to be more relevant than earlier stages for conscious perception (Cohen & Newsome, 2004). At these stages, how are features such as motion and color paired? Feature pairings could occur via a zipper mechanism that ties together the moment-by-moment representations of the motions and color. This might be a preattentive binding mechanism intended to maintain object-level representations, or potentially it could even be triggered by attention. Either way, the pairing reported should be the predominant feature pairing (shifted by any different latencies for color and motion) present in the stimulus or around the time of the ring.

The independent features alternative is that features are sampled when cued for report (either by stimulus transients or an external transient), and it is these features, those predominantly cued by the ring when it is present, that are reported. Our data (Fig. 7) strongly favor independent access.

Despite the independence of motion and color found here, some features are clearly recoded into higher-level descriptions, such as when a curve within a face outline becomes a smile or frown and is no longer accessible as the low-level curvature feature (Suzuki & Cavanagh, 1995). The lack of independent access to the curve in that case indicates that the features have been reorganized under different labels and need translation for recovery. Rapid alternation of feature pairs in some cases leads mid-level vision to construct separate surface representations (Clifford, Spehar, & Pearson, 2004b; Holcombe & Cavanagh, 2001). These surfaces persist temporally through the rapid alternation (Holcombe, 2001) and seem independently accessible to attention, allowing the features on a particular surface to be sampled and paired.

Features that are spatially separated and do not appear to be part of the same object are likely to be accessed independently and potentially asynchronously by attention, as were color and motion here. This independence may contribute to co-localization illusions such as the flash-lag effect, where the position of the moving object is perhaps not accessed until after the flash. The independent access theory advocated by this paper would predict that when features are paired not by attention but by lower-level processes with higher temporal resolution, no flash-

lag effect would occur. The global form detectors that detect structure in Glass patterns seem to meet this criterion (Clifford et al., 2004a) and as expected, its inputs are not affected by the flash-lag effect (Linares & Lopez-Moliner, 2007).

Where low-level mechanisms are not available as in our experiments, our data suggest that the visual system may rely on a minimal but general binding mechanism where two features are experienced as paired because attention samples both those features when selecting a particular location.

Acknowledgments

Appendix A. Alternative data plot

See Fig. A1.
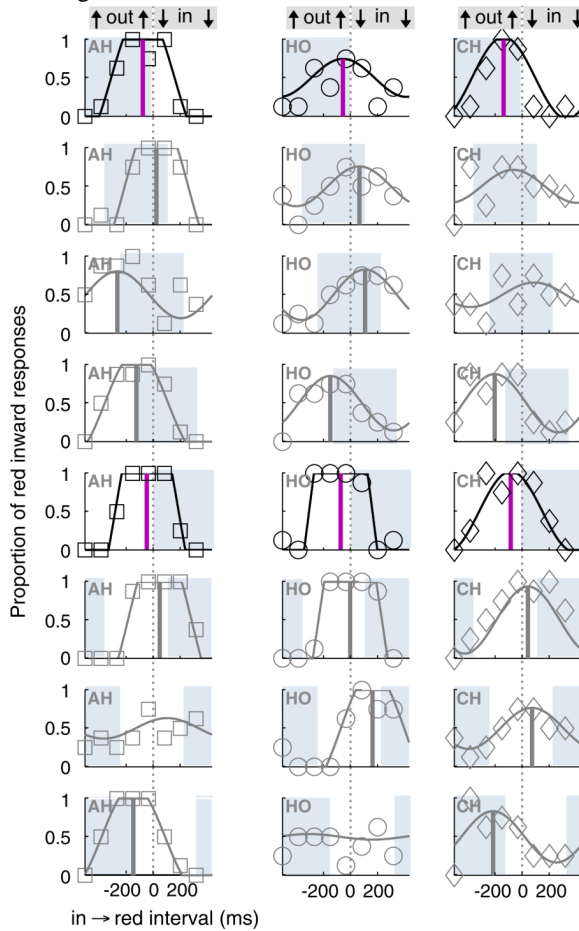
Fig. A1. Alternative to Fig. 7 for plotting the data of E2. When the ring cued the interval of outward motion (top row of subplots, for 3 observers), perceptual asynchrony (best in → red interval) was close to zero or negative, as shown by vertical lines in psychometric curves. The period relative to inward motion cued by the ring is depicted by the light shading. Descending down the subplots shows the data when the ring is presented with progressively more inward motion until the fifth (bolded) row where it contained entirely inward motion. The best time

to change the color shifts to follow the time of the ring. This type of analysis, using the technique of previous literature, does not depict well the best time of the ring relative to the color or motion. One issue is that the analysis treats each ring timing as an independent experiment, a second issue is the changes in amplitude of the curves. A more appropriate analysis is depicted in Fig. 7.

References

Adams WJ, Mamassian P. The effects of task and saliency on latencies for colour and motion processing. Proceedings of the Biological sciences/The Royal Society 2004;271(1535):139–146.

Arnold D. Perceptual pairing of colour and motion. Vision Research 2005;45(24):3015–3026. [PubMed: 16102796]

Arnold DH, Clifford CWG, Wenderoth P. Asynchronous processing in vision: Color leads motion. Current Biology 2001;11:596–600. [PubMed: 11369204]

Beaudot WH. Role of onset asynchrony in contour integration. Vision Research 2002;42(1):1–9. [PubMed: 11804626]

Bedell HC, Chung ST, Ogmen H, Patel SS. Color and motion: Which is the tortoise and which is the hare. Vision Research 2003;43:2403–2412. [PubMed: 12972391]

Blakemore C, Campbell FW. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. Journal of Physiology 1969;203:237–260. [PubMed: 5821879]

Blaser E, Papathomas T, Vidnyanszky Z. Binding of motion and colour is early and automatic. The European Journal of Neuroscience 2005;21(7):2040–2044. [PubMed: 15869499]

Burr DC, Ross J. Contrast sensitivity at high velocities. Vision Research 1982;22:479–484. [PubMed: 7112947]

Cavanagh, P.; Holcombe, AO. Distinguishing pre-selection from post-selection processing limits using a moving window of selection; Journal of Vision. 2005. p. 638a[Abstract] http://journalofvision.org/5/8/638/, doi:10.1167/5.8.638

Cavanagh, P.; Holcombe, AO.; Chou, W-L. Mobile computation: Perception on the fly. (submitted for publication)

Cheal M, Lyon D. Central and peripheral precuing of forced-choice discrimination. Quarterly Journal of Experimental Psychology: Human Experimental Psychology 1991;43A:859–880.

Clifford CWG, Arnold DH, Pearson J. A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. Vision Research 2003;43:2245–2253. [PubMed: 12885378]

Clifford, CW.; Holcombe, AO.; Pearson, J. Rapid global form binding with loss of associated colors.; Journal of Vision. 2004. p. 1090-1101.http://journalofvision.org/4/12/8/, doi:10.1167/4.12.8

Clifford CWG, Spehar B, Pearson J. Motion transparency promotes synchronous perceptual binding. Vision Research 2004b;44:3073–3080. [PubMed: 15474580]

Cohen MR, Newsome WT. What electrical microstimulation has revealed about the neural basis of cognition. Current Opinion in Neurobiology 2004;14(2):169–177. [PubMed: 15082321]

Comtois, R. VisionShell PPC [Software libraries]. Author; Cambridge, MA: 2003.

Croner LJ, Albright TD. Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. The Journal of Neuroscience 1999;19(10):3935–3951. [PubMed: 10234024]

Dakin SC, Bex PJ. Role of synchrony in contour binding: some transient doubts sustained. Journal of the Optical Society of America A-optics Image Science and Vision 2002;19(4):678–686.

Dobkins, KR.; Albright, TD. The influence of chromatic information on visual motion processing in the primate visual system.. In: Watanbe, T., editor. High-level motion processing: Computational, neurobiological and psychophysical perspectives. MIT Press; Cambridge, MA: 1998.

Dobson, AJ. An introduction to generalized linear models. Chapman & Hall; London: 1990.

Duncan J, Ward R, Shapiro K. Direct measurement of attentional dwell time in human vision. Nature 1994;369:313–316. [PubMed: 8183369]26 May

Efron, B.; Tibshirani, RJ. An introduction to the bootstrap. Chapman and Hall; Boca Raton: 1993.

Enns, JT.; Oriet, C. Perceptual asynchrony: Modularity of consciousness or object updating?; Journal of Vision. 2004. p. 27a[Abstract]http://journalofvision.org/4/8/27/, doi:10.1167/4.8.27

Friedman-Hill SR, Robertson LC, Treisman A. Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. Science 1995;269:853–856. [PubMed: 7638604]11 August

Holcombe AO. A purely temporal transparency mechanism in the visual system. Perception 2001;30(11): 1311–1320. [PubMed: 11768486]

Holcombe AO, Cavanagh P. Early binding of feature pairs for visual perception. Nature Neuroscience 2001;4(2):127–128.

Holcombe AO, Kanwisher N, Treisman A. The midstream order deficit. Perception and Psychophysics 2001;63(2):322–339. [PubMed: 11281106]

Houck MR, Hoffman JE. Conjunction of color and form without attention: Evidence from an orientation-contingent color aftereffect. Journal of Experimental Psychology: Human Perception and Performance 1986;12(2):186–199. [PubMed: 2940323]

Humphrey GK, Goodale MA. Probing unconscious visual processing with the McCollough effect. Consciousness and Cognition 1998;7:494–517. [PubMed: 9787058]

Leventhal AG, Thompson KG, Liu D, Zhou Y, Ault SJ. Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. The Journal of Neuroscience 1995;15 (3 Pt 1):1808–1818. [PubMed: 7891136]

Linares, D.; López-Moliner, J. Perceptual asynchrony between color and motion with a single direction change.; Journal of Vision. 2006. p. 974-981.http://journalofvision.org/6/9/10/, doi:10.1167/6.9.10

Linares D, Lopez-Moliner J. Absence of flash-lag when judging global shape from local positions. Vision Research 2007;47(3):357–362. [PubMed: 17207831]

Melcher D, Papathomas TV, Vidnyanszky Z. Implicit attentional selection of bound visual features. Neuron 2005;46(5):723–729. [PubMed: 15924859]

Moradi F, Shimojo S. Perceptual-binding and persistent surface segregation. Vision Research 2004;44 (25):2885–2899. [PubMed: 15380994]

Morgan MJ, Castet E. Stereoscopic depth perception at high velocities. Nature 1995;378:380–383. [PubMed: 7477373]

Moutoussis K, Zeki S. A direct demonstration of perceptual asynchrony in vision. Proceedings of the Royal Society of London B 1997;264:393–399.

Nakayama K, Mackeben M. Sustained and transient components of focal visual attention. Vision Research 1989;29(11):1631–1647. [PubMed: 2635486]

Nishida S, Johnston A. Marker location not processing latency determines temporal binding of visual attributes. Current Biology 2002;12(5):359–368. [PubMed: 11882286]

Posner MI. Orienting of attention. Quarterly Journal of Experimental Psychology 1980;32(1):3–25. [PubMed: 7367577]

Rivest J, Cavanagh P. Localizing contours defined by more than one attribute. Vision Research 1996;36 (1):53–66. [PubMed: 8746242]

Serences JT, Yantis S. Spatially selective representation of voluntary and stimulus-driven attentional priority in human occipital, parietal, and frontal cortex. Cerebral Cortex 2007;17:284–293. [PubMed: 16514108]

Suzuki S, Cavanagh P. Facial organization blocks access to low-level features: An object inferiority effect. Journal of Experimental Psychology: Human Perception and Performance 1995;21:901–913.

Treisman A, Gelade G. A feature integration theory of attention. Cognitive Psychology 1980;12:97–136. [PubMed: 7351125]

Vul E, MacLeod DI. Contingent aftereffects distinguish conscious and preconscious color processing. Nature Neuroscience 2006;9(7):873–874.

Wagner G, Boynton RM. Comparison of four methods of heterochromatic photometry. Journal of the Optical Society of America 1972;62:1508–1515. [PubMed: 4643012]

Werkhoven P, Snippe HP, Toet A. Visual processing of optic acceleration. Vision Research 1992;32(12): 2313–2329. [PubMed: 1288008]

Wojciulik E, Kanwisher N. Implicit but not explicit feature binding in a Balint's patient. Visual Cognition 1998;5(12):157–181.

Zeki SM. Functional specialization in the visual cortex of the rhesus monkey. Nature 1978;274:423–428. [PubMed: 97565]

Zeki S. The disunity of consciousness. Trends in Cognitive Sciences 2003;7(5):214–218. [PubMed: 12757823]
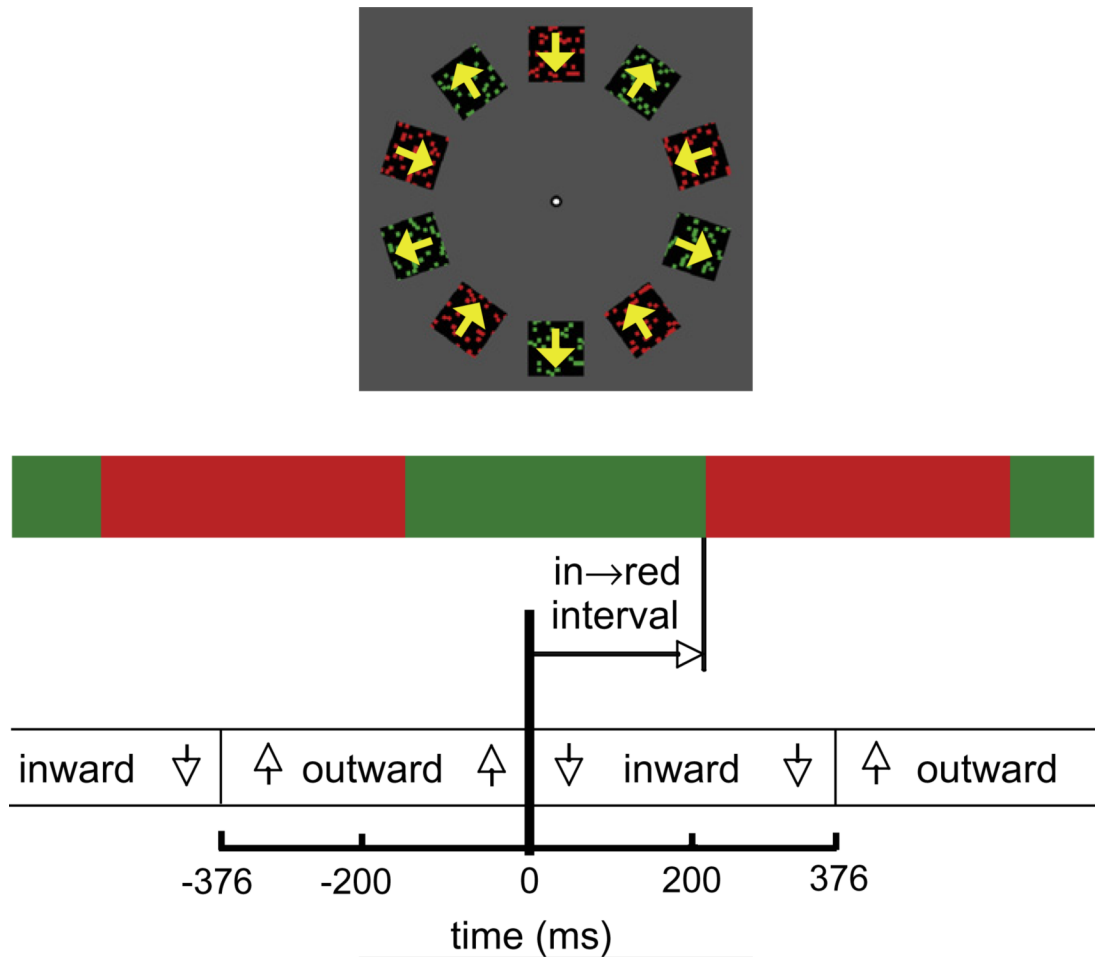
Fig. 1.
In all experiments, dot fields alternate in motion direction between inward and outward, and in color between red and green. In Experiments 1 and 4, adjacent dot fields are one-half cycle out of phase as shown at top. Bottom, relative timing between the motion and color change (in → red interval) is varied to determine the best timing to produce the optimal pairing of color and motion. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
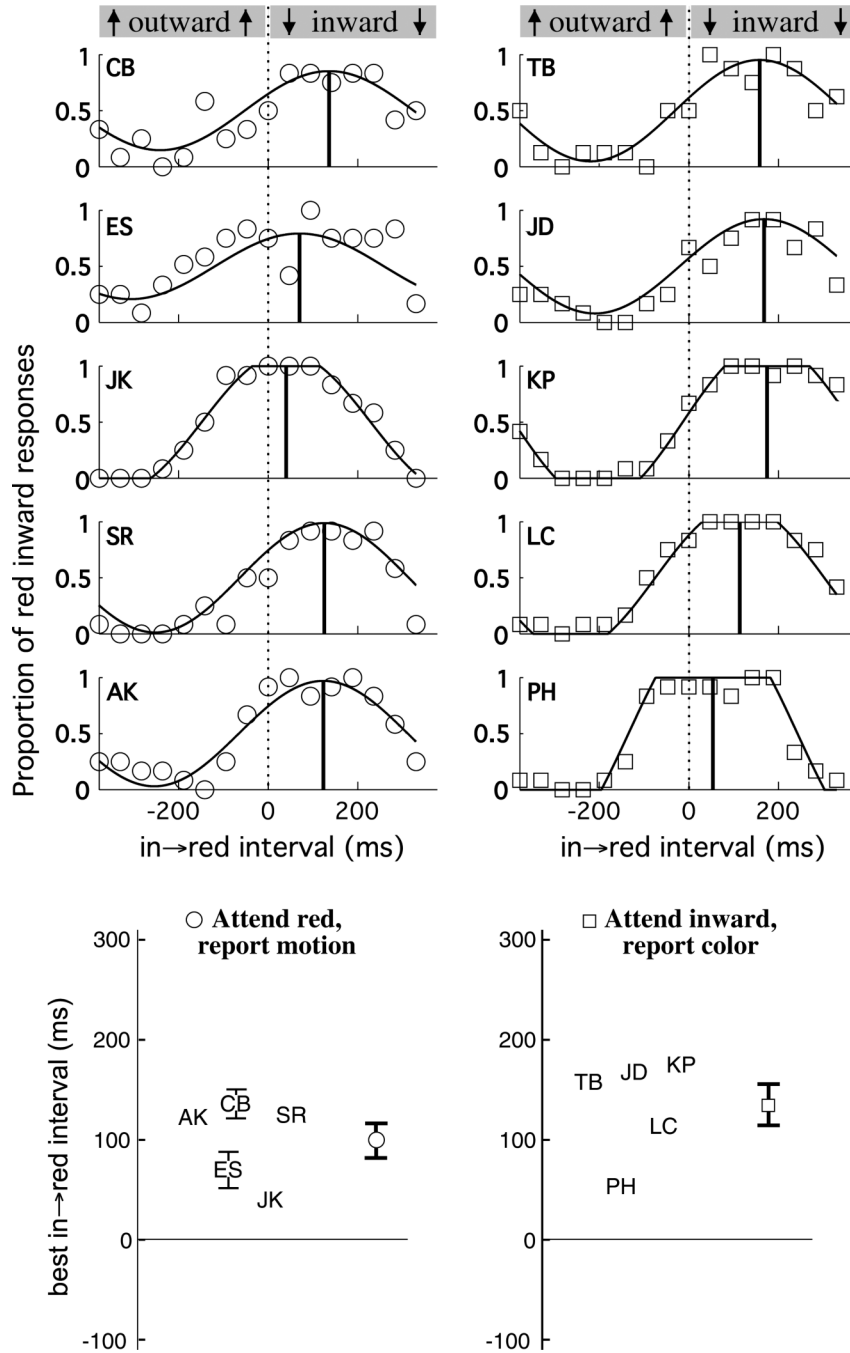
Fig. 2.
Effect of endogenous attentional strategy (Experiment 1). Five observers whose data is plotted on the left show similar perceptual asynchronies to those on the right who are given the opposite attentional instructions. Individual psychometric functions at top are summarized at bottom. ±1 SE shown by bars (where larger than the size of the data symbols), estimated via bootstrapping for this and all subsequent plots. In bottom plots, the bold symbols show the average of the five observers in each condition. The effect of manipulating the attentional strategy is 35 ms, not nearly enough to eliminate or reverse the perceptual asynchrony.
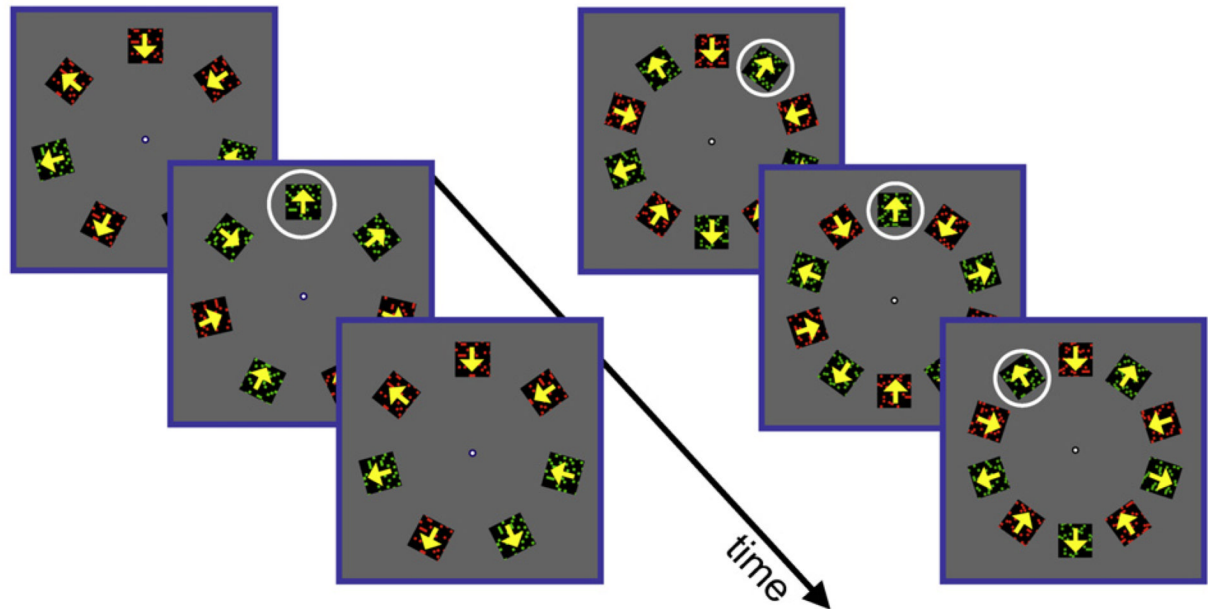
Fig. 3.
Experiment 2 uses a cueing ring presented for one-half cycle. Observers report the predominant feature pairing perceived during the cued interval. In Experiment 3, the same task is performed but the cueing ring steps about the circular array, at each step highlighting the same half-cycle of the alternation. Results from the two experiments are very similar.
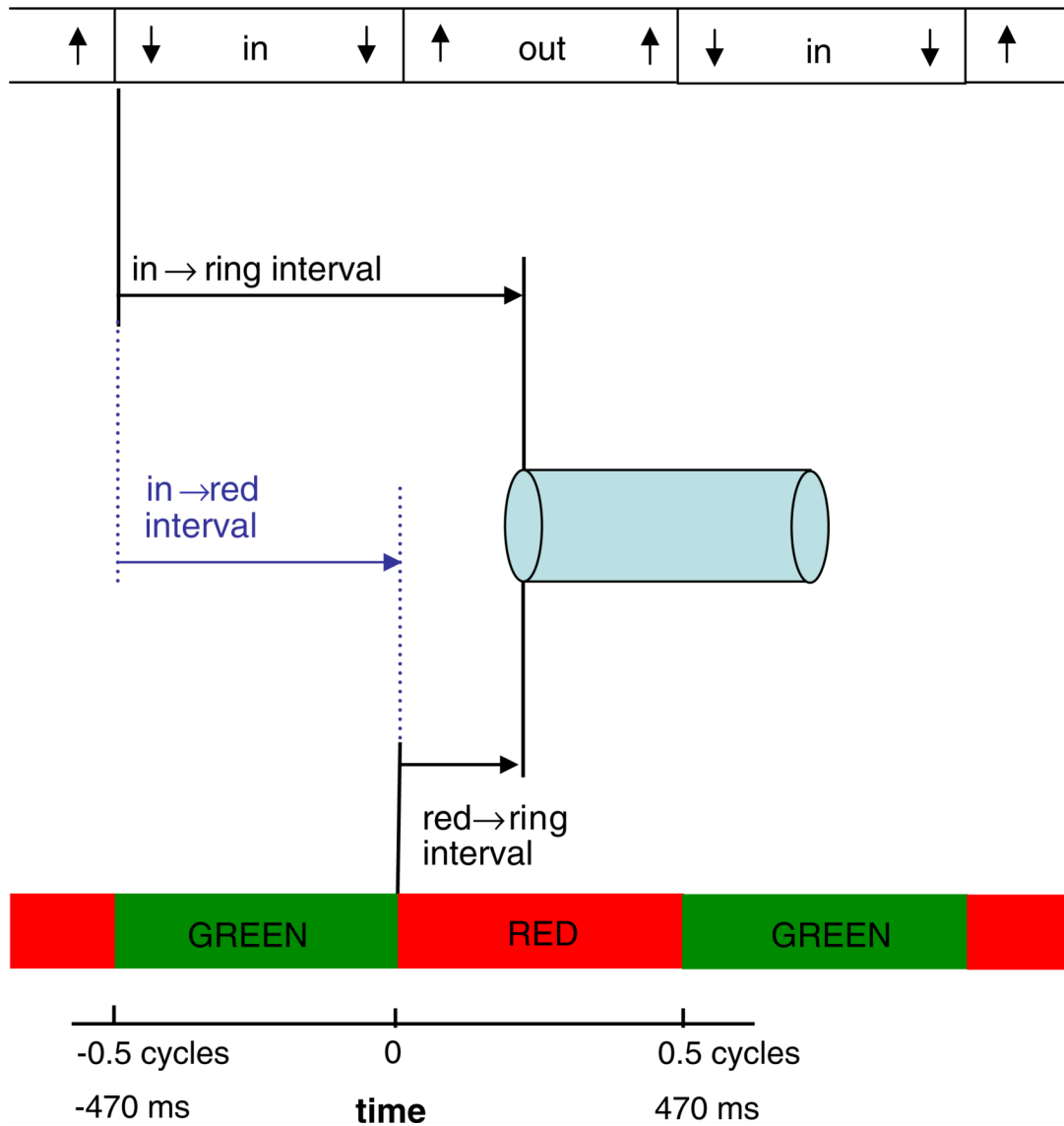
Fig. 4.
In Experiments 2 and 3, varying the timing of the ring relative to the color (red → ring interval) and the motion (in → ring interval) allows determination of the roles of the timing of color relative to motion (in → red interval), versus the timing of the attentional cue relative to each. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
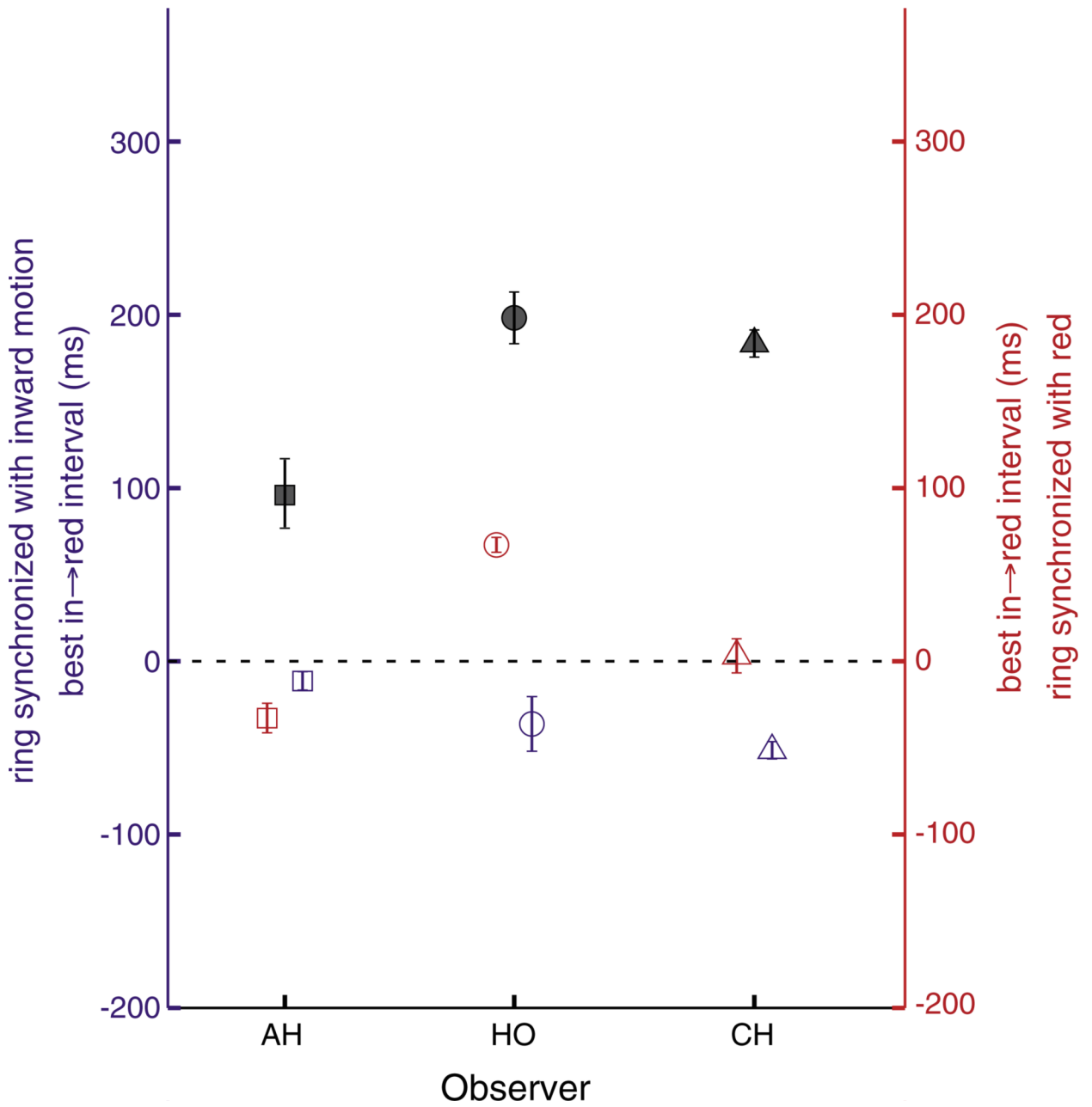
Fig. 5.
From Experiment 2, results based on a subset of the conditions – those when the cueing ring was synchronized with inward motion (purple symbols) and those when it was synchronized with the red (red symbols). The best relative timing between the color and motion is determined by fitting a capped cosine to the data, and extracting the timing of the peak. For each observer, the resulting best in → red interval is close to zero, whereas without the ring (black symbols, data from comparable stimulus of Experiment 4) the best in → red interval is greater than 100 ms. Psychometric functions for all the ring timings are displayed in Appendix A. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
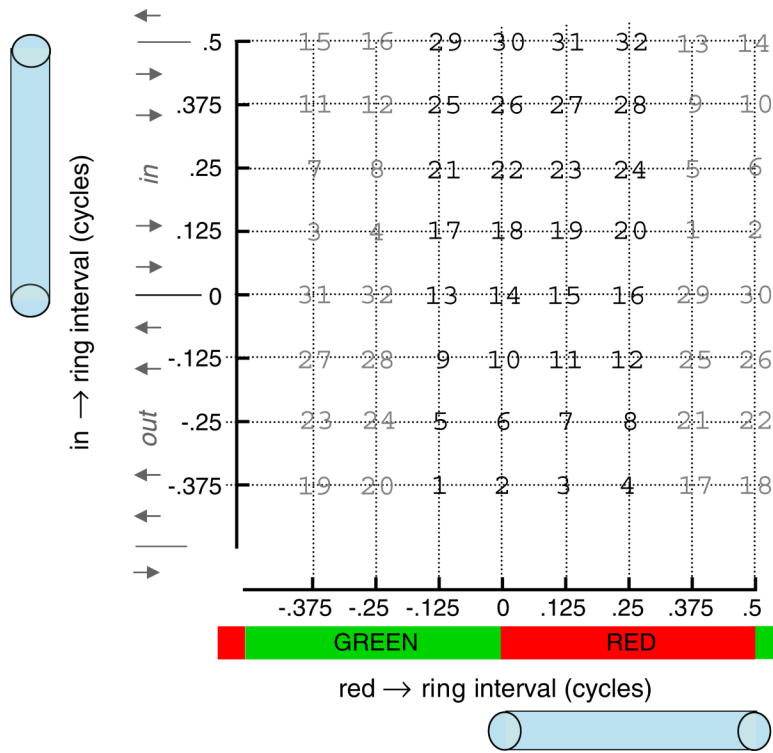
Fig. 6.
Design of E2 and E3, to illustrate how the data were collapsed. Eight red → ring and eight in → ring intervals were used, yielding the 64 conditions lying at the intersections of the dotted grid above. Binding theories are indifferent to whether the ring encircles green or encircles red, if the timing of the ring relative to a color change is the same and if the relative timing of the motion is the same. In other words, red-in and green-out stimuli are equivalent for our purposes. The data can then be simplified by averaging together conditions labeled by the same number, leaving only the 32 boldface conditions, which will be plotted in the figures below. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
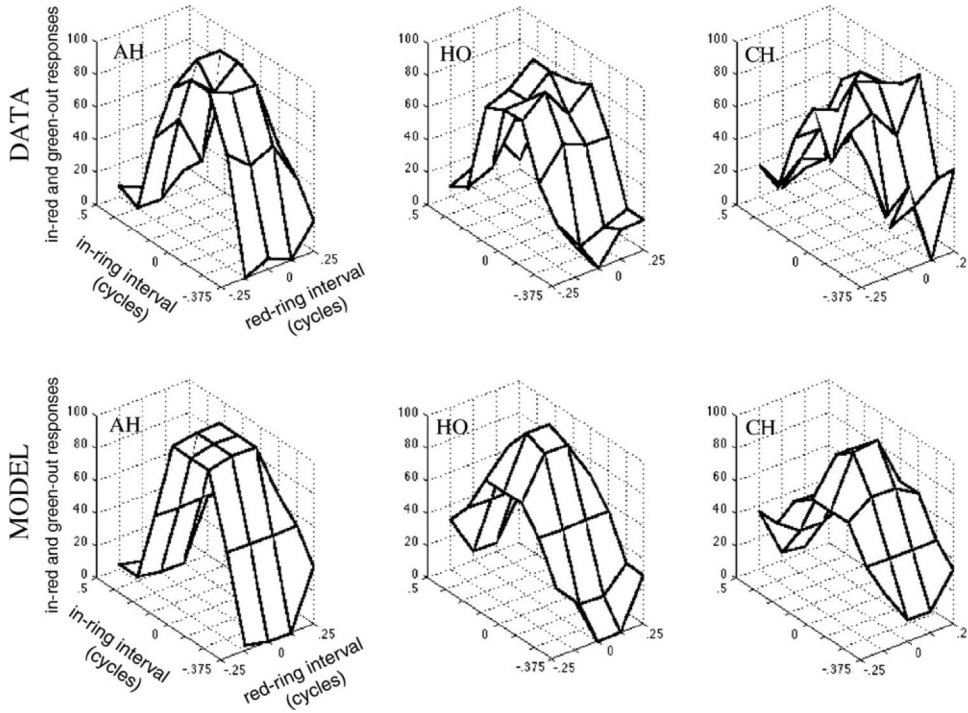
Fig. 7.
Top row is data (collapsed as shown in Fig. 6) from Experiment 2 showing the percent in →
red or green-out responses as a function of the ring timing relative to the color and motion, for
three observers. The data form a hill shape, with the position of the peak close to the origin,
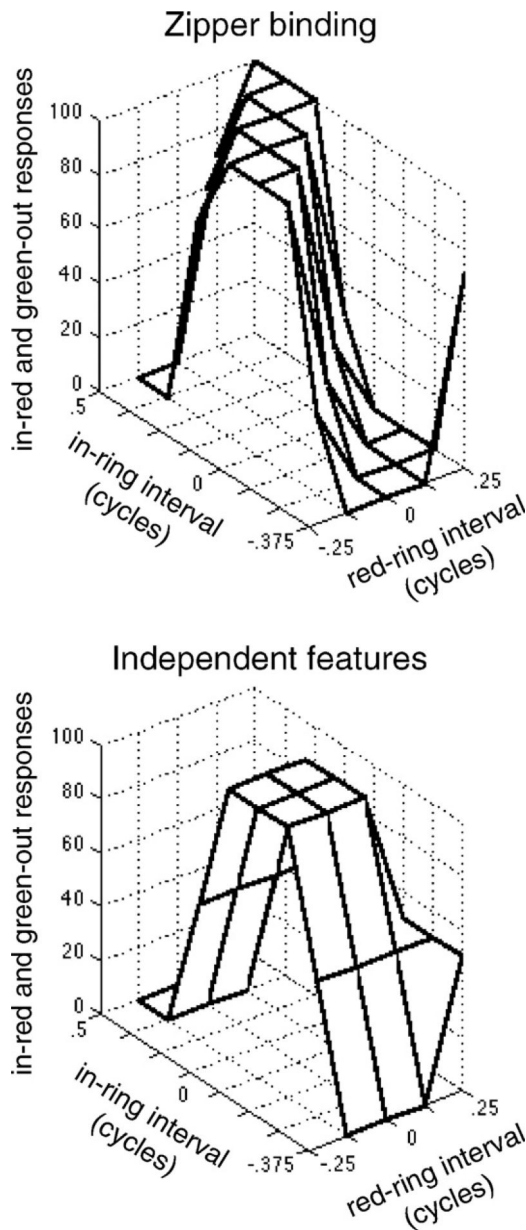indicating little to no asynchrony between ring, color, and motion (Table 1).

Fig. 8.
(Top) The prediction of the zipper binding model is that the data will form a diagonal ridge. Position of the ridge is determined by the relative latency of the color and motion (the pictured position corresponds to a motion latency of 150 ms). (Bottom) The independent features model predicts a single peak corresponding to when the ring cued a single feature on each dimension. As the ring diverges from this point and increasingly cues multiple features on each dimension, performance declines. Center of the peak indicates optimal red → ring interval and in → ring interval, pictured with both assumed to be 0.
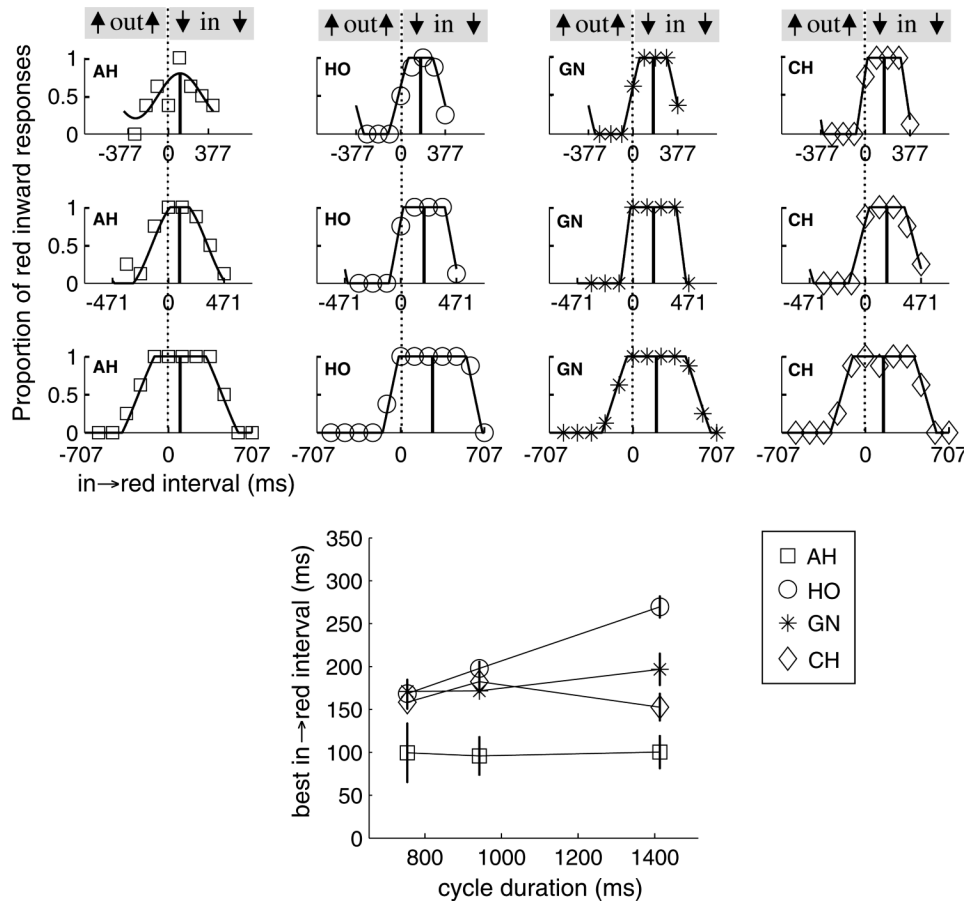
Fig. 9.
Effect of alternation rates on the best in → red interval, for values that span the rates used in the present paper. Each column shows psychometric functions for one observer, at three different rates (rows). At bottom, best relative timing between the onsets of inward motion and red color is plotted as a function of rate.

Table 1

Best-fitting parameters for data of Experiment 2

Participant	Red→ring (ms)	In→ring (ms)	Difference (in→red) (ms)
AH	28 ± 6	0 ± 9	−28 ± 10
HO	66 ± 10	85 ± 12	19 ± 12
CH	75 ± 11	19 ± 13	−56 ± 16

Numbers following ± are one standard error estimated by nonparametric bootstrapping.

Table 2

Fitted parameters for independent feature model for data of Experiment 3

Participant	Red→ring (ms)	In→ring (ms)	Difference (in→red) (ms)
AH	−13 ± 8	−13 ± 8	0 ± 11
HO	−28 ± 15	−28 ± 16	0 ± 22
CH	2 ± 6	46 ± 4	44 ± 7
GN	−15 ± 5	57 ± 4	72 ± 8

Numbers following ± are one standard error estimated by nonparametric bootstrapping.