To appear in Vision Research

Perceptual-binding and persistent surface segregation

Farshad Moradi[1], Shinsuke Shimojo[1,2]

[1]*Computation and Neural Systems Program, California institute of Technology, Pasadena, CA 91125*

[2]*NTT Communication Science laboratories (NTT Com. Sci. Lab.), Atsugi, Kanagawa, 243 Japan.*

Abstract

Visual input is segregated in the brain into subsystems that process different attributes such as motion and color. At the same time, visual information is perceptually segregated into objects and surfaces. Here we demonstrate that perceptual segregation of visual entities based on a transparency cue precedes and affects perceptual binding of attributes. Adding an irrelevant transparency cue paradoxically improved the pairing of color and motion for rapidly alternating surfaces. Subsequent experiments show: (1) Attributes are registered over the temporal window defined by the perceptual persistence of segregation, resulting in asynchrony in binding, and (2) attention is necessary for correct registration of attributes in the presence of ambiguity.

Keywords: color-motion binding; perceptual-asynchrony; attention; figure-ground; motion-transparency

1 Introduction

Neurophysiologists have shown that different aspects of visual stimuli (e.g., color, motion, etc.) are registered in separate cortical subsystems (Felleman & Van Essen, 1991, Livingstone & Hubel, 1988, Milner & Goodale, 1995, Mishkin, Ungerleider & Macko, 1983, Ungerleider & Mishkin, 1982, Zeki, 1978) (but see Lennie, 1998). Localized brain damage in human patients can result in isolated deficits in perception of a single attribute (Damasio, Yamada, Damasio, Corbett & McKee, 1980, Goodale & Milner, 1992, Meadows, 1974, Zeki, 1991, Zihl, von Cramon & Mai, 1983). Recent functional brain imaging studies also support specialization of cortical functions (Hadjikhani, Liu, Dale, Cavanagh & Tootell, 1998, Kanwisher, McDermott & Chun, 1997).

Although visual attributes are processed and represented in functionally distinct brain regions, they compose a coherent percept. Two types of computation are required in order to integrate different aspects of visual input into a unified percept. First, the visual system should be able to partition the visual scene into individuated entities such as surfaces and objects (segregation or parsing, (Spelke, Gutheil & Van de Walle, 1995)). Second, it should determine which combination of attributes is associated with each entity (feature-binding). Spatial or temporal coincidences of attributes are significant only if they come from the same entity. Even if attributes are superimposed, there is no guarantee that they come from the same single object. If an object casts a shadow on a surface the visual system discounts the shadows in perceiving the lightness of the surface. Thus, the shadow and the surface are perceived at the same location, but are not bound together. Figure-ground segregation can also influence how attributes are bound. For

example, when a figure (e.g., red square) moves on a uniform background, local motion signals at the figure-ground boundary are bound to the object perceived as the figure, and not the background[1] (Ramachandran & Anstis, 1986, Tommasi & Vallortigara, 1999). These observeations clearly support the idea that segregation and binding are related. Nonetheless, it is not known if object or surface segregation is interlinked with binding or if it is only necessary when there are ambiguities in the scene, that is, without considering segregation attributes can bind in different and incompatible ways.

We address this issue by examining whether an explicit surface segregation cue (motion or depth transparency) can facilitate color-motion (or color-disparity) binding. Color and motion are processed in distinct cortical areas (Zeki, 1978). Cells in the area MT of the monkey show high sensitivity for motion direction, whereas their sensitivity for color is low. In contrast, few cells in area V4 are direction selective (Cowey, 1994, Felleman & Van Essen, 1987). Consider a display containing red dots moving left, and green dots moving right (either simultaneously or at different times). Observers are asked to report the direction of the red dots. The task itself does not necessitate segregation of red and green dots into distinct surfaces. Thus if observers require an explicit segregation cue (e.g., transparency) to do the task, then we have established that binding and surface segregation are related. This issue is examined in Experiments 1-3.

The dependence does not imply that binding and segregation are the same process. Presumably there are certain aspects that dissociate segregation and binding. Experiments 4-6 address whether surface segregation based on transparency precedes binding (i.e., the scene is first segregated, then attributes are assigned to each object), or if binding and segregation occur concurrently at the same level of processing. We also examine if feature-binding could occur pre-attentively. There is a rich literature about feature-binding and attention. We would like to know if attentional mechanisms in feature-binding are also involved in surface segregation, and to what extent.

2 General Methods

2.1. Observers & Apparatus

Volunteers with normal or corrected-to-normal vision from the California Institute of Technology participated in the experiments. Participants were naïve to the purpose of the study. Subjective equiluminant green (CIE x = 0.29, y = 0.59) was measured for maximum red intensity (CIE x = 0.62, y = 0.33, Y = 21.83) using a technique based on minimizing the flicker between red and green at 14 Hz (Wagner & Boynton, 1972).

Experiments were programmed using the Psychophysics Toolbox extensions (Brainard, 1997, Pelli, 1997) on a Windows PC. The stimuli were presented on the 17" CRT monitor (Dell Trinitron Ultrascan 1000HS) at 85Hz viewed binocularly from a distance of 54cm. In each trial, the stimulus was displayed for about 1.2 s. The participant was asked to press a key to indicate the direction of movement of the red dots (Experiment 1-2) or the color of the moving dots (Experiment 4-6). Conditions were randomized within

[1] This display is ambiguous: it is also possible that a green surface with a square-shaped hole moves on a red background. Alternatively, red and green may both move. However, in the absence of other cues the dominant percept is red moving and green stationary.

each block. No feedback was given to the observers. The procedures will be described separately in each section.

3 Color-motion pairing and perceptual transparency

3.1 Experiment 1a

We tested whether explicit surface segregation (motion transparency) facilitates binding. Observers were asked to view displays of either two transparent surfaces interleaved with blank periods (Figure 1a), or two alternating single-colored surfaces (Figure 1b). They were required to report the direction in which the surface defined by red dots moved. This task essentially probes the perceptual binding of color and direction of motion (Moutoussis & Zeki, 1997). Particularly, in each trial we measure which direction is perceptually bound to a particular color (red).

Both surfaces appear at the same location, thus location cannot be used to pair motion and color. When two moving surfaces are superimposed, transparency subserves as an explicit cue for segregation (Nakayama, Shimojo & Ramachandran, 1990, Watanabe & Cavanagh, 1993). In contrast, the two surfaces are not perceptually segregated in the alternating condition. We empirically verified that the two surfaces are segregated in the transparent trials (Figure 1a), but not in the alternating trials (Figure 1b).

In theory, surface segregation is not necessary for performing the task. There are no ambiguities in the display that would necessitate surface segregation. Computationally, segregation is less efficient than directly looking for the correspondence between color and direction of motion. Observers might alternatively be able to selectively attend to a particular color or direction to carry out the task. However, if binding requires segregation then we predict that color-motion pairing will be impaired if the segregation between the two surfaces becomes weak.

3.1.1 Stimuli & Procedure

Stimuli are illustrated in Figure 1a,b. Random dot patterns (red/green) were presented inside a 5.7° x 5.7° square area on a black background. Each surface consisted of dots (3.5 arcmin) with 2.5% density moving coherently with velocity of 4.85° /s on the black background. Fourteen naïve observers binocularly viewed the display for 1180 ms from a distance of 54 cm.

The red and green dots were turned on and off either simultaneously (transparent condition) or in opposite phase (alternating condition). Thus, the red and green surfaces comprised different dots, moving in opposite directions independent of the other surface. Participants were asked to press a key to indicate the direction of the red dots (left vs. right) after each trial. Trials were randomized for direction, transparency, and interval (SOA of 60, 120, 240, 360 ms). Each participant performed 20 trials per condition.

3.1.2 Results

Figure 2 depicts pairing performance (proportion of correct response to the total number of trials) as a function of interval and condition. Observers correctly reported the direction of the red surface in most trials (>98%) for transparent surfaces, regardless of the frequency of alternation (Figure 2a). In contrast, subjects performed near chance

(50%) when two surfaces were alternated every 120 ms. The interaction between transparency and interval was highly significant (3-way ANOVA, $F(3,40) = 19.56$, $p < 0.0001$). The pattern of results was consistent across observers (Figure 2b). There was a significant difference between transparent and alternating conditions ($F(1,40) = 81.35$, $p < 0.0001$), and between different intervals ($F(3,40) = 17.2$, $p < 0.0001$).

In the transparent condition both surfaces are displayed at the same time and the relative motion between the two surfaces may provide an additional cue for binding and consequently confound the results. Can the higher performance in the transparent condition be attributed to the presence of relative motion in this condition? To rule out this possibility we introduced relative motion in the alternating condition. Nine naïve observes were tested in a separate experiment in which a fixed gray random dot surface was superimposed over the original alternating stimuli. The average performance for the 120 ms interval was 59.4 ±4.8 % (mean ± S.E.M.), which did not show any significant improvement compared to the original experiment (57.2 ± 5.8 %, $p > 0.38$, one tailed t-test).

3.2 Experiment 1b

Can observers' failure in the case of 120 ms interval be due to asynchronous processing of color and motion (Moutoussis & Zeki, 1997)? If color is processed faster than motion then alternating motion before color (Figure 1f) should compensate the difference in latencies and consequently should improve pairing.

To test this possibility, we varied the phase between color and motion. Four naïve observers participated in this experiment. Stimuli and setup were similar to the alternating condition in Experiment 1, except that color and motion were alternated with different phases (20 trials/observer for each data point). Three intervals (94, 120, and 140 ms) were tested. The order of trials was randomized. Observers were asked to report the direction of the red surface.

3.2.1 Results

The performance was around chance level for all phases and intervals, and did not improve by changing the phase between color and motion (Figure 3). Evidently, impaired pairing in the case of 120 ms interval cannot be explained in terms of different latencies for motion and color processing.

3.3 Discussion

Color-motion pairing was impaired in the alternating condition for the 120 ms interval, whereas it was intact in the transparent condition. For longer alternation intervals, colors and directions appeared separable to some extent, that is, the alternation rates were well within the temporal resolution for single attributes. However, observers still failed to correctly pair them in a significant number of trials.

The results support the claim that if binding and surface segregation are related, then providing an explicit segregation cue should facilitate binding. In the transparent condition the observer perceives two distinct surfaces. When the two surfaces alternate every 120 ms, the observer does not perceive distinct surfaces. More rapid alternations can be accompanied by a sense of transparency and segregation (Holcombe, 2001). For

the 60 ms interval, observers reported that the alternating surfaces were perceived as transparent. Consequently, the performance was high and comparable to the transparent condition.

What causes the impairment of veridical motion color binding? Failure to bind alternating features cannot be attributed to a fixed neural latency difference for color and motion (Moutoussis & Zeki, 1997), as performance was high for the shortest interval (60 ms), never dropped below the chance level, and presenting one attribute in advance did not improve the pairing for the 120 ms interval. The performance increased with longer alternation intervals (Figure 2a), suggesting that observers could utilize temporal cues at lower rates (Nishida & Johnston, 2002). Presumably, observers discern the color and direction that appear at the same time. At higher rates they fail to keep track of the order of the attributes. Nonetheless, in the absence of temporal binding cues in the transparent condition observers performed almost perfectly at all presentation rates. Thus, temporal cues are not necessary for feature-binding, even though they may play a role in the case of slowly alternating stimuli.

It is worth emphasizing the significance of the difference between performance in the transparent and alternating conditions. Previous studies have reported dependencies between pairing and frequency of alternations (Clifford, Arnold & Pearson, 2003, Holcombe & Cavanagh, 2001, Moutoussis & Zeki, 1997, Nishida & Johnston, 2002), and suggested different mechanisms to explain the dependencies. However, none of the suggested models mentioned earlier address the role of transparency or segregation in binding.

3.4 Experiment 2a

The results of Experiment 1 suggest that segregation cues improved the performance associated with feature-binding. There are two possible arguments against this interpretation: First, binding mechanisms for transparent surfaces might be inherently different from those for sequential alternation. This was partly ruled out by the control experiment in 3.1.2. Second, although a fixed latency difference by itself cannot account for the results, in Experiment 1 we cannot rule out a contribution of the different temporal dynamics for color and motion pathways (Clifford et al., 2003). Thus, the different temporal relation between the two surfaces in the transparent and alternating conditions is a possible confounding factor.

To address these issues, we introduced transparency in the alternating condition. The performance in the absence of transparency is low for the 120 ms interval. Does surface transparency - without modifying the temporal structure of the stimuli - improve pairing in the alternating condition? In particular, how does pairing performance improve when transparency serves as a cue for a) direct segregation between red and green surfaces, or b) indirectly (i.e., transparency is irrelevant to the task) as a constraint for segregation.

3.4.1 Stimuli and Procedure

Two horizontally moving random dot patterns (red, green) were presented alternatively, similar to alternative condition in Experiment 1a. Six naïve observers were asked to report the direction of the red dots. In half of the trials (no-cue condition), no transparency cue was included in the stimuli. In the rest (segregation cue condition), five

percent of the dots on each surface were gray (CIE x = 0.29, y = 0.29, Y = 21.93) and remained visible during the whole trial (1.2 s), moving to the same direction with the same velocity as the rest of the dots on each surface. Trials were randomized for direction, transparency, and interval (SOA of 60, 120, 240, and 360 ms).

3.4.2 Results

The stimulus in the no-cue condition was identical to the alternating condition in Experiment 1a. As expected, in the absence of any explicit surface segregation cue the pairing was impaired for the 120 ms and 240 ms intervals (Figure 4). Adding transparency cues significantly improved performance (3-way ANOVA, $F(1,15) = 34.35$, $p < 0.0001$). There was also a significant interaction between the interval and presence of a cue ($F(3,15) = 3.59$, $p < 0.05$). Results confirm that explicit surface segregation by motion transparency can facilitate binding.

3.5 Experiment 2b

Experiment 1 was modified by adding two vertically moving (orthogonal to the target) or stationary surfaces with different colors (blue and gray, Figure 1c,d). One target-distracter pair was alternated with the other target-distracter pair with a variable interval (Figure 1d). The transparency cue is irrelevant for discriminating between red and green surfaces. We postulate that the surface segregation constraints posed by transparency are enough to facilitate feature binding. If the addition of the irrelevant transparent surfaces indeed improves the performance of motion-color binding compared to Experiment 1, our claim would be supported.

3.5.1 Stimuli & Procedure

The stimuli and procedure were identical to Experiment 1 except that distracter surfaces with the same dot-density were added in order to introduce a transparency cue (Figure 1c, d). The two additional surfaces had a color of either blue (CIE x = 0.17, y = 0.11, Y = 19.4) or gray (CIE x = 0.29, y = 0.29, Y = 21.9), moving vertically (up or down, 4.85° /s). In the alternating condition, the red and green surfaces were presented alternatively superimposed on one of the distracter surfaces. In the transparent condition, all four surfaces were simultaneously displayed, followed by a blank screen for the same duration. The sequence was repeated for 1180 ms. Fourteen naïve participants (same as Experiment 1a) were asked to press a key to indicate the direction of the red dots (left vs. right) after the trial. The order of the Experiments 1a and 2b was counter-balanced between these participants.

The coincidence of the orthogonal patterns may arguably provide an additional cue for binding, that is, the target direction can be inferred from the direction of any of the distracters. To avoid this possibility, we randomized the sequence in which distracting colors appeared in each trial for a subset of observers, so that the distracters did not contain information about the target. In another control experiment, we used identical stimuli as before, except the blue and gray surfaces were stationary. In Experiment 1a we showed that adding stationary dots per se has little effect on pairing performance. If we get a different result here then it would show that the specific segregation constraints are consequential for binding. This experiment was conducted on five (4 naïve) participants.

3.5.2 Results

In the transparent condition, the performance was slightly lower compared to Experiment 1, demonstrating the increased difficulty of the task. Overall, observers still performed better in the transparent condition compared to the alternating condition (3-way ANOVA, $F(1,27) = 21.27$, $p < 0.0001$). The performance depended on the interval ($F(3,27) = 4.5$, $p < 0.011$), although this effect is smaller compared to Figure 2a. Although the patterns of results seem different in the transparent and alternating conditions, the interaction between transparency and interval was not significant ($F(3,27) = 1.59$, $p > 0.2$).

Despite the increased complexity of the display observers performed significantly better in the alternating condition for the 120 ms interval compared to Experiment 1 (Figure 5a, $p < 0.0001$, HSD post hoc test). The higher performance due to additional segregation cues corroborates our interpretation of Experiment 1 that segregation cues improve feature binding. Furthermore, the results show that feature-binding is improved by transparency -regardless of its task relevance- as long as it poses grouping and segregation constraints between the alternative choices.

Figure 5c depicts the performance of each subject in Experiment 2 compared to Experiment 1. We analyzed results of both experiments together and found significant interaction between number of surfaces, interval, and transparency (4-way ANOVA, $F(3,37) = 7.34$, $p < 0.001$). That is, adding irrelevant distracter surfaces significantly affected the difference between transparent and alternating conditions. In this case, pairing was improved for the alternating condition, and slightly deteriorated for the transparent condition.

It can be argued that orthogonal motion increases the saliency of the target surface, thereby enhancing performance. In the control experiment, alternating red and green surfaces were displayed with and without irrelevant stationary transparent surfaces. The results are shown in Figure 5b. Irrelevant stationary distracters also resulted in improvement in feature-binding[2]. Observers performed significantly better for the 120 ms interval ($p = 0.0028$, one-tailed t-test). The control experiment eliminates the possibility that orthogonal motion contributed to the better performance of feature binding.

3.6 Discussion

Although observers carried out essentially the same task in Experiments 1 and 2, and the pattern of presentation of red and green surfaces were identical, the correct feature binding for the alternating condition was considerably higher in Experiments 2a and 2b. Transparency between distracter and target surfaces provides a cue for segregation in the alternating condition. Even though this cue is not directly relevant to the task in Experiment 2b and increases the complexity of the stimulus, it facilitates pairing presumably by imposing a constraint for parsing input. Experiment 1 indicates that in the alternating condition (120 ms interval) the target surface cannot be effectively segregated

[2] This result may seem inconsistent with the control experiment in 3.1.2. However in 3.1.2 both red and green surfaces were segregated from the same stationary surface, where as here they are segregated from two different surfaces. Thus, the perceptual interpretation and internal representation of the two conditions are different.

from the subsequent surface which moves in the opposite direction (thus perception of the same surface moving back and forth). In Experiment 2b, the target is tailed by either opposite motion, or orthogonal motion. However, the target cannot be grouped with both (transparency constraint). If the target direction is even briefly grouped with dots that are not moving in the opposite direction then the observer would be able to infer the correct direction. This would nicely account for the result obtained here.

3.7 Experiment 3: pairing color and disparity

In the previous two experiments, we showed that the binding between motion and color is facilitated in the transparent condition compared with alternate presentation of each surface. Is this a general principle in binding or limited to color and motion? To examine the generality of the effect, we examine the binding between depth and color. Unlike direction of motion, many cells in area V4 are selective to disparity (Felleman & Van Essen, 1987, Watanabe, Tanaka, Uka & Fujita, 2002). Hence, binding disparity and color might involve a different mechanism than motion-color binding. We segregated the two surfaces by putting them in different depth planes. Two random-dot stereograms (green vs. red) were superimposed to form perceptually transparent surfaces, or were presented in alternation (Figure 1g).

3.7.1 Stimuli and Procedure

Participants viewed two superimposed random dot stereograms (6.7° x 6.7°, density = 4%, dot size = 2 arcmin, disparity = 0 vs. 20 arcmin) through a mirror stereoscopic system. The distance between each image and the observer's eyes was 46 cm. To help fusion and provide a reference, a static zero-disparity black and white random dot pattern (width = 1.65°) surrounded the stimuli. Observers were asked to report whether the red surface was in front of the background or not. Trials were randomized for depth, transparency, and interval. Three naïve observers and one author (FM) participated in this experiment.

3.7.2 Results and discussion

Observers performed almost perfectly in the case of transparent surfaces. In the alternating condition, the error rate was small when rapid alternation of the planes induced perceptual transparency (except for one subject), but the responses deteriorated significantly for longer intervals (Figure 6). It is worth mentioning that the red and green planes are presented at *spatially* (depth-wise) distinct locations. Yet, observers could not fully exploit the depth cue when the interval between alternations was about 120 ms or more.

In general, the results show a similar pattern to Experiment 1. Pairing between color and disparity is efficient when the pairing is supported by a transparency (segregation) cue. In contrast, pairing is impaired when the same stimuli are presented for the same duration but there is no explicit segregation cue.

4 Attention and conjunctive representation of attributes

Experiments 1-3 demonstrate that binding is easy when an explicit surface segregation cue (transparency) is present. Thus, we established that binding and segregation are

related. In the following experiments we examine whether feature-binding and segregation occur at the same level of visual processing, or they involve different but perhaps serially linked processes.

4.1.1 Binding and attention

Illusory conjunctions can occur in normal observers under increased attentional load (Treisman, 1977). Spatial (Treisman, 1988, Treisman & Gelade, 1980), feature (Saenz, Buracas & Boynton, 2002, Saenz, Buracas & Boynton, 2003, Simons & Chabris, 1999, Treue & Martinez Trujillo, 1999), surface (Mitchell, Stoner, Fallah & Reynolds, 2003, Valdes-Sosa, Cobo & Pinilla, 2000), and object-based attention (Duncan, 1984, Duncan, 1993, Duncan & Nimmo-Smith, 1996, Kahneman & Henik, 1981) have been suggested to mediate binding of high-level feature representations. In these accounts, segregation is imposed by top-down attention. Evidence from patients with parietal damage (Ashbridge, Cowey & Wade, 1999, Friedman-Hill, Robertson & Treisman, 1995, Robertson, Treisman, FriedmanHill & Grabowecky, 1997), transcranial magnetic stimulation (Ashbridge, Walsh & Cowey, 1997, Walsh, Ellison, Ashbridge & Cowey, 1999), and functional brain mapping in normal observers (Corbetta, Shulman, Miezin & Petersen, 1995, Shafritz, Gore & Marois, 2002) indicate the involvement of the parietal lobe in feature binding. Thus, top-down mechanisms mediated by the parietal cortex appear to be important for segregation, feature binding, or both.

4.1.2 Pre-attentive mechanisms in feature- binding

There is evidence suggesting the possibility that a weak form of feature binding can be executed pre-attentively, based on coarse location cues (Cohen & Ivry, 1989, Keele, Cohen, Ivry, Liotti & Yee, 1988), or implicit mechanisms (DeSchepper & Treisman, 1996, Egly, Robertson, Rafal & Grabowecky, 1995). Neurons in inferior temporal cortex respond to complex combinations of features such as shape and color (Tanaka, 1993). These studies indicate a complex representation of visual entities in which combination of multiple attributes is explicitly represented. Such representation is suggested to be pre-attentive and early (with the exception of inferior temporal neurons) before different attributes of visual input diverge into distinct cortical areas.

We suspect that some of the controversy surrounding the role of attention in feature-binding might reflect the difference between binding and segregation. In the following experiments we therefore dissociate segregation and binding. In Experiment 4 we verify that surface segregation is fast, consistent with a computational process that does not require feedback. In Experiment 5, we provide evidence that binding between different features (motion and color) follows surface segregation with considerable temporal delay. In Experiment 6, we demonstrate that registration of feature-conjunctions requires top-down attention. Results of Experiments 4-6 contrast fast surface segregation and delayed attentive feature binding.

4.2 Experiment 4: brief motion display

Two similar random dot patterns are superimposed, and one of them briefly moves. When both surfaces are stationary, observers cannot distinguish between them (Figure 7a), but once one surface starts moving the two surfaces become perceptually separable. As the motion stops, the separation disappears. How fast is the above process?

It is reasonable to assume that segregation and binding require some time, in particular if they involve attention. Thus, we set out to measure the minimum duration of transparent motion that permits segregation of a target surface and registration of attributes pertaining to it.

4.2.1 Stimuli and Procedure

Initially, two stationary random dot patterns (red, green, the same density and color as Experiment 1) were superimposed on a black background. The dots subtended a 5.7° x 5.7° square region. Around 520 ms after the onset of the trial, dots of one of the two surfaces inside a randomly located square region (2.8° x 2.8°) moved briefly (4.85° /s) and then stayed at their final location for another 520 ms (Figure 7a). Four naïve participants were asked to report the color and direction of the motion (4-choices). The duration of motion was varied from 24 ms to 120 ms (including the first and last frames). Location, color, and direction of motion were randomized and were not known beforehand.

4.2.2 Results

Direction was accurately reported (>98%) for all durations. Observers also accurately reported the color of the dots for all durations (Figure 7b, chance level is 50%). The location of the target was not known before motion began. A shift in attention conceivably requires 50 ms or more (Shepherd & Muller, 1989). Yet, all participants performed better than 85% correct for the shortest duration (24 ms), which was a one-pixel displacement. These results are consistent with the view that segregation of the target does not require top-down attention. Performance was slightly better for longer presentation times (60 ms -120 ms), which can be attributed to the increased saliency of the motion for 60 ms compared to 24 ms.

There are at least two possible explanations for these results: either the brain can bind color and motion information pre-attentively (Holcombe & Cavanagh, 2001), or the surface could form a trace in observer's brain (similar to iconic memory). In the later case, the internal representation of the surface lasts longer than the attribute (i.e., motion) that had defined it, which permits subsequent processing without further input.

4.3 Experiment 5: Persistence of motion-defined surface

We examined if color and motion attributes are paired and encoded together for brief motion display used in Experiment 4. It may seem counterintuitive that while a presentation duration as short as 24 ms was sufficient for correctly pairing color and motion in Experiment 4, a 120 ms interval in the alternating condition in Experiment 1 was not. A critical difference in these two experimental conditions is that in the alternating condition in Experiment 1, a new surface was presented at every interval. Thus, the persistence of the same dots on the screen after the motion (or alternatively the presence of the dots before the motion onset) might have contributed to the results in Experiment 4.

We assessed the contribution of preceding and subsequent retinal input by modifying the paradigm as follows: During the motion the patterns turned gray (preserving luminance and dot locations). After the motion, colors were presented again. Participants were asked to report the color and the direction of motion of the moving dots, regardless

of any color that may appear before, or after motion. If binding is early or instantaneous then we predict that color would be perceived as gray. However, if the representation of the target surface persists after motion offset then it would be affected by subsequent colors and that would be reflected in the responses.

4.3.1 Stimuli and procedure

The stimuli were the same as in Experiment 4. Dots of one of the two surfaces inside a square region subtending 2.8° moved for 60 ms. To discourage attending to local features, the location of the square was randomized across trials. Dots turned gray (CIE x = 0.29, y = 0.29, Y = 21.93) either during or after motion for 60 ms. In half of the trials, dots turned back to their original color. In the remaining trials (reversed condition), the colors were swapped (i.e., dots that were originally red become green and vice-versa, Figure 7c). Examples of stimuli are provided in the supplementary material [1].

Four naïve observers were asked to report the color of the moving dots, ignoring the color that appears before or after the motion (3-AFC: red, green, or gray). Observers were explicitly informed that the color may change before or after motion.

In a separate experiment, we varied the duration of motion (36, 60, or 120 ms). Five naïve observers were asked to report the color (2AFC: red, green) and direction (left, right) of the dots that had moved.

4.3.2 Results

Figure 7e depicts the results of the experiment. Observers failed to notice that the moving dots were gray. In fact, gray was mostly reported (in less than one-fifth of trials) when it appeared after motion offset. The phenomenally perceived color often matched the color physically presented following the motion. There was significant interaction between reversal and perceived color (3-way ANOVA, $F(2,36) = 13.67$, $p < 0.0001$), regardless of whether gray appeared during or after motion ($p > 0.3$).

A similar pattern was observed for 36 ms and 120 ms durations. In the majority of the trials, the reported color of the target dots matched the color that appeared after motion offset (i.e., the new color). Observers reported the new color significantly more than predicted by chance ($p < 0.0001$, Figure 7d), although their performance was somewhat worse than in Experiment 4 (Figure 7b). The performance depended on the duration of motion ($F(2,22) = 9.76$, $p < 0.001$), and was lower for 36 ms duration ($p < 0.005$ HSD post hoc test), but was the same for 60 ms and 120 ms motion ($p > 0.9$). Even for the shortest duration, observers still performed significantly better than chance in reporting the color that was applied to the surface after motion offset. The direction of the target was reported correctly in almost all trials, regardless of the reversal of the color.

Interestingly, the probability of reporting the new color was lower when it was not the same as the old color for all durations studied ($F(1,22) = 11.13$, $p < 0.003$), indicating that the color of the surface before motion onset influences its perceived color.

4.3.3 Discussion

These results demonstrate that the perceived color of a briefly moving surface is mostly determined by what is displayed a hundred milliseconds or so later. Positional cues cannot account for such result, because the location of the target was not known a priori,

and moving dots were superimposed on non-target stationary dots. Even if observers could individualize a single dot (i.e., as a single object), they clearly misbound color and motion. Overall, Experiment 5 suggests that color information is integrated over a temporal window which extends about ~100 ms and is biased toward later times relative to motion. When a new surface appears at time t_0, its initially perceived color is the temporally weighted average in the interval $[t_0, t_0+a]$. Although the color information becomes available later in time, it is perceptually attributed to the onset of the surface at t_0. In the absence of segregation based on transparency, the subsequent colors presented on the screen affect the perceived color of each surface. This temporal shifting and smearing may contribute to the reported perceptual asynchrony in other studies (Moutoussis & Zeki, 1997) and the failure of observers to pair correctly in the alternating conditions in Experiment 1. This issue is discussed in 5.2.

4.4 Experiment 6: Conjunctive search for motion-color

Although a pre-attentive account for segregation based on transparency cues can explain the results in Experiment 4, attention may be necessary to access the attributes of the surface. This was examined using a visual search paradigm: Four items (moving surfaces) were presented on the screen. Three items had the same color and participants were instructed to locate the odd one. To examine whether conjunctive search for motion-color is serial or parallel, we varied the duration of the presentation. Experiment 4 demonstrates that observers could report the color even for stimuli as brief as 24 ms. If binding is pre-attentive as proposed by (Holcombe & Cavanagh, 2001), then the search should be easy and parallel, and the performance should not be affected by increasing the number of targets for the short motion display. However, if the task requires serial deployment of attention to the four items, then the observers' performance will be low for the 24 ms motion trials, and will increase if the presentation time is increased.

It has been suggested that the visual system can direct attention to stimuli with common motion (Driver, McLeod & Dienes, 1992, McLeod, Driver, Dienes & Crisp, 1991). One may expect then that performance increases when all patterns move in the same direction compared to trials when each item moves in a different direction (that is, if attention to a particular direction of motion is involved in binding).

We contrast the odd color search with a guided search task to control for the particular stimulus parameters. In a separate experiment (guided search), the target was defined by upward motion among horizontally moving distracters. Observers were asked to report the color of the target. Thus, the direction of motion provides the cue for the location of the target. Since the upward motion pops out, attention needs to be directed to only one location, unlike the conjunctive search task in which at least three locations should be examined to find the odd target. Because the stimulus parameters are the same in both experiments, any difference in the performance should be attributed to the difference in attentional requirements of the two tasks.

4.4.1 Stimuli and procedure

Stimuli and setup were similar to Experiment 4, except that dots moved in four 2.3° x 2.3° square regions in the four quadrants (Figure 8a). In each square, all dots with a specific color moved coherently. The rest of the dots (different color) remained stationary. In one of the quadrants, the color associated with moving dots was different from others,

and five naïve participants were instructed to report that quadrant (4AFC). One of the patterns always had a different color than the other three. In half of the trials, dots moved in the same direction in all four squares. In the remaining trials, each location moved in a different direction.

The two superimposed random dot patterns were displayed for 520 ms before and after motion. The duration of motion was varied from 24 ms to 355 ms. Observers were asked to report the location of the odd pattern.

In the guided search experiment, dots moved vertically inside one square and horizontally inside the others. Participants were asked to report the color of the dots moving vertically. The stimulus parameters were the same as in the search experiment.

4.4.2 Results

In the conjunctive search task, participants performed near chance for the shortest duration. Their performance increased with longer search display durations, consistent with a serial strategy for searching (Figure 8b). Observers correctly reported the target location in about 78% of the trials for the longest duration (355 ms), and only 58% when the motion duration was 120 ms. In the guided search task observers correctly reported the color of the target in more than 92% of the trials for the same motion duration. Given this performance level we estimated that observers should have been able to perform better than 77% in the search task for 120 ms motion if one assumes that binding does not require attention.

Earlier reports suggest that visual search for shape is performed in parallel among stimuli with a common motion (McLeod, Driver & Crisp, 1988, McLeod et al., 1991). However, we did not find any difference between the condition in which all targets moved in the same direction, and the condition in which each item had a different direction. In both conditions, the conjunctive search for color-motion requires considerably longer exposure for four items than one item (Experiment 4), compatible with a serial search strategy. Thus, attention to the common direction of motion does not facilitate binding between color and motion.

A similar argument can be made about grouping. When all dots move in the same direction, they can be grouped together as one surface. If observers could attend to this surface then the odd location would have become immediately apparent, independent of the set size (Nakayama & Silverman, 1986). However, the result indicates that search was serial. Thus the segregation trace observed in Experiments 4-6 is local, presumably before stimuli are grouped into a surface.

5 General Discussion

5.1 Summary

We examined the role of surface segregation, temporal cues, and attention to a particular feature, in perceptual binding. Participants were asked to pair color and direction of motion for two alternating or superimposed surfaces, at the same location. While subjects could hypothetically use temporal cues or attend to a particular feature, they were severely impaired when shown alternating surfaces at a presentation rate of 4.25 Hz (120

ms interval). Visual stimulus synchrony has been suggested to facilitate binding (Usher & Donnelly, 1998). Our results indicate that binding cannot be solely based on temporal coincidence at 4.25 Hz. At lower rates of alternation (~2 Hz or less), temporal cues to some extent can be used to pair alternating attributes.

Pairing improved at a higher rate (8.50 Hz) and when surfaces were presented simultaneously. A similar result was demonstrated in the case of disparity and color. Transparency had little effect on binding when it did not invoke segregation of surfaces with opposite motion direction (Experiment 1a). Nonetheless, pairing improved if transparency cues invoked segregation of the target and distracter surfaces (Experiment 2a,b), indicating that different attributes were consistently redistributed to two or more apparently superimposed layers (Watanabe & Cavanagh, 1993). Thus, the visual system is able to construct reference representations (i.e., of visual entities such as objects and surfaces) that encode combinations of attributes, after which binding becomes possible. Directing attention to one surface facilitates access to its attributes, and reduces interference from attributes that belong to others. Although attention might be necessary for registration of surface attributes (Experiment 6), Experiment 4 suggests that segregation of surfaces and their attributes precedes attentional binding (Nakayama, He & Shimojo, 1995).

Surface representations are presumably invoked in low-level visual areas, before location, color, motion, disparity, etc. diverge into distinct and functionally separate cortical modules. Evidence form electrophysiology suggest that the late response profile of neurons in V1 reflect figure-ground separation (Lee, Mumford, Romero & Lamme, 1998). Thus, dynamic neural assemblies in V1 might be a possible candidate for the online representations of objects and surfaces. These representations are presumably created by top-down feedback from area MT which is involved in motion transparency (Castelo-Branco, Formisano, Backes, Zanella, Neuenschwander, Singer & Goebel, 2002, Grunewald, Bradley & Andersen, 2002, Muckli, Singer, Zanella & Goebel, 2002). Attention is necessary to either construct or access these early representations. Improved pairing in Experiments 1-3 can be explained by assuming that attributes of transparent surface are separated at an early level and integrated separately for each surface. Hence, subsequent attributes that do not belong to the same surface will not interfere with veridical binding.

Feature-binding improved when distracting surfaces were added to the alternating condition. Pairing attributes for alternating surfaces is computationally easier than pairing transparent surfaces, and is definitely easier without distracting surfaces. Yet, observers performed better in the more difficult task than the easier one (Figure 5). Despite increasing the attentional load, the distracters improved subjective segregation between target stimuli, facilitating subsequent binding of color and motion.

It can be argued that in Experiment 1 observers rely on the color and direction of a few dots rather than the whole surface. One could see leftward-moving red dots in some local areas, and rightward-moving green dots in other local areas. Although this argument may explain how the participants performed accurately in the transparent condition, it fails to account for the difference between the alternating and transparent conditions. Local cues are present for the same duration in both conditions. Yet, participants performed

significantly better in the latter case. It is therefore unlikely that observers used local cues (i.e., a few dots) for pairing.

We further examined the relationship between surface segregation and binding by presenting two static superimposed surfaces and moving one of them briefly. Observers could correctly segregate the target surface and discriminate its color. However, we demonstrated that the color does not even need to be present during the motion (Experiment 5). Color and motion are bounded together as long as they are perceived as attributes of the same object. Experiment 5 indicates that feature-binding occurs later in time, after segregation. Experiment 6 also demonstrates that attention is necessary for binding in the displays used. Probably, the persisting low-level representations preserve attributes of interest for later analysis by attentional mechanisms.

5.2 Temporal asynchrony and neural latency

It is suggested that high-level selective representations (such as those in color-selective and motion-selective areas) suffer from different latencies (Arnold & Clifford, 2002, Arnold, Clifford & Wenderoth, 2001, Moutoussis & Zeki, 1997), or different temporal dynamics (Clifford et al., 2003) of independent streams of visual information. Moutoussis and Zeki (1997) reported a 50-100 ms bias where color change is perceived to occur earlier than motion change. They argued that the bias reflects the differences in the latencies of color and motion processing that precedes binding. Presumably, such asynchronies mostly arise after early areas such as V1. Thus, Holcombe and Cavanagh (2001) tried to dissociate early and late binding by examining the temporal asynchrony in binding different aspects of visual stimuli. Using rapidly alternating colored orthogonal gratings, they found that color and orientation can be paired correctly for very short periods. Since their finding is not compatible with a model with late binding and different latencies for color and orientation they conclude that these features are encoded in combination.

Our results offer an alternative view which is not based on the latency difference. Experiment 1b and 5 indicate that in the absence of transparency or other explicit segregation cues, color is integrated over time. In both cases, the visual system fails to prevent the integration of consecutive stimuli. At higher alternation rates (8 Hz), surfaces become segregated based on apparent transparency from motion. At lower alternation rates (2 Hz), binding can be based on the temporal pattern of stimuli. To individualize the color of each surface, integration with preceding and succeeding colors has to be prevented. We propose that the reversals of the direction of motion at low rates reset integration of colors [3], which permits separating colors and pairing them with motion, but necessitates that perception be delayed (Eagleman & Sejnowski, 2000) [4]. Higher

[3] Observers seem to be able to accurately report if changes in color and direction are synchronous or not. Therefore, they have access to the temporal information. We, however, assume that changing color, at least near equiluminance, is not salient and does not reset color integration. Experiment 5 indirectly supports this claim: most observers never or rarely observed the gray color when retrospectively asked. We also note that the reset in integration is not absolute and there is a small residual effect of the preceding colors. This residual effect is reflected in Figure 7d (original color vs. color reversed).

[4] Segregation may or may not reset motion processing. Either way, resetting color integration results in apparent asynchrony between motion and color because the phase-lag between segregation and motion is always fixed.

performance in the case of color-orientation and color- disparity indicates that orientation and disparity to some extend restrict temporal integration of color.

5.2.1 Postdiction vs. Different Neural Latencies

Experiment 5 demonstrates a perceptual lag between individualization of an object, and perceiving its attributes. A related effect has been described earlier for the location of a moving object at the time indicated by flashing a stationary marker (flash-lag illusion) (Eagleman & Sejnowski, 2000, Nijhawan, 1994, Nijhawan, 1997, Sheth, Nijhawan & Shimojo, 2000). The notion of different latencies has been used to explain the flash-lag illusion (Baldo & Klein, 1995, Purushothaman, Patel, Bedell & Ogmen, 1998, Whitney & Murakami, 1998). However, the latency difference cannot reconcile the flash-lag illusion and color-motion asynchrony first reported by Moutoussis & Zeki (1997). The apparent latency for moving objects in flash-lag illusion is shorter than stationary objects (consistent with a processing advantage for the moving stimuli), whereas in color-motion binding, pairing improves if the direction of motion reverses 50 ms -100 ms before changing the colors. That is, the apparent latency for motion is longer than color.

Conversely, the postdictive account is consistent with both phenomena. It has been suggested that in the flash-lag illusion, the perceived location of the moving object at the time of flash (temporal cue) is determined by the position of the moving object in a temporal window after the flash (Eagleman & Sejnowski, 2000). Similarly, Experiment 5 demonstrates that the perceived color attributed to motion is contaminated by subsequent colors in a 50-150 ms time window after motion offset. The surface defined by the motion persists and consequently permits integrating subsequent colors - after motion has stopped. In our view, the lag between segregation and binding and the persistence of the surface after motion can account for the reported asynchrony between color and motion. We do not need to assume different latencies for color and motion to explain asynchrony in binding, which is consistent with the dissociation between judgments of the synchrony of changes in color and motion versus judgments of pairing color and motion (Bedell, Chung, Ogmen & Patel, 2003).

It is worth mentioning that the persisting representation discussed here is different from shape from motion (Shioiri & Cavanagh, 1992). Inferring a subsequent color camouflaged by random dots requires more than shape boundaries: the visual system should be able to individualize dots that moved from those that didn't. The short time-scale and the spatial specificity suggest that this representation is low-level. Indirect evidence for such low-level representation comes from transcranial magnetic stimulation of the occipital cortex: visual input to the stimulated cortex is not perceived. The location of the scotoma is well predicted by V1/V2 retinotopic organization. Interestingly, the scotoma appears as the color that follows the stimulation (Kamitani & Shimojo, 1999). Thus, when occipital activity is disrupted, the new visual input overrides the old representation.

5.3 Conclusion

We have dissociated segregation of visual input and binding, and established that binding is affected by surface segregation. The results reveal that surface segregation precedes binding of color and motion. Combinations of visual attributes can be segregated almost effortlessly based on motion or depth transparency and assigned to

persisting and presumably low-level representations. Such representations might not be readily available to visual awareness, and attention is employed for conscious registration and binding of attributes pertaining to the same representation. Further studies are necessary to examine if other segregation cues affect or facilitate binding.

6 Acknowledgement

This research was supported by NIMH, NSF, and Keck foundation. FM was supported by a fellowship from Caltech. We thank C. Koch, R. Kanai, D. Wu, and A. Abbasian for feed-back on the manuscript.

7 References

Arnold, D.H., & Clifford, C.W. (2002). Determinants of asynchronous processing in vision. *Proc R Soc Lond B Biol Sci, 269* (1491), 579-583.

Arnold, D.H., Clifford, C.W., & Wenderoth, P. (2001). Asynchronous processing in vision: color leads motion. *Curr Biol, 11* (8), 596-600.

Ashbridge, E., Cowey, A., & Wade, D. (1999). Does parietal cortex contribute to feature binding? *Neuropsychologia, 37* (9), 999-1004.

Ashbridge, E., Walsh, V., & Cowey, A. (1997). Temporal aspects of visual search studied by transcranial magnetic stimulation. *Neuropsychologia, 35* (8), 1121-1131.

Baldo, M.V., & Klein, S.A. (1995). Extrapolation or attention shift? *Nature, 378* (6557), 565-566.

Bedell, H.E., Chung, S.T., Ogmen, H., & Patel, S.S. (2003). Color and motion: which is the tortoise and which is the hare? *Vision Res, 43* (23), 2403-2412.

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat Vis, 10* (4), 433-436.

Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., & Goebel, R. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc Natl Acad Sci U S A, 99* (21), 13914-13919.

Clifford, C.W., Arnold, D.H., & Pearson, J. (2003). A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. *Vision Res, 43* (21), 2245-2253.

Cohen, A., & Ivry, R. (1989). Illusory conjunctions inside and outside the focus of attention. *J Exp Psychol Hum Percept Perform, 15* (4), 650-663.

Corbetta, M., Shulman, G.L., Miezin, F.M., & Petersen, S.E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science, 270* (5237), 802-805.

Cowey, A. (1994). Cortical visual areas and the neurobiology of higher visual processes. In: M.J. Farah, & G. Ratcliff (Eds.), *The neuropsychology of high-level vision* (pp. 3-31). Hillsdale, New Jersey: Lawrence Erlbaum associates.

Damasio, A., Yamada, T., Damasio, H., Corbett, J., & McKee, J. (1980). Central achromatopsia: behavioral, anatomic, and physiologic aspects. *Neurology, 30* (10), 1064-1071.

DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: implicit coding without attention. *J Exp Psychol Learn Mem Cogn, 22* (1), 27-47.

Driver, J., McLeod, P., & Dienes, Z. (1992). Motion coherence and conjunction search: implications for guided search theory. *Percept Psychophys, 51* (1), 79-85.

Duncan, J. (1984). Selective attention and the organization of visual information. *J Exp Psychol Gen, 113* (4), 501-517.
Duncan, J. (1993). Coordination of what and where in visual attention. *Perception, 22* (11), 1261-1270.
Duncan, J., & Nimmo-Smith, I. (1996). Objects and attributes in divided attention: surface and boundary systems. *Percept Psychophys, 58* (7), 1076-1084.
Eagleman, D.M., & Sejnowski, T.J. (2000). Motion integration and postdiction in visual awareness. *Science, 287* (5460), 2036-2038.
Egly, R., Robertson, L.C., Rafal, R., & Grabowecky, M. (1995). Implicit processing of unreportable objects in Balint's syndrom. *36th annual meeting of the Psychonimic Society* (Los Angeles.
Felleman, D.J., & Van Essen, D.C. (1987). Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *J Neurophysiol, 57* (4), 889-920.
Felleman, D.J., & Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex, 1* (1), 1-47.
Friedman-Hill, S.R., Robertson, L.C., & Treisman, A. (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science, 269* (5225), 853-855.
Goodale, M.A., & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends Neurosci, 15* (1), 20-25.
Grunewald, A., Bradley, D.C., & Andersen, R.A. (2002). Neural correlates of structure-from-motion perception in macaque V1 and MT. *J Neurosci, 22* (14), 6195-6207.
Hadjikhani, N., Liu, A.K., Dale, A.M., Cavanagh, P., & Tootell, R.B. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Nat Neurosci, 1* (3), 235-241.
Holcombe, A.O. (2001). A purely temporal transparency mechanism in the visual system. *Perception, 30* (11), 1311-1320.
Holcombe, A.O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nat Neurosci, 4* (2), 127-128.
Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In: M. Kubovy, & J. Pomerantz (Eds.), *Perceptual Organization* (pp. 181-211). Hillsdale, NJ: Erlbaum.
Kamitani, Y., & Shimojo, S. (1999). Manifestation of scotomas created by transcranial magnetic stimulation of human visual cortex. *Nat Neurosci, 2* (8), 767-771.
Kanwisher, N., McDermott, J., & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci, 17* (11), 4302-4311.
Keele, S.W., Cohen, A., Ivry, R., Liotti, M., & Yee, P. (1988). Tests of a temporal theory of attentional binding. *J Exp Psychol Hum Percept Perform, 14* (3), 444-452.
Lee, T.S., Mumford, D., Romero, R., & Lamme, V.A. (1998). The role of the primary visual cortex in higher level vision. *Vision Res, 38* (15-16), 2429-2454.
Lennie, P. (1998). Single units and visual cortical organization. *Perception, 27* (8), 889-935.
Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science, 240* (4853), 740-749.
McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature, 332* (6160), 154-155.

McLeod, P., Driver, J., Dienes, Z., & Crisp, J. (1991). Filtering by movement in visual search. *J Exp Psychol Hum Percept Perform, 17* (1), 55-64.
Meadows, J.C. (1974). Disturbed perception of colours associated with localized cerebral lesions. *Brain, 97* (4), 615-632.
Milner, A.D., & Goodale, M.A. (1995). The Visual Brain in Action. (Oxford: Oxford University Press.
Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision - 2 cortical pathways. *Trends in Neurosciences, 6* (10), 414-417.
Mitchell, J.F., Stoner, G.R., Fallah, M., & Reynolds, J.H. (2003). Attentional selection of superimposed surfaces cannot be explained by modulation of the gain of color channels. *Vision Res, 43* (12), 1323-1328.
Moutoussis, K., & Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci, 264* (1380), 393-399.
Muckli, L., Singer, W., Zanella, F.E., & Goebel, R. (2002). Integration of multiple motion vectors over space: an fMRI study of transparent motion perception. *Neuroimage, 16* (4), 843-856.
Nakayama, K., He, Z.J., & Shimojo, S. (1995). Visual surface representation: a critical link between lower-level and higher-level vision. In: S. Kosslyn, & R.A. Andersen (Eds.), *Frontiers in Cognitive Neuroscience* (Cambridge, MA: MIT Press.
Nakayama, K., Shimojo, S., & Ramachandran, V.S. (1990). Transparency: relation to depth, subjective contours, luminance, and neon color spreading. *Perception, 19* (4), 497-513.
Nakayama, K., & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320* (6059), 264-265.
Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370* (6487), 256-257.
Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature, 386* (6620), 66-69.
Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr Biol, 12* (5), 359-368.
Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis, 10* (4), 437-442.
Purushothaman, G., Patel, S.S., Bedell, H.E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396* (6710), 424.
Ramachandran, V.S., & Anstis, S. (1986). Figure-ground segregation modulates apparent motion. *Vision Res, 26* (12), 1969-1975.
Robertson, L., Treisman, A., FriedmanHill, S., & Grabowecky, M. (1997). The interaction of spatial and object pathways: Evidence from Balint's syndrome. *Journal of Cognitive Neuroscience, 9* (3), 295-317.
Saenz, M., Buracas, G.T., & Boynton, G.M. (2002). Global effects of feature-based attention in human visual cortex. *Nat Neurosci, 5* (7), 631-632.
Saenz, M., Buracas, G.T., & Boynton, G.M. (2003). Global feature-based attention for motion and color. *Vision Res, 43* (6), 629-637.
Shafritz, K.M., Gore, J.C., & Marois, R. (2002). The role of the parietal cortex in visual feature binding. *Proc Natl Acad Sci U S A, 99* (16), 10917-10922.
Shepherd, M., & Muller, H.J. (1989). Movement versus focusing of visual attention. *Percept Psychophys, 46* (2), 146-154.

Sheth, B.R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nat Neurosci, 3* (5), 489-495.
Shioiri, S., & Cavanagh, P. (1992). Visual persistence of figures defined by relative motion. *Vision Res, 32* (5), 943-951.
Simons, D.J., & Chabris, C.F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception, 28* (9), 1059-1074.
Spelke, E.S., Gutheil, G., & Van de Walle, G. (1995). The development of object perception. In: S.M. Kosslyn, & D.N. Osherson (Eds.), *Visual Cognition,* 2 (pp. 297-330). Boston, MA: MIT Press.
Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science, 262* (5134), 685-688.
Tommasi, L., & Vallortigara, G. (1999). Figure ground segregation modulates perceived direction of ambiguous moving gratings and plaids. *Vision Res, 39* (4), 777-787.
Treisman, A. (1977). Focused attention in perception and retrieval of multidimensional stimuli. *Percept Psychophys, 22* (1), 1-11.
Treisman, A. (1988). Features and objects: the fourteenth Bartlett memorial lecture. *Q J Exp Psychol A, 40* (2), 201-237.
Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognit Psychol, 12* (1), 97-136.
Treue, S., & Martinez Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399* (6736), 575-579.
Ungerleider, L.G., & Mishkin, M. (1982). The two cortical visual systems. In: D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549-586). Cambridge, Masachussetts: MIT Press.
Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature, 394* (6689), 179-182.
Valdes-Sosa, M., Cobo, A., & Pinilla, T. (2000). Attention to object files defined by transparent motion. *J Exp Psychol Hum Percept Perform, 26* (2), 488-505.
Wagner, G., & Boynton, R.M. (1972). Comparison of four methods of heterochromatic photometry. *J Opt Soc Am, 62* (12), 1508-1515.
Walsh, V., Ellison, A., Ashbridge, E., & Cowey, A. (1999). The role of the parietal cortex in visual attention--hemispheric asymmetries and the effects of learning: a magnetic stimulation study. *Neuropsychologia, 37* (2), 245-251.
Watanabe, M., Tanaka, H., Uka, T., & Fujita, I. (2002). Disparity-selective neurons in area V4 of macaque monkeys. *J Neurophysiol, 87* (4), 1960-1973.
Watanabe, T., & Cavanagh, P. (1993). Surface decomposition accompanying the perception of transparency. *Spat Vis, 7* (2), 95-111.
Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci, 1* (8), 656-657.
Zeki, S. (1991). Cerebral akinetopsia (visual motion blindness). A review. *Brain, 114 (Pt 2),* 811-824.
Zeki, S.M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature, 274* (5670), 423-428.
Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain, 106 (Pt 2),* 313-340.

8 Figure Captions

Figure 1. Experimental Paradigms
Color-motion binding for transparent (a) and alternating (b) surfaces in Experiment 1. Observers were asked to report the direction of the red dots. c, d) Experiment 2b: the two additional surfaces (gray, blue) are irrelevant to the task. Participants were asked to report the direction of the red dots. c) All four surfaces were superimposed (transparent) d) two superimposed surfaces were alternated e-g) A representation of color and motion (or disparity) alternations in time in different experiments. e) Experiment 1a. f) Experiment 2a: Temporal asynchrony between color and motion (compare to Moutoussis & Zeki 1997). In this example, an ideal observer should report red is moving right for a phase difference of 0 and green for a phase difference of π. Results are plotted against phase for all observers in Figure 3. g) Binding between disparity and color information (Experiment 3). Two surfaces were presented at zero or 20' disparity planes.

Figure 2. a) Average performance (proportion of correct responses to the total number of trials) plotted as a function of rate (interval between alternations) in Experiment 1. b) Individual plots of nine participants (same axes as in a). Each dot is the average of 20 trials.

Figure 3. Asynchronous presentation of color and motion (Experiment 1b) did not improve pairing for 94 (diamond), 120 (square), or 140 ms (triangle) intervals. The performance of a model observer (based on temporal information) is depicted as an ogival curve.

Figure 4. Results of Experiment 2a. Pairing accuracy is plotted as a function of alternation interval and presence of a transparency cue. Adding a transparency cue (5% of the dots) significantly improved pairing in the alternating condition. * $p < 0.05$.

Figure 5. Pairing is more reliable when an irrelevant segregation cue (orthogonally moving or stationary distracting surfaces) is present (Figure 1c,d). Participants were asked to report the direction of red dots.
a) Increasing attentional load can paradoxically improve feature-binding between color and motion (see 3.6). Four surfaces are displayed instead of two in Experiment 1. Open squares in the graph depict the performance in the 2-surface condition for comparison. Pairing accuracy (proportion of correct responses to the total number of trials) is plotted as a function of interval. b) A similar result was obtained in five observers using stationary surfaces (Experiment 2b control). c) Results of 14 participants (Experiment 2b, alternating condition) are plotted against their performance in Experiment 1a (alternating condition) for the 120 ms and 240 ms alternating conditions.

Figure 6. Pairing accuracy (proportion of correct responses to the total number of trials) is plotted as a function of interval in Experiment 3. a) Binding between disparity and color information (Experiment 3, n = 4). Two surfaces were presented at zero or 20' disparity planes. b) Individual plots of each participant.

Figure 7. Persistence of surface after motion.

a) After 530 ms, all green dots inside a square (invisible boundaries) move left for 26-120 ms (Experiment 4). b) Performance vs. duration of motion for three observers. The performance is high even for the shortest duration. c) Experiment 5: all dots turned gray before motion onset. The dots that were originally green moved left. After the motion, dots either turned back to their original color or switched colors. d) Observers nearly always reported the final color of the dots (considered as the correct response). Dotted line depicts performance in Experiment 4. e) Four naïve observers were asked to report the color of the moving pattern. The moving dots were turned gray for 60 ms during or after motion. The green bar represents the trials in which the original color was reported, the gray represents trials in which observers reported gray.

Figure 8. a) Visual search stimulus (Experiment 6). In one of the quadrants, moving dots have a different color. b) The serial nature of the conjunctive search, as marked by an increase in performance with time, suggests that attention is necessary to register surfaces' color. Open circles (o) depict the performance of subjects in reporting the color of the target defined by upward motion (guided search experiment).

9 Supplementary material

1. Examples of stimuli used in experiment 5. The observer is required to report the color of the moving dots. In the first trial, all dots turn gray during the motion and go back to their original color afterwards. In the second trial, dots turn gray during the motion, but the color switches after the motion offset. In the last trial, dots turn gray only after the motion (control experiment).

Figure 1

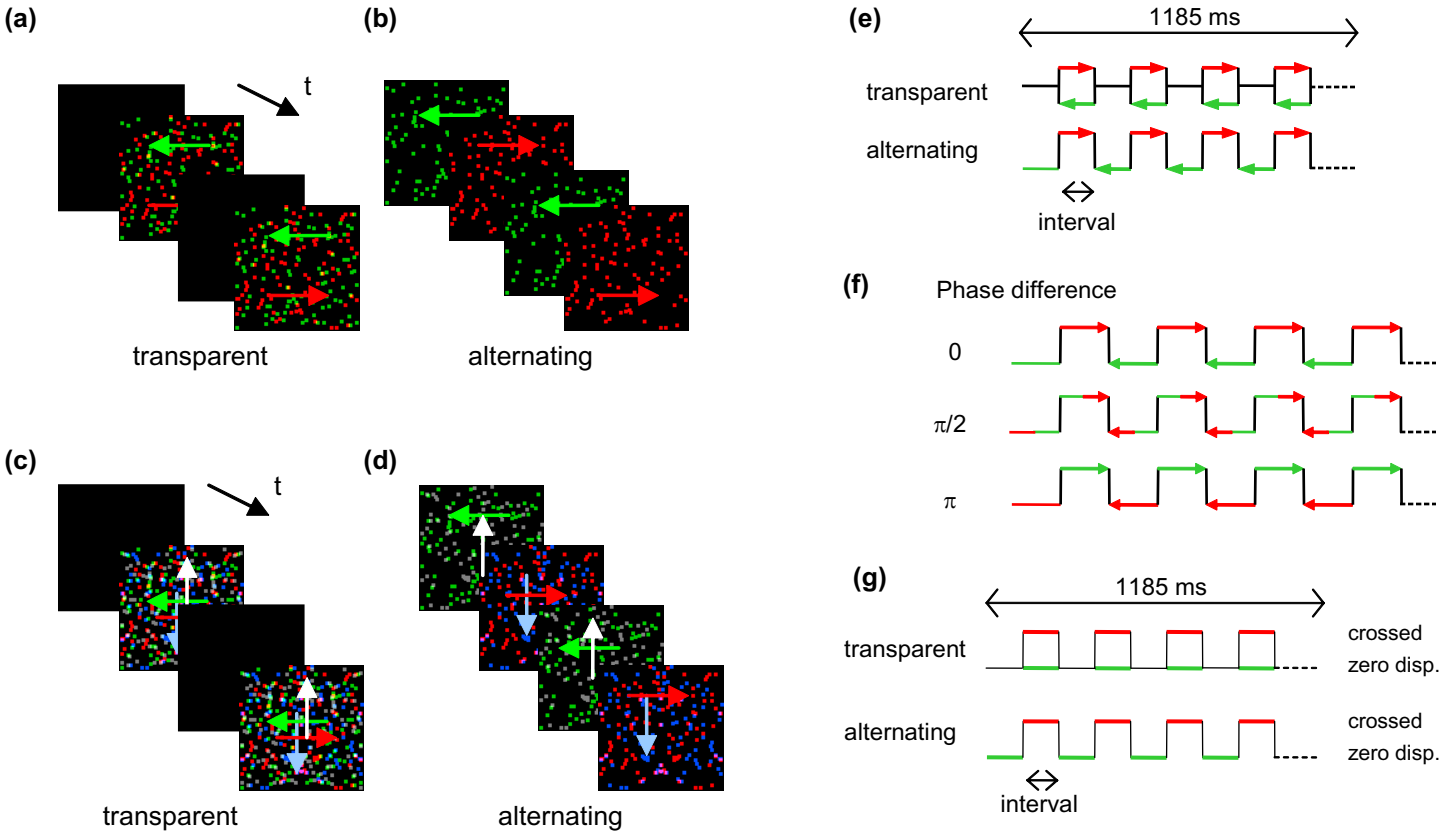

(a)
t
transparent
(b)
alternating
(c)
t
transparent
(d)
alternating
(e)
1185 ms
transparent
alternating
interval
(f)
Phase difference
0
π/2
π
(g)
1185 ms
transparent
crossed
zero disp.
alternating
crossed
zero disp.
interval

Figure 2

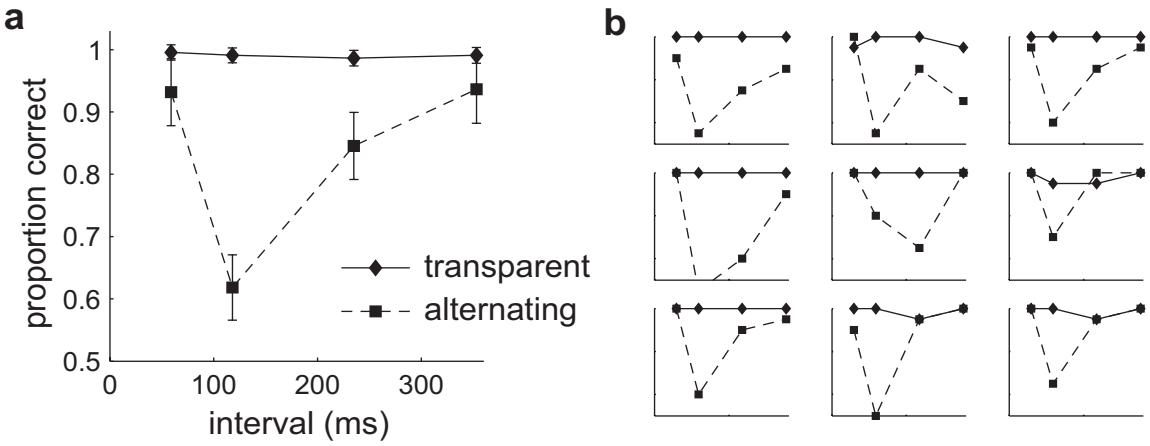


Figure 3

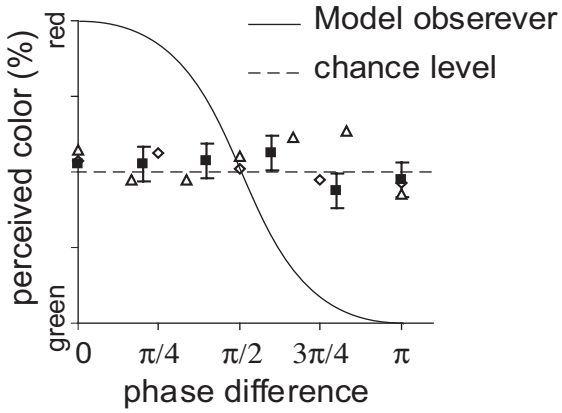


Figure 4

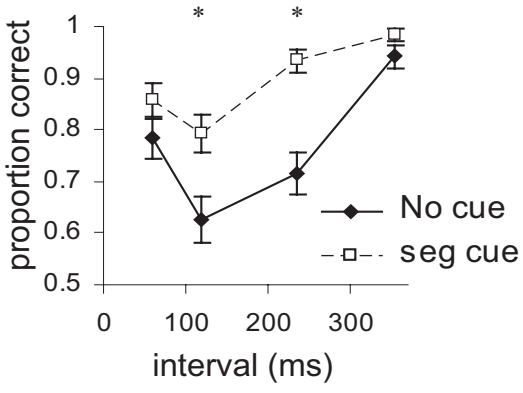


Figure 5

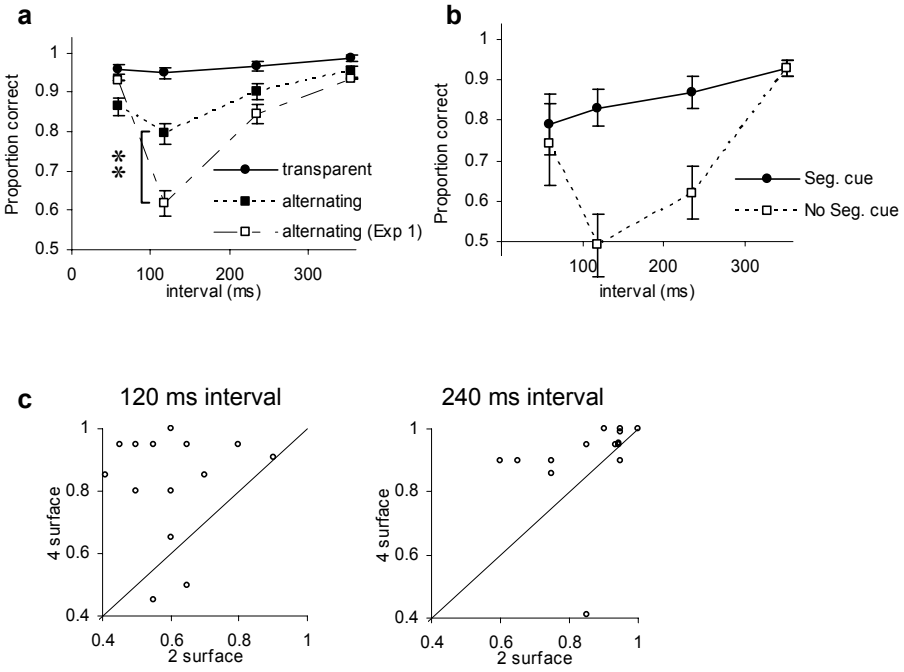


Figure 6

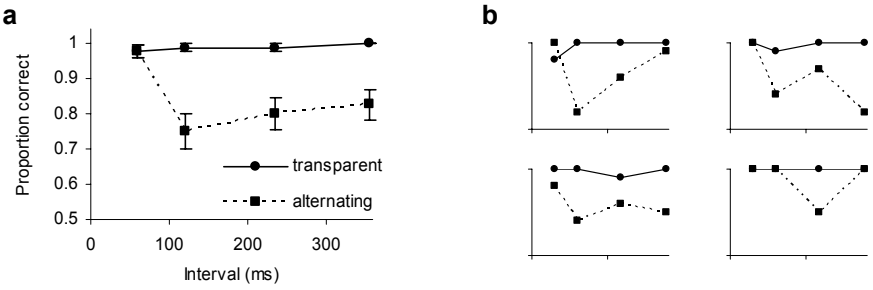


Figure 7

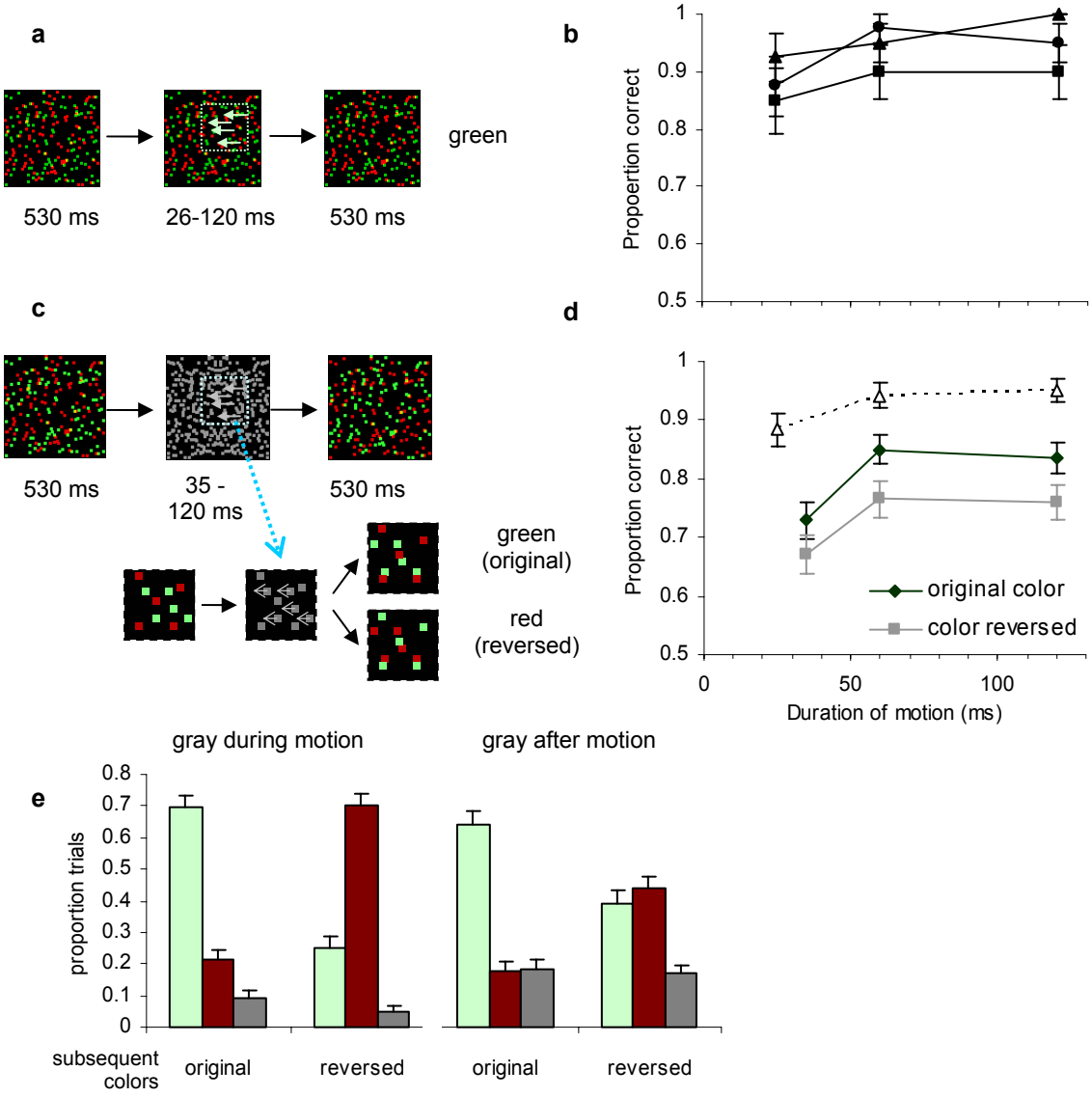
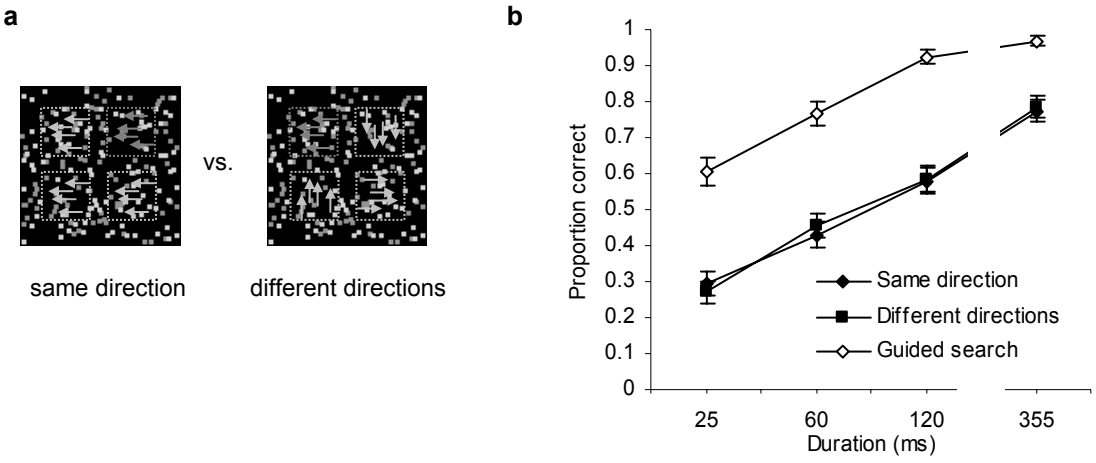

a
530 ms
26-120 ms
530 ms
green
b
Propoertion correct
1
0.9
0.8
0.7
0.6
0.5
c
530 ms
35 -
120 ms
530 ms
green
(original)
red
(reversed)
d
Proportion correct
original color
color reversed
0
50
100
Duration of motion (ms)
gray during motion
gray after motion
e
proportion trials
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
subsequent
colors
original
reversed
original
reversed

Figure 8


a
vs.
same direction
different directions
b
Proportion correct
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
25
60
120
355
Duration (ms)
Same direction
Different directions
Guided search