

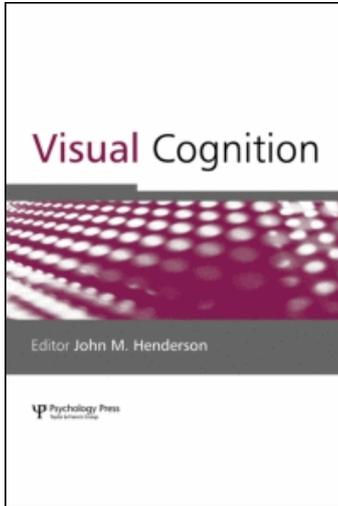
This article was downloaded by: [University Of British Columbia]

On: 19 August 2010

Access details: Access Details: [subscription number 917249135]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Visual Cognition

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713683696>

The role of temporal synchrony in perceptual object formation and updating

Chris Oriet^a; James T. Enns^b

^a University of Regina, Regina, Saskatchewan, Canada ^b University of British Columbia, Vancouver, British Columbia, Canada

First published on: 28 June 2010

To cite this Article Oriet, Chris and Enns, James T.(2010) 'The role of temporal synchrony in perceptual object formation and updating', Visual Cognition, 18: 8, 1179 – 1213, First published on: 28 June 2010 (iFirst)

To link to this Article: DOI: 10.1080/13506281003791009

URL: <http://dx.doi.org/10.1080/13506281003791009>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The role of temporal synchrony in perceptual object formation and updating

Chris Oriet

University of Regina, Regina, Saskatchewan, Canada

James T. Enns

University of British Columbia, Vancouver, British Columbia, Canada

The appearance and disappearance of an object in the visual field is accompanied by changes to multiple visual features at the object's location. When features at a location change asynchronously, the cue of common onset and offset becomes unreliable, with observers tending to report the most recent pairing of features. Here, we use these last feature reports to study the conditions that lead to a new object representation rather than an update to an existing representation. Experiments 1 and 2 establish that last feature reports predominate in asynchronous displays when feature durations are brief. Experiments 3 and 4 demonstrate that these reports also are critically influenced by whether features can be grouped using nontemporal cues such as common shape or location. The results are interpreted within the object-updating framework (Enns, Lleras, & Moore, 2010), which proposes that human vision is biased to represent a rapid image sequence as one or more objects changing over time.

Keywords: Feature binding; Object representation; Attention; Object updating.

When an object appears and then disappears from view, not only do its features tend to occur in the same spatial location, but its features also appear and disappear together in time. To understand the problem this poses for the visual system, the *physical* proximity of an object's features in space and time must be contrasted with the mediation of its perception at a *neural*

Please address all correspondence to Chris Oriet, Department of Psychology, University of Regina, 3737 Wascana Parkway, Regina, Saskatchewan, Canada S4S 0A2. E-mail: chris.orient@uregina.ca

Portions of this work were presented at the annual meeting of the Vision Sciences Society, May 2004, Sarasota, FL, USA, and at the annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Sciences, July 2005, Montreal, Quebec, Canada. This work was supported by Natural Sciences and Engineering Research Council of Canada grants to CO and JTE.

level. Object perception in the brain involves interaction among neural circuits in functionally distinct and often spatially disparate brain regions, leading to the *binding problem* (see Robertson, 2003; Treisman, 1996, for reviews). To further complicate matters, many of these brain regions receive sensory signals from the same physical event at different times and the presence of neural activity in itself does not always correlate with awareness (Lamme, 2006). For example, within 100 ms following the onset of a visual event, neural signals in the higher-order parietal and temporal cortices send feedback signals that modify activity in the primary visual cortex (Lamme & Roelfsema, 2000), making it difficult to disentangle the contribution of sensory signals from internally-generated expectations. Yet, somehow, these neural processes are able under many circumstances to yield the perceptual experience of objects as unitary and coherent entities. In this paper, we examine the role played by the concurrent temporal onset and offset of visual features in the formation of object representations, focusing especially on the role played by geometric and chromatic factors when temporal information is no longer a reliable guide for binding features to one another.

Although the features of an object are often spatially and temporally contiguous, neither of these physical cues is essential or definitive for object perception. Consider the case of a rabbit with its white, winter coat, hiding from a predator in a snowy terrain. While motionless, the rabbit is invisible to the predator who might incorrectly assign the features of the rabbit (e.g., its long ears, its body, and its tail) to the surface features of the snow; despite the spatial contiguity of the rabbit's body parts, no representation of a rabbit is formed. However, once the rabbit begins to move and its various features move together, it suddenly becomes visible against the snow. Thus, in the absence of strong shape and colour cues for grouping features into objects, the temporal cues of synchronous onsets, motion, and offset can become effective for object perception (Lee & Blake, 1999; Sekuler & Bennett, 2001; Usher & Donnelly, 1998).

Temporal cues in the absence of correlated spatial cues also give rise to feature bindings and object perceptions, although some of these interpretations lead to predictable illusions. Consider metacontrast masking, in which two spatial patterns are presented at different points in time (typically the onset of a target shape is followed after 50–100 ms by the onset of a masking shape that fits snugly around the contours of the target shape). Although the visual system registers the unique temporal onset of each of the two patterns, as evidenced by speeded responses made to the onset of the target or the mask shape (Fehrer & Raab, 1962) and by measuring the associated electrical activity of the brain directly (Bridgeman, 1980), the viewer forms a conscious object representation of only the second, larger, mask shape. Yet, because no perceptual representation corresponding to the target shape is formed, the features that make it distinct from the mask (e.g., its smaller size,

colour, texture) are often incorporated into the perceptual experience that is assigned to the mask shape, such that viewers experience a mask shape that seems to expand briefly in size when it appears, taking on some of the colour or texture of the target shape that is invisible (Enns, 2002; Herzog & Koch, 2001; Wilson & Johnson, 1985).

Although the inherent ambiguity of object perception under conditions of unreliable spatial and temporal cues has been known for a long time, the literature on binding shape and colour features across space (beginning with Treisman & Gelade, 1980) is much larger than the corresponding literature on binding the same features over time. Because temporal synchrony of neural responses has been proposed as a mechanism for solving the binding problem (Singer & Gray, 1995), important trends and controversies are now emerging in the growing literature devoted to temporal aspects of binding (e.g., Elliot, Shi, & Kelly, 2006; Engel, Fries, König, Brecht, & Singer, 1999; Hommuk & Bachmann, 2009; Kanwisher & Driver, 1992; Moutoussis & Zeki, 1997a, 1997b; Palanca & DeAngelis, 2005).

One theoretical framework proposed for understanding the coordination of spatial and temporal binding begins with the premise that the visual system is biased to represent a rapid image sequence as one or more objects changing over time (Enns, Lleras, & Moore, 2010; Lleras & Enns, 2004; Moore & Enns, 2004). This is consistent with the heuristic that the world is generally a stable place; our visual-motor interactions with objects usually involve changes that come about as a consequence of our own motion or the motion of the object with which we are interacting. Thus, when the features making up an object undergo a change, the representation of that object formed by the visual system must be updated to reflect the change, to ensure that the most recent appearance of an object is represented. Alternatively, if there is sufficient evidence that the change to the scene involves the introduction of a new object, rather than a change in the characteristics of an old one, then a new object representation must be formed and incorporated into a model of the current environment.

Whether a change in the features of an object lead to updating of an existing object representation or to the formation of a new one has recently been shown to play a role in a variety of perceptual phenomena. One is the *attentional blink* that occurs when two targets must be identified in rapid succession (Raymond, Shapiro, & Arnell, 1992). The difficulty in identifying the second target is greatly reduced when the second target is seen as belonging to the same object representation as the first target (Raymond, 2003). A second is the *flash-lag effect* that occurs when comparing the position of a moving object with the occurrence of a brief flash (Nijhawan, 1994). In this case the perceptual lag is greatly reduced when the motion path is seen as comprised of two separate objects rather than only one (Moore & Enns, 2004; Moore, Mordkoff, & Enns, 2007). A third is visual *backward*

masking (Breitmeyer, 1984), which is also greatly reduced when the target and mask are perceived as belonging to different object representations (Di Lollo, Enns, & Rensink, 2000; Enns, 2002; Jiang & Chun, 2001). In all three of these cases, whether the visual system represents a rapid sequence of images as either a single object undergoing change or as a series of different objects strongly influences the accuracy of reporting a visual target.

What are the factors influencing whether object updating or new object formation occurs? According to the object updating framework, two main principles govern the perception of images that change over time:

1. The formation of a new object (object registration) will occur instead of updating of an existing object (object updating) as a direct consequence of the number and strength of cues present in an image signalling the likelihood of a new object. These cues include violations of spatial contiguity, geometric and chromatic similarity, and temporal simultaneity (Bennett, Lleras, Oriet, & Enns, 2007; Lleras & Enns, 2004; Moore et al., 2007).
2. New object registration is more likely than object updating when focused attention is allocated to the location of an image change before it occurs. Stated conversely, attention at the location of a change protects the representation of the first image from masking by the second, thus permitting new object formation and preventing object updating (Enns, 2004; Enns & Di Lollo, 1997). This role of attention has been demonstrated in metacontrast masking (Enns, 2004), object substitution masking (Di Lollo et al., 2000), and the attentional blink (Brehaut, Enns, & Di Lollo, 1999). In all these cases, focusing attention in advance of an image change allows an object representation to be formed more quickly, thereby making the newly formed representation less vulnerable to updating from a subsequent change in the image and freeing up resources for the formation of a second object based on the changed image.

The flash-lag illusion (Nijhawan, 1994) serves as a convenient illustration of the operation of both principles. In a typical version of this illusion, a target shape travels along a predictable path and at some point a second object appears and disappears suddenly (the flash) adjacent to the target. Observers are required to indicate where the target is located at the time of the flash. The standard finding is that observers localize the target further along its path than it actually is at the time of the flash. An object updating account, based on the first principle, is that the spatial, temporal, and chromatic differences between the moving and the flashed shapes signal the appearance of a new object at the time of the flash, whereas the smooth motion and stable appearance of the moving target signal the same object

changing only in its spatial position over time. However, the registration of a new object not only takes some time, but often also attracts focused attention. According to the second principle, by the time attention can shift back to the moving object its location feature will have been updated and it is this updated location that is compared in experience to the temporal onset of the flashed shape (Baldo & Klein, 1995).

Moore and Enns (2004) tested the first principle of the object updating framework by implementing large changes to the moving target shape at the time of the flash. These changes (in size or colour) to the moving target served as cues that, at the time of the flash, not only did a new object appear in the flashed location, but another new object appeared in the path of the moving target. Consistent with this hypothesis, participants experienced the flash-lag illusion when the moving target shape remained unchanged, but they did not experience the illusion when a large change occurred in the moving target at the time of the flash. Because the shape or colour change that occurred along the path of motion was not seen as belonging to the original moving target, its location feature was not updated after the flash had been registered.

Predictions made by the second principle of the object updating framework—focusing greater attention on the target reduces the likelihood of object updating (reducing the illusion), whereas withdrawing attention increases its likelihood (increasing the illusion)—have also been confirmed. Reducing the predictability of the flash location (Baldo, Kihara, Namba, & Klein, 2002; Namba & Baldo, 2004; Vreven & Verghese, 2005), reducing the task-relevance of the flash (i.e., increasing the likelihood it will capture attention; Chappell, Hine, Acworth, & Hardwick, 2006), and dividing attention between the flash-lag task and another attention-demanding task (Sarich, Chappell, & Burgess, 2007), all serve to strengthen the illusion, as is expected if object updating is more likely when spatial attention is directed away from the target.

In the present series of experiments, we apply the object updating framework to the perception of displays in which the temporal onsets and offsets of features are unreliable, in order to study the process of binding shape and colour features to objects in the absence of reliable temporal cues. The special status of objects in attentional selection is well established (e.g., Desimone & Duncan, 1995; Egly, Driver, & Rafal, 1994). Rather than presenting a new model of object representation, our goal is to better understand the spatiotemporal factors that influence processes of object formation. In all experiments, participants view sequences of images in which the relevant features begin and end either synchronously (onsets and offsets are in phase) or asynchronously (onsets and offsets are out of phase). In temporally asynchronous sequences, a given feature (e.g., the orientation horizontal or vertical) is physically paired with each of two values of a

second feature (e.g., the colours red and blue). When this asynchrony is maximal, the two values of the second feature are each paired with the first for equal lengths of time (e.g., vertical bars are red for the same amount of time they are blue). Observers are then asked to report the colour (report feature value; e.g., red) paired with the orientation (defining feature value; e.g., vertical). If temporally asynchronous displays pose no special problem for the visual system—other than reducing the overall reliability of feature binding—then observers should choose randomly between the two colours physically paired with a defining orientation, and should report each colour with equal frequency. If, on the other hand, the visual system tends to represent the feature pairing based on its most recent updating of the appearance of the same object—as predicted by the object updating framework—then the colour that was paired last in the sequence with a defining orientation feature should be reported at rates greater than expected by chance. A comparison of the proportion of observed reports of the last feature compared to the proportion expected by chance yields an index of the difficulty of binding features with asynchronous temporal offsets.

EXPERIMENT 1

We tested the first principle of object updating by comparing reports of feature pairing when participants were viewing displays in which all features appeared and disappeared simultaneously (synchronous displays) with displays in which features changed at different times (asynchronous displays). Figure 1A shows the overall spatial layout of the displays. Eight moving squares were arranged in two rows of four squares each. Adjacent squares moved in opposite directions, so that at any given time four squares were moving upward and four were moving downward. Figure 1B illustrates how colour and motion features changed asynchronously in some displays, and Figure 1C illustrates the time course of feature changes in the four conditions that were compared. Holcombe and Cavanagh (2001) reported that participants could accurately pair visual features changing synchronously at very high rates of alternation, provided that the features to be bound appeared in the same location. However, when the colour and orientation to be paired appeared in different locations, thresholds for pairing with 75% accuracy increased by a factor of 10. In Experiment 1 we asked participants to pair colour and motion in cycling displays, in which the features were each presented for 300 ms, and the onset and offset of features was either synchronous (Figure 1C, left side) or asynchronous (Figure 1C, right side). Asynchronous features were out of phase with one another by half a cycle (i.e., by 150 ms).

Thus, for one group of participants, motion was assigned as the defining feature and colour was the report feature (colour group); for the other group, this assignment was reversed (motion group). This between-groups factor therefore tested the hypothesis that shifting attention from a defining to a report property would increase the probability of object updating. This manipulation has been used previously in reports of perceptual simultaneity (e.g., Arnold, 2005; Clifford, Arnold, & Pearson, 2003; Holcombe & Cavanagh, 2008). The prediction here was that reports should be biased towards the later of the two features paired for equal durations with the defining feature value, regardless of whether the report feature was colour or motion.

To summarize, we predicted that when temporal onsets and offsets were synchronous, participants would accurately report the value of the report feature that was actually presented with the defining feature, regardless of which feature was assigned to each role (e.g., when red was paired with upward motion, reports would be accurate regardless of whether participants were asked to report colour or motion direction). With asynchronous temporal onsets and offsets, however, we expected the role assigned to features in the participant's task to influence pairing reports. For example, as shown in Figure 1C for the asynchronous condition labelled red-blue/down-up (upper-right panel), when participants in the colour group shifted attention from the defining motion feature of "up" to the report feature of colour they should tend to report the colour "blue" rather than "red", because "blue" was the later of the two colours that were paired with "up". In contrast, when participants in the motion group examined the same displays, now shifting their attention from the defining motion feature of "red" to the reporting feature of motion, they should tend to report the direction "up" rather than "down", because "up" was the later of the two colours that were paired with "red". What is remarkable is that the colour group is predicted to pair "up" with "blue", whereas the motion group is predicted to pair "up" with "red", even though both groups are viewing the same display sequences. An alternative outcome for this experiment, one not favouring the object updating hypothesis, would be that participants would make the same colour and motion pairing in both conditions, perhaps because one of the features is registered more slowly than the other, leading to perceptual lags caused by the transmission of feature information rather than because of shifting attention (e.g., Moutoussis & Zeki, 1997a).

Method

Participants. Twenty-four students from the University of British Columbia volunteered in exchange for partial course credit. All reported

normal or corrected-to-normal vision. Twelve were assigned to the colour group and twelve to the motion group.

Stimuli and procedure. Displays were presented on a Macintosh eMac computer with the colour screen set to a refresh rate of 60 Hz. As shown in Figure 1A, displays consisted of eight squares, arranged in two rows of four columns, with the two different types of items in each display arranged so that no adjacent items were identical. The centre-to-centre distance between display items was 3.0 degrees of visual angle and each item occupied a square region of 1.3° . Each square in these displays was either blue or red, and each square was also moving upward or downward, as illustrated by the arrows in Figure 1B. In any given display frame, four of the items were identical in colour and motion direction; the other four were of the opposite colour and motion direction. Display frames were 150 ms in duration, such that each feature was presented for 300 ms, and an entire cycle of four frames took 600 ms. Displays continued to cycle until the participant pressed a response key.

Four different display sequences were presented in random order across trials. Two of these were synchronous displays, as shown in Figure 1C (left side). In these sequences, the upward moving squares were always blue and the downward moving squares were always red (blue-blue/down-down), or the upward moving squares were always red and the downward moving squares were always blue (red-red/up-up). In the asynchronous displays, shown in Figure 1C (right side) the same features were presented for the same durations, but the changes in colour and motion were offset by a phase shift of 90° (one 150 ms frame; red-blue/down-up or blue-red/up-down).

Participants in the colour group were instructed to indicate the colour of the squares when they were moving upward; those in the motion condition were instructed to indicate the direction of motion when the squares were red. The response keys “z” (blue or horizontal) and “/” (red or vertical) were used to make these responses and there was no speed stress in the task. On average participants made these decisions within 2–3 s of display onset. Each participant completed at least 100 trials (four blocks of 20 trials separated by a short self-paced break) with several participants in each group completing 200 trials (eight blocks).

Results

The mean proportion of colour and motion reports is shown in Figure 2. Colour reports were coded arbitrarily as proportion red (the complement is proportion blue), motion direction reports were coded as proportion upward (the complement is proportion downward) for the four types of display

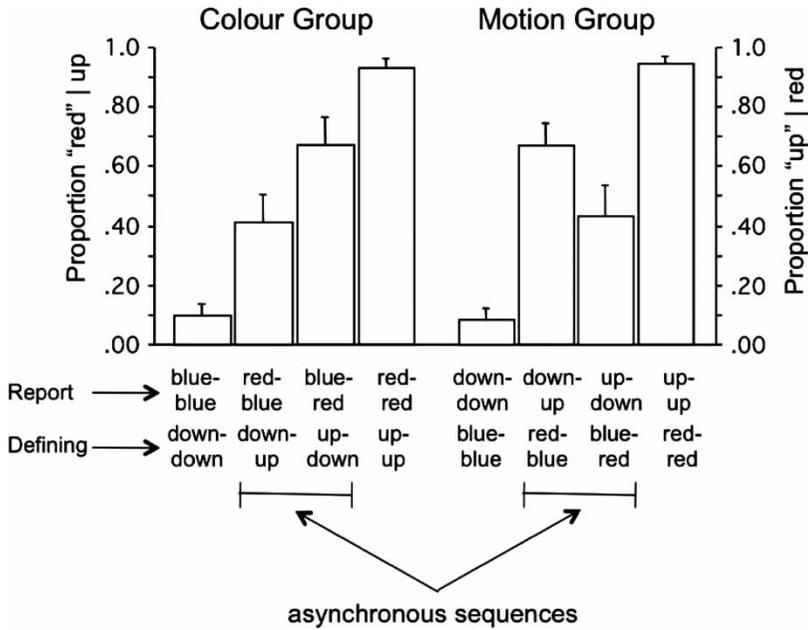


Figure 2. Experiment 1: Mean proportion “red” reports (given an upward direction of motion) in the colour group and the mean proportion “up” reports (given a red colour) in the motion condition. The four conditions in each group are labelled according to the feature values that were paired in the cycling displays shown in Figure 1.

sequences. The main finding was a significant tendency to report the second of the two paired report feature values more frequently than the first paired report feature in asynchronous displays, even though they were each paired with a defining feature value for the same length of time. This tendency was observed both for participants reporting the colour of the moving squares (colour group in Figure 2) and for those reporting the motion direction of the red squares (motion group in Figure 2).

These observations were supported by a mixed-model ANOVA involving the between-participants factor of group (report colour vs. report motion) and the within-participants factor of condition (blue-blue, red-blue, blue-red, red-red). A significant Group \times Condition interaction, $F(3, 66) = 3.30$, $MSE = 0.075$, $p < .02$, indicated that the tendency to report feature values in asynchronous displays (the middle two bars in each panel of Figure 2) varied with the report participants were asked to make. Direct comparisons involving the two asynchronous conditions showed the colour group tended to report the last of two colour values that were paired with the defining feature of vertical motion more frequently than expected by chance, $t(66) = 2.34$, $p < .05$. When viewing the same displays, the motion group tended to

report the last of two motion directions paired with the defining feature of red, also more frequently than expected by chance, $t(66) = 2.00$, $p < .05$.

Discussion

As previously reported by Holcombe and Cavanagh (2001), reports of feature pairing were very accurate for synchronous displays at the rate tested (300 ms of simultaneous feature presentation or 1.67 Hz). Reports of feature pairing for asynchronous features, however, favoured the later of the two values paired with the defining feature. If observers simply could not discriminate which colour or motion was paired with the defining feature and were choosing feature values at random, the proportion of “red” and “upward” reports for the asynchronous conditions shown in Figure 2 (two middle bars) should have approached 50%. Also, if detection of one of the defining features (e.g., up motion) was processed more slowly than detection of the other defining feature (e.g., red colour), then the later colour should have been reported by the colour group and the earlier motion direction should have been reported by the motion group. Instead, the second of the two colours paired with “up”, and the second of the two motions paired with “red” was reported at rates that differed both from rates of the first feature and from rates expected by chance. According to object updating theory, such a pattern is expected if the visual system is biased to update an object undergoing a change with the most recently available value of a changing feature.

In this experiment, the amount of time each report feature value was paired with the defining feature was equal (half the total duration of the defining feature). As such, the duration over which features were physically paired differed for synchronous and asynchronous conditions. For example, in the synchronous condition “red” and “up” were paired for 300 ms; in the asynchronous condition they were paired for only 150 ms before either the colour or direction of motion changed. Given that the features were also maximally out of phase with one another (90°) this could be considered an extreme case. In the next experiment we systematically varied the temporal offset between feature changes in smaller steps in addition to varying the overall feature duration. This allowed us to (1) examine the influence of finer gradations in phase-offset and (2) directly compare feature reports in synchronous and asynchronous displays when the durations of the two features paired were equated in total time.

It is possible that the last-feature report tendency observed in Experiment 1 simply reflects a tendency to report the most recently seen feature pairing, irrespective of whether the changed feature is assigned to an existing object representation or to a new one. We tested for this possibility in Experiment

2. If participants simply remember the most recently seen feature pairing better than the previous one, this tendency should increase as features are paired for longer durations. Note that object updating makes the opposite prediction, in that last-feature reports should decrease as feature duration increases. This is because increasing the time that two features are paired should increase the likelihood of consolidating a stable object representation before one of the feature values changes.

EXPERIMENT 2

Experiment 2 was designed as a second test of the first principle of object updating, namely, that asynchronous feature changes bias the visual system to update an existing object rather than to form a new one. We did this by manipulating the phase-shift of the feature changes (in 45° steps) independently from the duration of each feature value (64 or 128, 256, 512 ms). Note that an asynchronous cycling display poses two different problems for the visual system. The brief duration of each feature presentation means that report errors may occur simply because not enough time has elapsed for a feature to be registered before its values have been replaced. The asynchronous nature of the changes means that some cue other than temporal contiguity of onset or offset must be used to establish temporal congruence of the features. By varying phase and feature duration independently, we could decouple the contributions of asynchrony per se from those of brief feature durations on the difficulty of feature binding.

This design also allowed us to compare conditions in which two features were paired for an identical duration but were presented together in quite different contexts. For example, a 150 ms simultaneous presentation of red and upward motion could occur in the slower, asynchronous displays (upward moving squares are blue for 150 ms then red for 150 ms), or in faster, synchronous displays (upward moving squares are red for 150 ms, but never turn blue during upward motion). If the asynchronous displays used in Experiment 1 were difficult only because feature values are paired for a shorter duration than in synchronous displays, performance in these two conditions should not differ. Rather than repeating our tests with colour and motion, as in Experiment 1, here we paired colours (blue, red) with orientation (horizontal, vertical) in order to generalize the point that asynchronous feature presentations pose a problem of ambiguity for the perceptual system. Participants were instructed to report the colour of the rectangles whenever they were in vertical orientation.

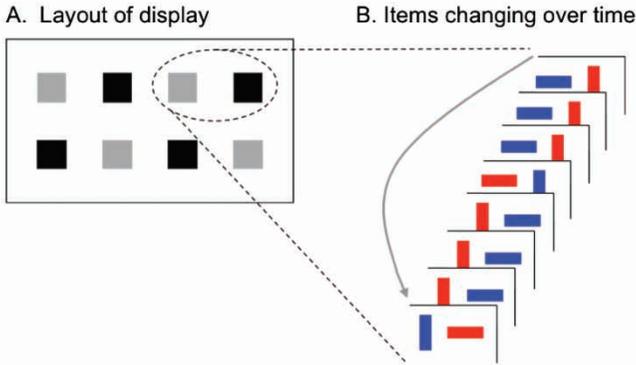
Method

Participants. Thirty-three undergraduate students (27 female) volunteered their participation in exchange for partial course credit. Data from three participants were excluded from consideration because they seemed to misunderstand the instructions, selecting the colour paired with the horizontal orientation rather than the correct (vertical) orientation on more than 90% of synchronous trials. Fifteen of the remaining participants were tested using frame durations of 32 ms, 64 ms, and 128 ms (University of British Columbia) and 15 participants were tested with frame durations of 16 ms, 64 ms, and 128 ms (University of Regina).

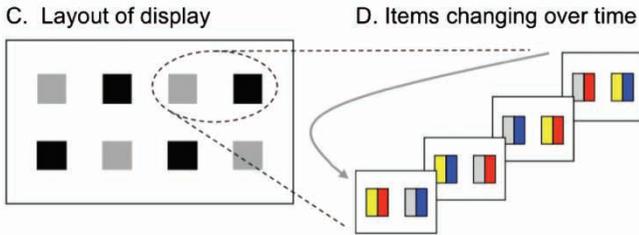
Stimuli and apparatus. Display sequences consisted of a spatial array of rectangular items, laid out in two rows of four columns, so that no adjacent items were identical in orientation or colour in any single frame. The centre-to-centre distance between display items was 3.0 degrees of visual angle and each item occupied a square region of 1.3° . Each rectangle in these displays was either blue or red, and each rectangle was oriented either vertically or horizontally, as illustrated in Figure 3. These displays continued to cycle until the participant pressed a response key.

Procedure. Each display cycle consisted of four frames in which a given element was vertical followed by four frames in which it was horizontal. Colour was alternated in a similar way, with each element presented in red for four frames followed by blue for four frames. Displays differed in the phase relationship between these two features and in the duration of each display frame. Phase was varied in 8 steps of 45 degrees, ranging from 0 to 315 degrees; each phase shift was chosen quasirandomly, with equal probability, with the restriction that the same phase shift could not appear on more than four consecutive trials. Frame durations were manipulated in multiples of 16 ms, corresponding to the vertical refresh rate of the monitors used in testing. The shortest possible feature duration was 64 ms (i.e., four 16 ms frames, to allow for varying phase relationships in 45 degree steps). For one group, the feature duration was 128 ms, 256 ms, and 512 ms; for the other group it was 64 ms, 256 ms, and 512 ms; this duration was chosen randomly on each trial. The fact that these groups differed only on the shortest feature duration meant that we could also test whether performance on the middle feature duration was influenced by a range effect; namely, the context of the other durations experienced in the same session.

Experiment 2



Experiment 3



E. Grouping Conditions

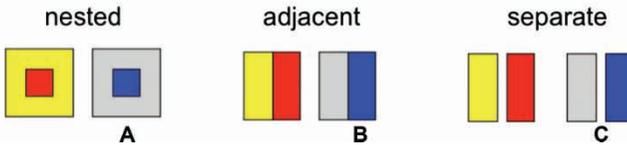


Figure 3. (A) Spatial layout of the display shown in Experiment 2, showing how the two different items alternate across the eight positions (indicated by grey and black squares). (B) Illustration of how items in Experiment 2 change colour and orientation over time in the asynchronous conditions with a phase shift of 45° (blue for the first of four frames when the bar was vertical and red for the remaining three frames). Each feature was visible for either a short duration (64 or 128 ms, consisting of 4 frames lasting either 16 or 32 ms), a medium duration (256 ms, consisting of 4 frames of 64 ms), or a long duration (512 ms, consisting of 4 frames of 128 ms). (C) Spatial layout of the display used in Experiment 3, showing how the two different items alternate across the eight positions (indicated by grey and black squares). (D) Illustration of how items change colour and orientation over time asynchronously in Condition B of Experiment 3. (E) The three conditions tested in Experiment 3, differing in how strongly the defining feature (yellow shape) is perceptually grouped with the report feature (blue or red shape): Nested (A), adjacent (B), and separate (C).

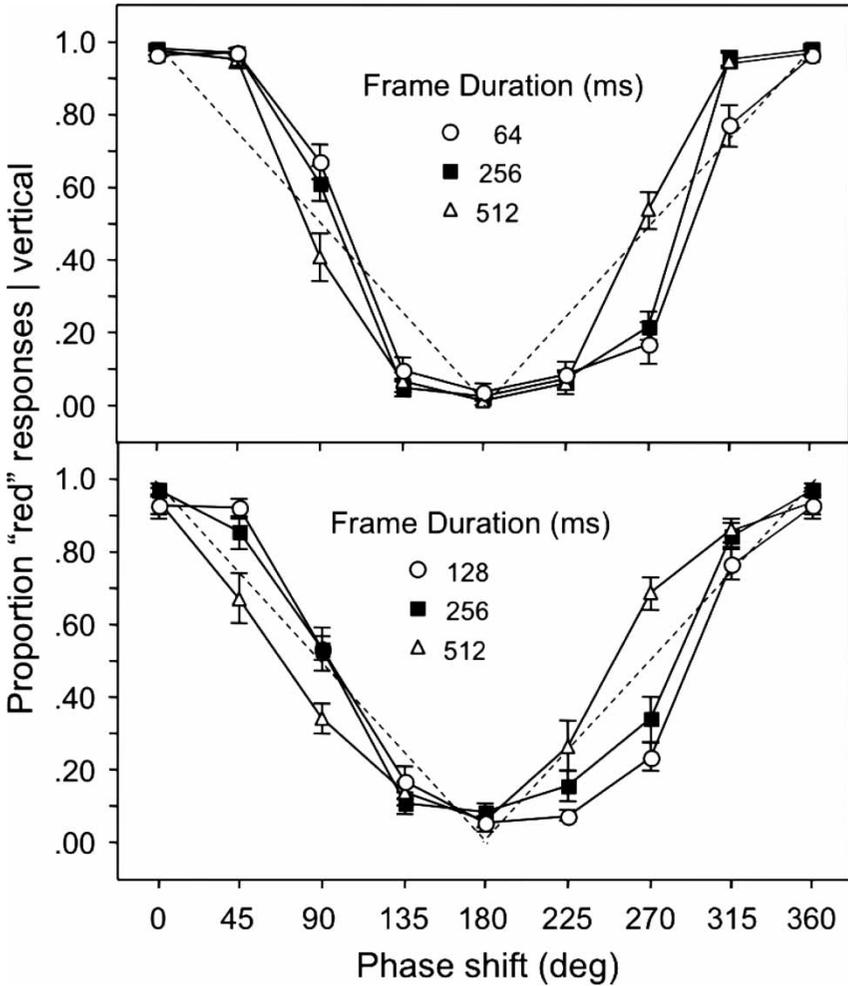


Figure 4. Experiment 2: Mean proportion of "red" reports when the bar was vertical, as a function of phase shift and feature duration for the group with 64 ms as shortest feature duration (upper panel) and for the group with 128 ms as shortest feature duration (lower panel). At 0° phase shift, bars were red for entire time they were vertical; at 45° they were blue for the first of four frames and red for the remaining three frames; at 135° they were blue for the first three frames and red for the remaining frame. At the 180° phase shift, bars were blue for the entire time they are vertical. The dashed lines in each panel illustrate the expected responses if perception corresponded to the relative physical durations of each colour.

Results

Figure 4 shows the mean proportion of “red” reports (chosen arbitrarily; the complement is the proportion of “blue” reports) for the two groups of participants, plotted as a function of the phase relationship between features and feature duration. The dashed line in the figure shows the proportion of “red” reports expected if responses simply tracked the relative amount of colour paired with each period of vertical orientation.

The main finding was that frame duration influenced perceptual reports for asynchronous displays (45° , 90° , 135° , 225° , 270° , 315° phase shift conditions) more than for synchronous displays (0° , 180° conditions). This observation was supported by a mixed design ANOVA involving the between-participants factor of group (shortest duration = 64 ms, 128 ms) and the within-participants factors of phase shift (0, 45, 90, etc.) and duration (short, medium, long). This analysis revealed an overall interaction of Group \times Phase shift \times Duration, $F(14, 392) = 2.24$, $MSE = 0.015$, $p < .007$, as well as significant Phase shift \times Duration interactions in each group considered separately, $F(14, 196) = 11.4$, $MSE = 0.013$, $p < .001$, and $F(14, 196) = 11.8$, $MSE = 0.017$, $p < .001$, for the 64 ms and 128 ms group, respectively.

More detailed comparisons showed that for the group tested with 64 ms as the shortest duration (upper panel in Figure 4), there was a significant difference between observed reports and those expected by chance (based on the actual percentage of time the vertical shape was red, as indicated by the broken line in Figure 4) in both the short and medium feature durations for the 90° phase shift condition (mean difference = .17 and .11, respectively) and for the 270° phase shift condition (mean difference = .33 and .28, respectively), all $t(14)$ values > 2.3 , all p -values $< .04$. For the group tested with 128 ms as the shortest feature duration (lower panel in Figure 4), “red” reports occurred less frequently than expected by chance at the 270° phase shift for both short (mean difference = .26), $t(14) = -6.78$, $p < .01$, and medium feature durations (mean difference = .16), $t(14) = -2.45$, $p < .03$. In contrast, for the longest feature duration (512 ms in each group), perceptual reports favoured the colours shown at the onset of the vertical orientation, which is a reversal of the effect observed in Experiment 1. This effect was not reliable for the 64 ms group, but was reliable for the 128 ms group at both 90° (mean difference = $-.16$), $t(14) = -3.69$, $p < .003$, and 270° phase shifts (mean difference = $-.19$), $t(14) = 4.10$, $p < .002$. Finally, frame duration had no significant influence in either of the two synchronous conditions in either group (0° , 180° conditions), all p -values $> .20$.

Another comparison examined the group differences in the Phase shift \times Duration interaction. As shown in Figure 4, reports of “red” in the 256 ms condition varied considerably depending on whether the shortest feature

duration was 64 or 128 ms. An analysis excluding the shortest feature duration in each group (because this value differed for the two groups) revealed a significant Group \times Phase shift \times Duration interaction, $F(7, 196) = 2.15$, $MSE = 0.010$, $p < .05$. Inspection of the means indicated that the difference in observed and expected “red” reports at a feature duration of 256 ms was larger overall when seen in the context of much shorter frame durations (shortest frames = 64 ms) than in the context of longer ones (shortest frames = 128 ms).

Finally, comparing the shortest feature duration in each group (64 ms vs. 128 ms) afforded a unique opportunity to test the claim that asynchronous trials pose a problem for the visual system independently of feature duration. With a 64 ms feature duration and synchronous trials (0° and 180° phase shifts), vertical orientation was paired with a single colour for 64 ms. This can be compared with a 128 ms feature duration on maximally asynchronous trials (90° and 270° phase shift), in which each colour is paired with vertical orientation for exactly this same duration (64 ms). These two conditions are equal, both in the absolute durations of the feature pairings and in the rate at which the bar changes at least one of its features (i.e., once every 64 ms). Yet, on synchronous trials (64 ms) the colour physically paired with vertical was chosen 95.9% of the time whereas on asynchronous trials (128 ms) the first colour paired with vertical (for 64 ms) was chosen only 35.7% of the time, $t(28) = 11.4$, $p < .001$, with the second colour pairing receiving the majority of reports (64.3%).

Although speed of response was not emphasized to participants in this experiment, we incidentally measured the time participants took to make their reports, which allowed us to examine how their decision times were related to relative report frequency. These analyses all showed that decision time generally decreased as report frequency increased for a given feature value (i.e., as perceptual certainty increased), giving us increased confidence that the colour reports were reliably related to participants’ experiences and not complicated by guessing strategies.

Discussion

In contrast to expectations based on either veridical perception or random responding, colour reports were disproportionately influenced by the second colour appearing on the vertical bar. For example, when the vertical bar was red for the last 75% of its period (45° phase shift condition) or blue for the last 75% of its period (225°), the bar was reported as red (or blue) about 90% of the time. In conditions in which the bar was red for only the last 50% of its vertical period, red reports were again overrepresented when red was the last colour (90° phase shift) with participants reporting the bar as red about 68%

of the time and underrepresented (about 20%) when red appears as the first colour (270° phase shift).

The strength of the tendency to report the later appearing colour paired with vertical was qualified, however, by an interesting context effect. When the shortest feature duration tested in the experiment was 64 ms, colour reports in the 256 ms feature duration condition strongly favoured the second of the two colours paired with vertical. However, this tendency was greatly attenuated (and, at the 90° phase shift, eliminated entirely) at the same feature duration when the shortest feature duration tested in the experiment was 128 ms. The tendency to report the later appearing colour was therefore increased when displays changed more rapidly (shorter feature durations) and, at the other extreme, was nearly eliminated when displays had a slow rate of alternation (longest feature duration). However, regardless of feature duration, a difference between observed and expected reports was much more likely to occur with asynchronous than with synchronous displays, even when frame duration was held constant. When the vertical shape appeared in only one colour, this colour was correctly reported as paired with vertical, even at the shortest feature duration.

Comparing performance across conditions that equated the physical duration of the paired feature values reinforced our interpretation that asynchrony per se is problematic for the visual system. Specifically, a vertical bar that was blue for the first 64 ms of its 128 ms vertical period and red for the remaining 64 ms was reported as red on the majority of trials; a vertical bar that was blue for 64 ms and then changed to a horizontal red bar was almost always (correctly) reported as blue. Comparing these two conditions also highlights the fact that the time needed to switch attention from the defining feature to the report feature per se cannot fully explain the last-feature report tendency. When reliable spatiotemporal cues are present (as on synchronous trials), observers accurately pair the only colour presented with vertical; if the speed of the attention switch was the only relevant factor, the other colour should have been reported as often in the 64 ms synchronous condition as the later colour was reported in the 128 ms asynchronous condition (in which a single colour was paired with vertical for 64 ms).

The results of Experiment 2 also counter the possibilities that the greater number of “last” colour reports reflect a response bias rather than a perceptual bias, or that they index a tendency to better recall the most recently seen pairing of features. With a very long feature period of 512 ms, colour reports tended to favour the *first* of the two colours paired with vertical, rather than the second. A general response bias would not have been affected by this change in feature duration. From an object updating perspective, however, this finding is expected, because as a pair of features appears together for a longer duration, it will be possible to consolidate a

stable object representation before the feature changes. According to the second principle of object updating, attention to the report feature prior to a change in its value should buffer it from masking and, as such, updating should not occur.

The results of Experiments 1 and 2 confirm that asynchronous onsets and offsets of features pose difficulties for a visual system trying to bind features into a stable object representation. The solution taken by the visual system—pairing the most-recently presented feature value with a defining feature—reflects the tendency for object updating rather than new object registration. Synchronous displays involving the same brief durations of feature presentations do not pose the same problems, suggesting that temporal coincidence is used by the visual system as a cue for feature binding (Lee & Blake, 1999; Sekuler & Bennett, 2001; Usher & Donnelly, 1998). The results of the feature duration manipulation were also instructive: Increasing the duration of the feature pairing beyond a critical value (i.e., between 256 and 512 ms) was sufficient to allow for the consolidation of an object representation, preventing object updating, even for asynchronous displays. Notably, at these longer feature durations participants did not report the most recently seen feature pairing, as predicted by a mere bias to report the most recently paired feature.

In the next experiment we systematically varied the difficulty of switching attention from the defining feature to the report feature, in an effort to manipulate the probability of object updating occurring. Of course, if the last feature report tendency reflects better recall of more recently seen feature pairings, this manipulation should have no effect on performance. Additionally, we tested whether the reporting tendencies observed in Experiments 1 and 2 would persist even when the defining and report features involve the same perceptual dimension (colour).

EXPERIMENT 3

If object updating plays a causal role in the tendency to report the second of two paired feature values as concurrent with a defining feature value, then manipulating the speed of an attention switch should influence this tendency. According to the second principle given in the introduction, the probability of object updating is greater when attention is not available for consolidating an object representation than when attention is focused on the location of a change. This is presumably because focused attention increases the temporal resolution of visual segmentation processes (Enns, Brehaut, & Shore, 1999; Shore & Spence, 2005). Increasing the time needed to switch attention from the defining feature to the report feature, then, should increase the difficulty of consolidating a stable object representation and increase the likelihood of

object updating. Conversely, the more rapidly attention is deployed to the report feature, the less likely it is that its value will have changed and, correspondingly, the less likely it is that object updating will occur.

In Experiment 3, we varied the time needed to switch attention from the defining feature to the report feature by varying the extent to which the two feature values were bound within an object representation. We expected that when the defining and report features were perceptually *integral*, using Garner's (1974) term, it would be more difficult to segregate them for the purposes of report than when the defining and report features were perceptually *separable*. Pairing integral features should increase the time needed to switch attention from the defining feature to the report feature, thereby increasing the likelihood that the value of the report feature will have changed when attention is redeployed to the report feature. If so, we expected the last-feature report tendency to be largest when the defining and report features were integral rather than separable.

We varied feature integration by capitalizing on a previous report that two colours in a nested-feature relationship (e.g., the window and wall colour of a house) are perceptually grouped more strongly than are two colours on adjacent surfaces (e.g., the colours of each half of a duplex) or even on spatially separated surfaces (Wolfe, Friedman-Hill, & Bilsky, 1994). The stimuli in Figure 3E were designed to vary along this continuum of grouping strength. In each condition, one part of each display item alternated in colour between grey and yellow, whereas the other part alternated between red and blue. Participants were instructed to report the colour paired with yellow. What differed between the conditions was the degree to which the grey–yellow alternating portion was spatially grouped with the red–blue alternating portion.

As a check on the validity of our assumption that these displays varied in perceptual grouping, we first conducted visual search experiments in which the task was to report the colour (blue or red) of the one item in the display that was also yellow. The nontarget items were defined by grey in the regions occupied by yellow in the target item and they varied randomly in whether they were red or blue in the report region. Note that this search task involves the inverse perceptual problem of the task studied by Wolfe et al. (1994). There, participants were presented with many items that were either yellow or grey and red or blue, and were asked to find the single yellow-red item as rapidly as possible. This defines a standard "conjunction search". In the present task there was only a single yellow item among many grey items and the task was to report whether the other colour in the yellow item was blue or red. This means participants must first find the yellow item (defining feature) and then report the other colour (report feature). That step should be more difficult if two parts of an object must be first perceptually segregated or unbound in order to make the report.

Seven participants were tested on 160 trials with search displays containing either one or eight items in a static 2×4 spatial array. Accuracy of search was $> 92\%$. The results were clear in showing that discriminating blue from red in Condition A (mean correct RT = 570 ms) was significantly slower than in Conditions B (mean correct RT = 524 ms) and C (mean correct RT = 538 ms), leading to a main effect of condition, $F(2, 12) = 4.27$, $MSE = 9131.06$, $p < .04$. Direct comparisons confirmed that responses in Condition A were reliably slower than in Conditions B and C combined, $t(12) = 2.16$, $p < .05$, but that responses in Conditions B and C did not differ from each other, $t(12) = 0.78$. Search times were also slower when the display contained eight items rather than only one, $F(1, 6) = 84.78$, $MSE = 2909.99$, $p < .001$, but this effect did not interact with condition, $F(2, 12) = 1.44$, $MSE = 797.79$, $p > .27$. These results confirm that the task was most difficult for the items in Condition A, consistent with the prediction that the two parts of the object would be grouped most strongly in this condition (Wolfe et al., 1994).

Method

A total of 51 university students who had not participated in any other experiment reported in this paper volunteered in exchange for partial course credit. Seventeen participants were assigned to each of the conditions. Display sequences were constructed as in Experiment 2, using the three types of display items shown in Figure 3. The feature duration was 150 ms and only the 0° and 180° phase shifts were tested. Otherwise the methods were as described in Experiment 2.

Results

The mean proportion of “red” reports is shown in Figure 5. Colour reports were coded arbitrarily as proportion of red reports (the complement is proportion blue) while the defining regions were yellow. The x-axis indicates when the stimulus was actually blue in both frames (blue-blue), when it was red only in the first frame (red-blue), when it was red only in the second frame (blue-red), and when it was red in both frames (red-red). There were two main findings. First, reports in each of the asynchronous conditions favoured the feature present during the second frame of the defining feature, indicating that the tendency to report the second of two feature values paired with a defining feature value occurs even when both defining and report features involve the same dimension (i.e., colour). Second, this effect was larger in condition A (the one in which the two colours were perceptually coupled most tightly) than it was in Conditions B and C.

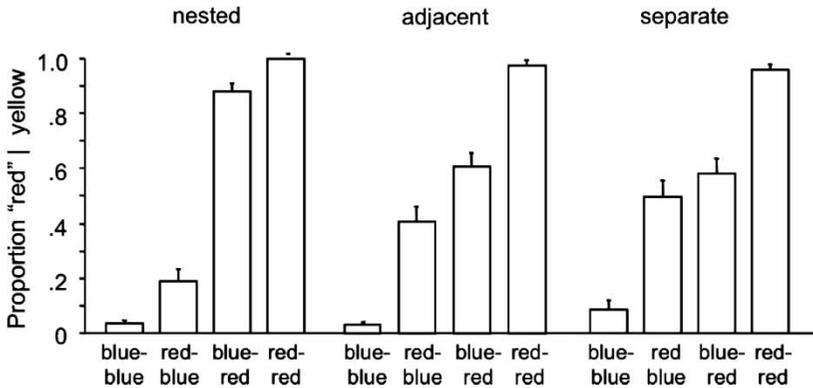


Figure 5. Experiment 3: Mean proportion of “red” reports when the other part of the item was yellow, as a function of how strongly the defining feature (yellow) is perceptually grouped with the report feature.

These observations were supported by a mixed-model ANOVA involving the between-participants factor of grouping (nested, adjacent, separated) and the within-participants factor of condition (blue-blue, red-blue, blue-red, red-red). A significant Grouping \times Condition interaction, $F(6, 144) = 14.13$, $MSE = 0.062$, $p < .001$, was observed, indicating that the tendency to report the later-appearing colour in asynchronous displays (the middle two bars in each panel of Figure 5) varied with grouping. Follow-up tests were carried out using Fisher’s LSD procedure. Direct comparisons of the last-feature reporting tendency indicated it was larger in Condition A (mean difference = .69) and greater than expected by chance, $t(144) = 8.08$, $p < .001$, than it was in Condition B (mean difference = .22), $t(144) = 2.58$, $p < .01$, or in Condition C (mean difference = .08), $t(144) < 1$.

Discussion

This experiment showed that the difficulty of binding two features with asynchronous temporal onsets is not unique to situations in which the two features to be compared involve different dimensions. Indeed, the tendency to report the later-appearing colour in Condition A (“report the colour of the centre square when the frame is yellow”) was much larger than the same tendency we measured in Experiments 1 and 2, in which different dimensions were involved.

The second important finding was that the magnitude of the report error was in direct proportion to the likelihood of object updating in these displays. When these two regions formed a well-integrated perceptual entity (as indicated by the conjunction search tasks of Wolfe et al., 1994, and by

our visual search experiment described in the introduction to the present experiment), the tendency to report the later of the two feature values paired with the defining feature was larger than when the two regions were less strongly grouped. According to the object updating hypothesis, spatially segregating the defining and report features decreased the difficulty of switching (and, correspondingly, the time needed to switch) between the features. This increased the likelihood that attention was available prior to a change in the value of the report feature, which in turn reduced the probability of object updating—and the likelihood of reporting the updated feature. Note that if last feature reports reflected better memory for more recently paired features, then the manipulation of the integrality of the defining and report features should have had no effect on performance, simply because the later of the two values of the report feature was paired with the target feature for the same duration across all three conditions.

Taken together, the results of Experiments 1–3 suggest that asynchronous onsets and offsets of features create a situation that makes it difficult to consolidate a stable object representation, which in turn increases the likelihood that object updating occurs. Conditions that favour the fast redeployment of attention to the report feature reduce the probability of object updating and, in turn, reduce the tendency to report the second of two values paired with the defining feature for an equal duration. In Experiment 4, we test a prediction that follows from the first principle of the object updating hypothesis, that object updating will be prevented if a cue present in the display reliably signals the appearance of a new object.

EXPERIMENT 4

If object updating plays a significant role in the reporting errors of Experiments 1–3, it should be possible to reduce those errors by providing the visual system with a reliable nontemporal cue to object formation. Specifically, according to the first principle of object updating, if reliable spatiotemporal cues signal that an updated feature value should be attributed to a different object, no updating of the initial representation formed by the visual system should occur. Alternatively, if participants simply recall the most recent feature pairing most strongly, then the presence of cues favouring the appearance of a new object over a change to an old one should have no effect on the last feature reporting tendency.

In Experiment 4 we systematically varied the spatial location of the features in the displays. This was a test of the idea that when the displays in Experiment 3 formed a well-integrated perceptual object (Figure 3E, Condition A) the object updating process led to an increase in participants' reports that the later appearing of two values were paired with the defining

feature. In our view, this occurred because the value of the report feature changed during the time that the defining feature was being registered, and the object representation was updated accordingly. Conversely, when the report and defining features were grouped less strongly (Figure 3E, Conditions B and C) object updating was less likely to occur because each alternating item could be represented individually. Establishing separate representations in these conditions increased the likelihood that attention could be redeployed to the report feature before it changed, with the last-feature report tendency being nearly eliminated in Condition C.

Experiment 4 compared three conditions. A Baseline condition was conceptually very similar to Experiment 3, in that colour was used for both the defining and the report features. It differed only in geometry, as illustrated in Figure 6. The defining feature was the colour of a stationary vertical bar that alternated between yellow and grey and the report feature was the colour of a partially occluded horizontal bar that alternated between blue and red. The critical comparison was between two conditions that involved displacement of the horizontal bar in alternation throughout the display sequence. In the synchronous jump condition, the horizontal bar jumped (it appeared in another location) in synchrony with every change in its own colour. For example, if the horizontal bar was blue it remained in one location for the entire blue period and then it appeared in the other location and it turned red. Subjectively, this was experienced as a bar that changed colour every time it changed location. A change in both colour and location should provide a strong cue to the presence of a new object, irrespective of the asynchronous onsets and offsets of the features to be paired. In the asynchronous jump condition this same bar jumped in synchrony with each change in the colour of the vertical bar (defining

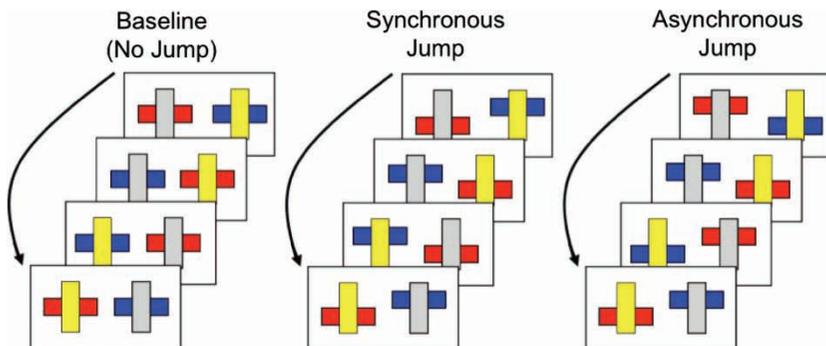


Figure 6. Experiment 4: Illustration of how items change colour and orientation over time. Conditions differed in whether the shape of the report feature was stationary (baseline), was relocated with the change in the report colour (synchronous), or was relocated out of phase with the change in the report colour (asynchronous).

feature). Thus, if the horizontal bar was blue in one location it remained blue even after it moved to another location, where it stayed blue for one frame before becoming red for one frame, and then moved as a red bar to the other location.

According to the object updating account, the tendency to report the second of the two colours paired with the target colour should be diminished in the synchronous jump condition relative to the baseline condition because the horizontal bar can be represented by the visual system as a distinct object every time it changes location (at which time it also changes colour; Lleras & Moore, 2003; Moore & Lleras, 2005). As such, the object representations at each of these locations should be less susceptible to updating based on their colour values at the other location. In contrast, the horizontal bar in the asynchronous jump condition should be maximally vulnerable to object updating by the new colour occurring in the same location. In this condition, there is no reliable cue to signal the presence of a new object that could compensate for the difficulty posed by asynchronous temporal onsets and offsets of the features to be paired, so the conditions that favour object updating are present.

Method

Fifteen students who had not participated in any of the previous experiments volunteered in exchange for partial course credit. The method was identical to that of Experiment 3, with the following exceptions. First, display items consisted of a vertical bar of one colour (yellow or grey) superimposed in front of a horizontal bar of another colour (blue or red), as shown in Figure 6. These bars were similar in their spatial dimensions to the bars used in Experiment 2.

The participants' task was to report the colour of the horizontal bar when the vertical bar was yellow. As in previous experiments, the colour of the horizontal bar (report feature) alternated either in phase or out of phase with the colour of the vertical bar (defining feature). Three different conditions were compared as outlined and illustrated in Figure 6: Baseline, synchronous jump, and asynchronous jump. These three conditions were tested in a counterbalanced order across participants and order played no significant role in the results (all $F_s < 1.0$).

Results

The mean proportion of "red" reports is shown in Figure 7. Colour reports were coded arbitrarily as proportion red (the complement is proportion blue); the defining regions were yellow. The x-axis indicates when the

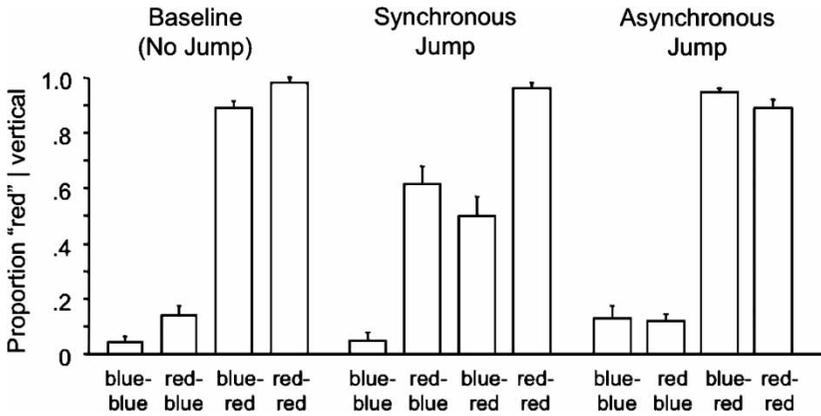


Figure 7. Experiment 4: Mean proportion of "red" reports when the shape was vertical, as a function of whether the shape of the report feature was stationary (baseline), was relocated with the change in the report colour (synchronous), or was relocated out of phase with the change in the report colour (asynchronous).

stimulus was actually blue in both frames (blue-blue), when it was red only in the first frame (red-blue), when it was red only in the second frame (blue-red), and when it was red in both frames (red-red). The main finding was the strong tendency to report the later feature value in the baseline condition and in the asynchronous jump condition but no such tendency in the synchronous jump condition (Figure 7).

This observation was supported by a repeated-measures ANOVA involving the factors of jump (none, synchronous, asynchronous) and condition (blue-blue, red-blue, blue-red, red-red). A significant Jump \times Condition interaction was observed, $F(6, 102) = 27.6$, $MSE = 0.026$, $p < .001$, indicating that the tendency to report the later appearing colour in asynchronous displays (the middle two bars in each panel of Figure 7) varied significantly with condition. Direct comparisons of the last-feature report in each of the three conditions were carried out using Fisher's LSD test. The results indicated a significant deviation between observed and expected reports (no bias = 0) in the predicted direction in the baseline condition (mean = .74), $t(102) = 13.30$, $p < .01$, and in the asynchronous jump condition (mean = .82), $t(102) = 22.80$, $p < .01$, but in the opposite direction in the synchronous jump condition (mean = -.11), $t(102) = 2.03$, $p < .05$.

Discussion

According to the object updating hypothesis, the difficulty of binding features with asynchronous temporal onsets and offsets arises because such

conditions favour object updating. Thus, providing a reliable nontemporal cue for assigning features to objects should greatly reduce the tendency for object updating to occur. In Experiment 4, otherwise very similar display sequences resulted in either a strong tendency to report the later of two report values paired with the defining feature, or no such tendency at all, depending on whether or not the report features that alternated over time occupied the same spatial locations. When they did occupy the same location (baseline and asynchronous jump), a strong last-feature report tendency was observed; when they occupied unique locations this tendency was actually reversed, with reports favouring the first of the two colours paired with yellow. The modulation of the last-feature report bias by this manipulation of a geometric cue is inconsistent with the idea that the bias results from better memory for the most-recently encountered feature pairing. The object updating account, on the other hand, explains this finding with its claim that the visual system treats the different features occurring in these two locations as corresponding to distinct objects. As such, in the presence of reliable cues for assigning new features to new objects, the visual system is less susceptible to the processes of object updating (i.e., feature replacement) that occur whenever it encounters dynamic scenes and instead a stable object representation in which the target feature is bound to the first of the two report features is formed.

GENERAL DISCUSSION

Attention has often been implicated in object perception, yet most of the work in this area has explored the role of spatial attention in conjoining features that occur in close proximity to one another (e.g., Treisman & Gelade, 1980). To successfully group features of an object that are changing over time, as occurs when an object is in motion, the visual system relies on the cooccurrence of feature appearance and disappearance in time (Lee & Blake, 1999; Sekuler & Bennett, 2001; Usher & Donnelly, 1998). The heuristic used by the visual system is that features that onset and offset together probably belong to the same object. Here we explored the consequence of disrupting this temporal cue by creating displays in which the features of objects onset and offset out of phase with one another, forcing attention to be switched from one feature to another in order to make a report of perceptual simultaneity.

According to the object updating hypothesis, asynchronous changes in feature values should promote object updating, similar to other situations in which spatial or temporal contiguity of features in dynamic visual displays promote object updating, as in the flash-lag effect (Moore & Enns, 2004), the attentional blink (Brehaut et al., 1999; Raymond, 2003), and backward

masking (Enns, 2004; Enns & Di Lollo, 1997). In support of this hypothesis, we found that when a defining feature value was paired with two report feature values of equal duration in a cycling display, participants tended to report the later of the two feature values. We argue that this occurred because the visual system must first segregate the defining feature value from other features in the display, which, in a continuously changing display, is an attention-demanding operation. To complete the report, attention then must be switched to the report feature value. Because this takes time, attention is momentarily unavailable for consolidating a stable object representation and object updating cannot be prevented; the value of the feature initially paired with the defining feature is updated, or replaced in the object representation by the more recent value. The consequence of this is that the later of the two report feature values paired with the defining feature is overrepresented when participants report on perceptual simultaneity in asynchronously changing displays.

The viability of the object updating account of this last-feature reporting tendency was tested in four experiments. In all experiments, the pattern of verbal reports with asynchronous displays did not mirror either the physical reality of these displays, nor did it reflect random guessing on the part of the participant. Instead, participants' reports conformed to a prediction based on the object updating hypothesis, which is that the most recently updated feature value is registered in the object representation. This same reporting tendency was not observed when synchronously changing visual features were paired (Experiments 1–4), even at high rates of alternation (Experiment 2). The tendency to report the second of two values paired with the defining feature in asynchronous displays was observed irrespective of which of feature was treated as the defining property and which as the report property (Experiment 1), and could not be attributed to differences in the length of time the defining feature and each report feature value were physically paired in synchronous and asynchronous displays (Experiment 2). The last-report tendency was also stronger with faster rates of alternation and reversed with very long feature durations (Experiment 2). Last-feature reports depended neither on the specific features to be paired (colour and motion in Experiment 1 vs. colour and orientation in Experiment 2), nor on whether the feature dimensions were the same or different (Experiments 3 and 4 vs. Experiments 1 and 2). Manipulations that increased the difficulty of switching attention from the defining to the report feature (Experiment 3) and decreased the availability of cues for assigning features to objects (Experiment 4) increased the tendency to report the later-appearing feature. Finally, providing reliable nontemporal cues for assigning features to objects (concurrency in colour and unique location cues) reversed last-feature reports in Experiment 4, even though the temporal cues for the onset and offset of the relevant features remained unreliable. The pattern of results

observed in all experiments is therefore broadly consistent with the predictions of the object updating account.

Taken together, the last-feature report tendency is likely a reflection of the difficulty participants have in forming a stable object representation under circumstances in which spatial and temporal cues provide conflicting information. Under normal circumstances, the temporal onset and offset of feature values are concurrent when they derive from the same object. The dark colour and the characteristic motion patterns of a bird both begin and end with the appearance and disappearance of the bird from view. Indeed, laboratory studies also confirm that temporal synchrony is a powerful factor influencing perceptual grouping, even when it is the only cue by which a visual display can be organized (Lee & Blake, 1999). The unusual circumstance that confronts the visual system in the task described here is that features are not beginning and ending concurrently with one another. This prompts the system to look for other clues to how the scene might be organized, such as grouping by spatial proximity and other geometric heuristics relevant to object perception. But, while this is going on, the scene continues to change, so the system becomes vulnerable to the feature “overwriting” that is normally beneficial as one views what first looks like a bird, and then becomes reinterpreted as an aeroplane, even if it only rarely goes on to become Superman.

Relation to the same-object cost

For many years attention researchers have been aware of a limit on object identification that is known as the two-object cost (Duncan, 1984). Namely, if participants try to report two features (e.g., colour and orientation), they are more accurate when both of these features belong to the same object than when they try to report the colour of one object and the orientation of another. The favoured interpretation of this finding is that attention is organized around “objects”, such that attending to an object activates awareness of all of its features. Reporting features of two different objects must therefore come about at the expense of sharing limited attentional resources between two objects.

Davis and colleagues have recently demonstrated that the two-object cost is not nearly as general as researchers have assumed (Davis & Holmes, 2005; Davis, Welch, Holmes, & Shepherd, 2001). Using novel solid shapes, rather than the overlearned, overlapping, outline drawings of most previous studies, these researchers have documented a strong same-object cost. Specifically, participants in their studies are less accurate in reporting two features of the same object than they are in reporting individual features from two separate objects. Clearly, more research will be required to fully

understand this dramatic reversal in the conventional wisdom regarding the conditions in which reporting on multiple features from the same object is easy or difficult. The results of the present study hint at the possibility that the last-feature reporting tendency could be used as an index of how tightly coupled the features of an object are, and it could therefore be used to predict the relative magnitude of the same-object cost in future studies.

Relation to the perceptual asynchrony effect

In the task examined in the current study, participants were required to pair two feature values that were alternating either synchronously or asynchronously over time. A similar task has been used to investigate the perceptual asynchrony effect (Moutoussis & Zeki, 1997a, 1997b), the finding that observers tend to report as coincident feature values whose onsets are actually about 50–100 ms out of phase with one another. The favoured interpretation of this finding (Arnold, Clifford, & Wenderoth, 2001; Clifford et al., 2003; Moutoussis & Zeki, 1997a, 1997b; but see Adams & Mamassian, 2004; Bedell, Chung, Ogmen, & Patel, 2003; Nishida & Johnston, 2002) is that awareness of motion or orientation lags awareness of colour, despite the fact that, at the neural level, motion and orientation are generally believed to be computed more rapidly than colour (Schmolsky et al., 1998).

Many of the findings of the present study are in broad agreement with results of studies of the perceptual asynchrony effect. That is, when displays were asynchronous, a lag of approximately 100 ms between feature changes resulted in a maximum number of last-feature reports (e.g., Experiment 2: The 135° phase shift condition with a feature duration of 128 ms corresponds to a lag of 96 ms; the 45° phase shift condition with a feature duration of 512 ms corresponds to a lag of 128 ms). However, numerous details of the present findings call for an account that goes beyond a simple lag in the processing or conscious awareness of visual features such as motion, colour, or orientation. First, factors other than temporal lag are needed to account for the variability in error rates for asynchronous conditions that all involved a lag of near 100 ms (i.e., the same lag led to last-feature reports that ranged from 68% to 95%). Most notably, the likelihood of reporting the later colour was strongly influenced by the context in which the pairing task was completed (i.e., whether the shortest feature duration tested in the experiment was 64 ms or 128 ms). Second, reversing the defining and report attributes reversed the direction of the effect for the same physical display sequences in Experiment 1 (i.e., last-colour reports occurred when upward motion was the defining property, whereas last-motion direction reports occurred when the colour red was the defining property), which has not been found in investigations of the

perceptual asynchrony effect (e.g., Clifford et al., 2003; Holcombe & Cavanagh, 2008). If observers generally become aware of colour before motion or orientation, as proposed by some theories of the perceptual asynchrony effect (Moutoussis & Zeki, 1997a, 1997b), then this reversal in our findings could never occur. Third, when the displays were synchronous (i.e., both features onset and offset together) participants were extremely accurate at pairing features that were actually concurrent, even at the shortest feature duration tested (as previously reported by Holcombe & Cavanagh, 2001); no lag was necessary for accurate pairing of colour and motion, or colour and orientation. Fourth, last-feature report biases were observed, and indeed exacerbated, even when the pairing task involved the same attribute (colour) as both the defining and report feature (as in Experiments 3 and 4) rather than two different attributes (as in Experiments 1 and 2).

The differences between our findings and those for the perceptual asynchrony effect may be reconciled in at least two different ways in future studies. One possibility is that the task used to measure the perceptual asynchrony effect, despite its superficial similarity to the task used in the present studies, indexes a different set of psychological processes. This possibility is implied by some authors who have reported that the perceptual asynchrony effect is unaffected by reversing the role of defining and report features (Arnold, 2005; Clifford et al., 2003; Holcombe & Cavanagh, 2008). A second possibility is that the two tasks index the same set of psychological processes, but emphasize different characteristics of the feature binding operation. That is, there may indeed be a variable lag time needed for different feature values to reach awareness, but in addition to this lag, perceptual reports may also be subject to a delay imposed by switching attention from a defining feature to a report feature. Moreover, the attention switching delay may be especially sensitive to the specific factors manipulated in the present study (including the present conditions of extreme temporal asynchrony), leaving the differential feature lag as a constant and relatively untouched factor in all of our experiments. We leave these theoretical possibilities as problems to be resolved by future studies. What the present set of experiments make clear, however, is that temporal synchrony seems to be as important as spatial and geometric cues (i.e., shared locations, shapes, and colours) for solving the binding problem in the perception of objects whose features are changing over time.

Limitations and future directions

The present findings highlight the importance of understanding how the spatial and temporal allocation of attention interacts in object representa-

tion. We note that in our pairing task, and in similar tasks (e.g., Moutoussis & Zeki, 1997a, 1997b), participants are asked to pair asynchronous features in displays that contain more than one object, such that across the entire display the probability of any pair of features (e.g., red paired with upward motion) is equated. That is, at any point in time, for every occurrence in the display in which red is paired with upward motion, there is a corresponding occurrence in which blue is paired with downward motion (e.g., Figure 1C). Implicit in the use of such multi-element displays is the assumption that focused spatial attention to the location of one object alternating in feature pairings would invoke different processes than distributed spatial attention across the entire display of objects. Of course, there is nothing to prevent participants from focusing attention to a single object. Holcombe and Cavanagh (2008) found that the frequently reported lag in perceiving motion relative to colour was greatly reduced when the target was spatially cued, suggesting that spatial attention may indeed be important in binding features with asynchronous temporal onsets and offsets. However, in their experiments participants also viewed multielement displays, so it remains to be seen whether the same result would obtain if spatial attention to the target were encouraged by including single-target displays, rather than by explicitly cueing participants to attend to a single item.

CONCLUDING REMARKS

A considerable literature has now been established exploring the processes that allow the visual system to select and bind features belonging to the same object in static displays (e.g., Treisman & Gelade, 1980). Increasingly, interest is turning to the question of how a stable object representation is consolidated when features of the object appear in dynamic displays, constantly changing over time. Our particular interest in this paper was to examine the artificial situation in which features belonging to the same object onset and offset at different points in time. Our motivation for examining such displays was to investigate the processes of object representation in the absence of this particularly powerful cue for grouping features into objects. We hypothesized that such situations would tax attentional resources, requiring a time-consuming attention shift from one feature to another; without the availability of attention for consolidating a stable object representation, object updating would occur. Based on comparisons with other phenomena in which object updating appears to play a role, we anticipated that pairing reports would tend to emphasize the most recently experienced pairing of features in the display, but that this tendency would disappear if other, reliable cues were reintroduced. These expectations were confirmed, emphasizing

once again the ubiquity of object updating as a general heuristic of the visual system when faced with unreliable temporal cues for assigning features to object representations.

REFERENCES

- Adams, W. J., & Mamassian, P. (2004). The effects of task and saliency on latencies for colour and motion processing. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *271*, 139–146.
- Arnold, D. H. (2005). Perceptual pairing of colour and motion. *Vision Research*, *45*, 3015–3026.
- Arnold, D. H., Clifford, C. W. G., & Wenderoth, P. (2001). Asynchronous processing in vision: Color leads motion. *Current Biology*, *11*, 596–600.
- Baldo, M. V. C., Kihara, A. H., Namba, J., & Klein, S. A. (2002). Evidence for an attentional component of the perceptual misalignment between moving and flashing stimuli. *Perception*, *31*, 17–30.
- Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature*, *378*, 565–566.
- Bedell, H. E., Chung, S. T. L., Ogmen, H., & Patel, S. S. (2003). Color and motion: Which is the tortoise and which is the hare? *Vision Research*, *43*, 2403–2412.
- Bennett, J. D., Lleras, A., Oriet, C., & Enns, J. T. (2007). A negative compatibility effect in priming of emotional faces. *Psychonomic Bulletin and Review*, *14*, 908–912.
- Brehaut, J., Enns, J. T., & Di Lollo, V. (1999). Visual masking plays two roles in the attentional blink. *Perception and Psychophysics*, *61*, 1436–1448.
- Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.
- Bridgeman, B. (1980). Temporal response characteristics of cells in monkey striate cortex measured with metacontrast masking and brightness discrimination. *Brain Research*, *196*, 347–364.
- Chappell, M., Hine, T. J., Acworth, C., & Hardwick, D. (2006). Attention “capture” by the flash-lag flash. *Vision Research*, *46*, 3205–3213.
- Clifford, C. W. G., Arnold, D. H., & Pearson, J. (2003). A paradox of temporal perception revealed by a stimulus oscillating in colour and orientation. *Vision Research*, *43*, 2245–2253.
- Davis, G., & Holmes, A. (2005). Reversal of object-based benefits in visual attention. *Visual Cognition*, *12*, 817–846.
- Davis, G., Welch, V. L., Holmes, A., & Shepherd, A. (2001). Can attention select only a fixed number of objects at a time? *Perception*, *30*, 1227–1248.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*, 481–507.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*, 501–517.
- Duncan, J. (1985). Visual search and visual attention. In M. I. Posner & O. Marin (Eds.), *Attention and performance XI* (pp. 85–106). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Egley, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*, 161–177.
- Elliot, M. A., Shi, Z., & Kelly, S. D. (2006). A moment to reflect upon perceptual synchrony. *Journal of Cognitive Neuroscience*, *18*, 1663–1665.

- Engel, A. K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and Cognition*, 8, 128–151.
- Enns, J. T. (2002). Visual binding in the standing wave illusion. *Psychonomic Bulletin and Review*, 9, 489–496.
- Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research*, 44, 1321–1331.
- Enns, J. T., Brehaut, J., & Shore, D. I. (1999). The duration of a brief event in the mind's eye. *Journal of General Psychology*, 126, 355–372.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8, 135–139.
- Enns, J. T., Lleras, A., & Moore, C. M. (2010). Object updating: A force for perceptual continuity and scene stability in human vision. In R. Nijhawan (Ed.), *Problems of space and time in perception and action* (pp. 503–520). Cambridge, UK: Cambridge University Press.
- Fehrer, E., & Raab, D. (1962). Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, 63, 143–147.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Lawrence Erlbaum Associates, Inc.
- Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences*, 98, 4271–4275.
- Holcombe, A. O., & Cavanagh, P. (2001). A purely temporal transparency mechanism in the visual system. *Perception*, 30, 1311–1320.
- Holcombe, A. O., & Cavanagh, P. (2008). Independent, synchronous access to color and motion features. *Cognition*, 107, 552–580.
- Hommuk, K., & Bachmann, T. (2009). Temporal limitations in the effective binding of attended target attributes in the mutual masking of visual objects. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 648–660.
- Jiang, Y., & Chun, M. M. (2001). The spatial gradient of visual masking by object substitution. *Vision Research*, 41, 3121–3131.
- Kanwisher, N., & Driver, J. (1992). Objects, attributes, and visual attention: Which, what, and where. *Current Directions in Psychological Science*, 1, 26–31.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10, 494–501.
- Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23, 571–579.
- Lee, S. H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science*, 284, 1165–1168.
- Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. *Journal of Experimental Psychology: General*, 133, 475–493.
- Lleras, A., & Moore, C. M. (2003). When the target becomes a mask: Using apparent motion to isolate the object component of object-substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 106–120.
- Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychological Science*, 15, 866–871.
- Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1171–1180.
- Moore, C. M., Mordkoff, J. T., & Enns, J. T. (2007). Path of least persistence: Object status mediates visual updating. *Vision Research*, 47, 1624–1630.
- Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 264, 393–399.

- Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 264, 1407–1414.
- Namba, J., & Baldo, M. V. C. (2004). The modulation of the flash-lag effect by voluntary attention. *Perception*, 33, 621–631.
- Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, 370, 256–257.
- Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Current Biology*, 12, 359–368.
- Palanca, B. J. A., & DeAngelis, G. C. (2005). Does neuronal synchrony underlie visual feature grouping? *Neuron*, 46, 333–346.
- Raymond, J. E. (2003). New objects, not new features, trigger the attentional blink. *Psychological Science*, 14, 54–59.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849–860.
- Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, 4, 93–102.
- Sarich, D., Chappell, M., & Burgess, C. (2007). Dividing attention in the flash-lag illusion. *Vision Research*, 47, 544–547.
- Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology*, 79, 3272–3278.
- Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychological Science*, 12, 437–444.
- Shore, D. I., & Spence, C. (2005). Prior entry. In L. Itti, G. Rees, & J. K. Tostós (Eds.), *Neurobiology of attention* (pp. 89–95). New York: Elsevier Academic.
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555–586.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6(2), 171–178.
- Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature*, 394, 179–182.
- Vreven, D., & Verghese, P. (2005). Predictability and the dynamics of position processing in the flash-lag effect. *Perception*, 34, 31–44.
- Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, 238, 778–780.
- Wilson, A. E., & Johnson, R. M. (1985). Transposition in backward masking: The case of the travelling gap. *Vision Research*, 25, 283–288.
- Wolfe, J., Friedman-Hill, S. R., & Bilsky, A. B. (1994). Parallel processing of part-whole information in visual search tasks. *Perception and Psychophysics*, 55, 537–550.

Manuscript received July 2009

Manuscript accepted March 2010

First published online June 2010