P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

Part V

Space–time and awareness

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

28

Object updating: a force for perceptual continuity and scene stability in human vision

JAMES T. ENNS, ALEJANDRO LLERAS, AND CATHLEEN M. MOORE

Summary

How does the visual system provide us with the perception of a continuous and stable world in the face of the spatial–temporal chaos that characterizes its input? In this chapter we summarize several programs of research that all point to a solution we refer to as *object updating*. We use this phrase because perceptual continuity seems to occur at an *object* level (as opposed to an image level or a higher conceptual level) and because our research suggests that the visual system makes a sharp distinction between the formation of new object representations versus the *updating* of existing object representations. We summarize the research that led us to this view in the areas of masking by object substitution, the flash-lag illusion, response priming, and an illusion of perceptual asynchrony.

28.1 Introduction

Biological vision is the marvelous ability of an organism to be informed about its surroundings at a distance and with a high degree of spatial and temporal resolution. This ability allows us to know where things are, what shape and color they are, and equally importantly, *when* they are there, so that we may interact with them appropriately. Yet, contrary to many people's implicit understanding of how biological vision is accomplished, it is *not* a process by which light, reflected from surfaces in the three-dimensional world, is recorded faithfully by the brain in order to reconstruct the nature of the surfaces that gave rise to the recorded pattern of light. To a layperson, this may seem like a plausible place to begin, but modern vision science has pointed to numerous reasons why this approach is a nonstarter. One reason is the computational complexity of the problem of inverse optics (reconstructing the three-dimensional world from a two-dimensional pattern of light on the retina). This has been shown to be an underdetermined problem, in that the time needed to solve it far exceeds the time limits under which biological organisms must act in order to survive (Tsotsos 1990). This forces visual systems to be highly selective in their processing of light.

A second problem with this naïve view is that the pattern of light falling on the retina is highly discontinuous, across both space and time, for many reasons. One is that the retina does not register light in a uniform way across its surface, being, for example, more sensitive

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

to color at the center and more sensitive to motion in the periphery. Second, there is only a relatively small spatial window in which the eye has a high degree of spatial resolution. As such, many eye movements are needed in order to "see" even the simplest of scenes, with the duration and order of these eye movements being extremely difficult to predict (Henderson & Ferreira 2004). But these biological solutions to accomplishing vision (i.e., the need for processing selectivity and discrete eye movements) also present scientists with a new set of problems to understand. For example, how does the visual brain create a sense of timely order out of what seems at first glance to be the spatial and temporal chaos of the visual input to the brain?

In this chapter we highlight one of the general solutions that we believe human vision has settled on in order to establish perceptual continuity and scene stability in the face of discontinuous inputs. We refer to this solution as *object updating* based on our findings that perceptual continuity seems to occur at the *object* level (as opposed to the image level or a higher conceptual level). This solution involves a sharp distinction between the formation of new object representations versus the *updating* of existing object representations. In what follows we present research that led us to this view.

28.2 Object updating

Any visual system that samples information more-or-less continually is faced with a fundamental problem. How should the system incorporate newly sampled information into representations it has formed from past samples? One possibility is that updating occurs via a point-for-point image-comparison process, whereby each lowest-level unit in the representation (retinal cell or pixel) is updated independently. Such a mechanism could be easily implemented as a parallel process, but it would be blind to any meaning in the scene, such as knowing which objects are present, where objects are in relation to each other, and whether any objects are moving.

An alternative possibility is that representations are updated through an object-based process, such that the meaningful units in the scene (objects) are taken into account, and changes are made only insofar as they occur to an object already represented. This distinction between image-based and object-based updating is analogous to differences between pixel-based (painting) and object-based (drawing) programs for computer graphics. In an image-based program, editing something on one object can inadvertently alter another object. For drawing programs, on the other hand, objects are selected and edited independently. Objects other than the one currently selected are protected from changes that are made to the selected object, even if the two objects overlap each other in the image space.

As already noted, image-based updating is appealing from a computational perspective because it can be implemented easily within models that embody the retinotopic registration of information in different visual brain areas. But this appeal must be weighed against the considerable costs it incurs for later, higher-order processing. Because image updating has no regard for the meaning of the scene in terms of surfaces and objects, many important distinctions are lost with each resampling cycle. Image-based updating would fail, for

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

example, to maintain region assignments to figure versus ground, edge assignments to luminance change versus surface orientation, as well as associations between discontinuous regions of a surface caused by occlusion. Given the importance of organized representations for disambiguating the retinal image, it is our view that image-based updating cannot be a complete solution to the problem of perceptual continuity.

This consideration of the difficulties encountered by the image-based approach makes object updating worth considering. The general idea is that, if currently sampled information is interpreted as deriving from an object that is already represented in the scene, then that information will be used to update that object representation. In contrast, if the new information is interpreted as deriving from a different object, then the original object representation will be spared from updating and will therefore remain unchanged in the face of new sensory information. Finally, if the information is perceived as deriving from a new object, then it may elicit the establishment of a new object representation in the scene. Notice that in any given sampling cycle, an old object could be in a new location and a new object could be in a location where an old object had been before. In this way, object-mediated updating is dissociable from image-based updating.

Our ideas on object updating first emerged from research on a type of backward masking that has come to be referred to as masking by *object substitution*, a term introduced by Enns and Di Lollo (1997). However, because the theoretical ideas behind this term are the focus of this chapter, we will begin with a less theory-laden description of the “four-dot masking” that led to these ideas. Backward visual masking, at its most general, refers to the observation that the presentation of a later shape (the *mask*) disrupts the processing of an earlier shape (the *target*) that would have been completely visible if presented in isolation. When we first introduced four-dot masking, the prevailing view of how backward masking is accomplished was that the contours of the mask (a) were confused with the contours of the target in early visual representations (integration), (b) interrupted the ongoing processing of the target contours (interruption), or (c) inhibited the emerging contours of the target (inhibition). However, the details of the four-dot masking procedure caused us to rethink how backward masking might be accomplished.

In a typical four-dot masking experiment (Di Lollo et al. 2000; Enns 2004) a display of shapes is presented (e.g., diamonds with missing corners, circles with gaps in one of four locations, or a variety of simple shapes) and one of the shapes is surrounded by four small dots (sometimes each dot is only one pixel in size). It is the participant’s task to identify the shape surrounded by these four dots. When these shapes and the four dots flash on and off simultaneously, it is a relatively easy task for the participant to identify the shape indicated by the dots. However, when only the four dots linger on the screen after the shapes have been erased, the shape surrounded by the dots becomes very difficult (and sometimes even impossible) to identify. It seems as though only the four dots were ever presented in that location (for a demonstration see www.sfu.ca/∼enzo/).

This surprising finding that a sparse pattern of only four dots surrounding the target are effective as a backward mask led us to rethink the prevailing theoretical accounts of backward masking. For one thing, no contour-based interference of any kind, whether

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

resulting from integration, interruption, or inhibition, seemed to be at issue because the contours of the dots were so small in comparison to the contours of the target they were masking, and so great was the distance of the mask contours from those of the target. A traditional understanding of contour interactions in early visual processing predicted that no contour effects should occur under these conditions (Di Lollo et al. 2000). Second, the fact that backward masking occurred even though the mask was visible from the onset of the target display was inconsistent with the then-understood temporal dynamics of masking. Theories of contour interruption and inhibition predicted that masking should be strongest when there was a critical delay between the emerging neural signal of the target and the later arriving neural signal corresponding to the mask (Enns & Di Lollo 2000).

But if contour interactions and temporal asynchronies in the neural signals associated with targets and masks are not responsible for dot masking, then what is? The hypothesis we entertained was that the lingering appearance of the four dots in the target display resulted in the formation of a different object representation for that location in the scene. Initially, the object representation for that location began with the inclusion of the contours associated with both the shape and the four dots. However, before that representation could be completely formed (at least completely enough to lead to a positive identification of the shape) its location in the display contained only the four dots. Thus, the four dots came to be seen by the visual system as an updated version of the object representation that was initiated by the appearance of the target shape and the dots. The continuity of the four dots over time, in comparison to the brief appearance of the target contours without their later confirmation, was taken as evidence by the visual system that only the four dots had ever existed in that location.

In this view of object updating as the mechanism of backward masking, the simultaneous onset of the mask and target elicits an initial grouped representation of their contours. When the target shape disappears before a stable representation of it is formed, but the mask lingers, the scene comes to be represented as the original object changing shape. Information that is then sampled from the lingering display containing only the mask is used to update the original representation. However, because it is the original representation that is needed to answer the question posed by the experimenter (i.e., what is the identity of the target?) this updating process results in poor performance. The relevant information for the psychophysical task is simply no longer included in the updated representation, presumably because this is a design feature of a biological visual system that updates information with a strong bias for object continuity. In keeping with this account, we also noted that attending to the location of the target, either previous to its appearance or rapidly on its presentation, facilitates the establishment of a stable representation (Di Lollo et al. 2000; Enns 2004). If a stable representation of the target can be established based on the brief information about it prior to its offset, then information derived from the lingering mask has less of an opportunity to disrupt critical target information.

The initial evidence that prompted us to think of four-dot masking in terms of object updating was that the mask was so sparse. If only contours were involved, then the target should have had plenty of room to shine through the space in between the widely separated

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

dots in the mask and thereby avoid any masking. But the masking we observed was also not at a purely abstract and conceptual level either. We noted that four dots that were not centered on the location of the target were quite ineffective in masking the target, as for example in the case of four dots that only underlined or lay beside the location of the target (Enns & Di Lollo 1997). Clearly, this form of masking was sensitive to some aspects of space, implying that it was not occurring because of a more general interruption of processing. However, its space sensitivity was also not as precise as individual contours either, because masking was quite effective when the four masking dots that began in the same location as the target lingered in an adjacent location eccentric to the center of gaze (Jiang & Chun 2001). Another clue to the object-based nature of the masking was that the strength of masking increased with the number of potential target objects in the display (Di Lollo et al. 2004). Masking strength also decreased if the location of the target or the identity of the target was cued in advance (Enns 2004), in keeping with the position that there is a limited capacity to represent objects in visual short-term memory (Vogel et al. 2001).

Lleras and Moore (2003) provided the first direct test of the hypothesis that four-dot masking was object-mediated. Their experiments were designed to test whether an object-level component of four-dot masking could be isolated from any spatially local interference that might also be contributing to it. Recall that four-dot masking occurs when the mask lingers on the screen for some time following the offset of the target. Under the object-updating account, the important aspect of the delay is that the object that corresponds to the target is delayed in the emerging mental representation of the scene, not that a masking object is physically delayed at the same location as the target. With this in mind, Lleras and Moore used apparent motion to delay the mask as an object in the scene without delaying it at the location where the target object appeared. In the critical condition, a given mask-target pair offset at the same time, but a short while later, the mask reappeared at a new location. This gave rise to the perception of a single object, moving and changing in shape over time. Thus, at the level of object representations of the dynamic scene, this was a delayed-offset condition because the object that started as the mask-target pair and changed to just the mask was delayed in the scene. At the local or image level, however, it was a simultaneous-offset condition because the mask and target offset at the same time at the same location. Yet, despite this simultaneous offset, substantial masking occurred.

In contrast, no significant masking occurred in a long-interval condition in which the same transient events as those in the short-interval condition occurred, but the interval was now too long to support apparent motion. What happened here instead was that the display was perceived as one object suddenly appearing at one location, turning off, and then a new object appearing at a new location. In the strongest instantiation of this research logic, the mask was only a single dot that never appeared at the location of the target, and yet it still produced significant masking (Lleras & Moore 2003, Experiment 4).

Moore and Lleras (2005) reported additional direct evidence that four-dot masking involved object-mediated updating. They demonstrated that targets could be protected from four-dot masking by introducing manipulations that facilitate the early establishment

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

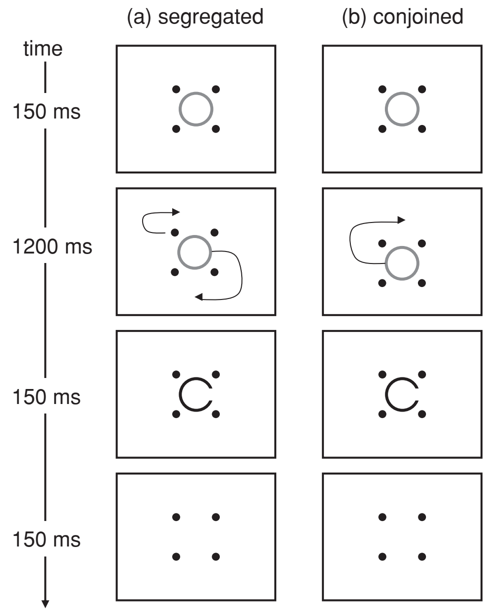

Fig. 28.1 An illustration of the display sequence for one of the items in Moore and Lleras (2005). (a) In the segregated condition the four dots and the circle moved independently of one another before the masking sequence. (b) In the conjoined condition the four dots and the circle moved in concert with one another.

of separate representations for the target and the mask. In other words, if the visual system treats the mask as a different object from the target, then the information sampled from a lingering mask will not be used to update the target representation, and whatever target information was established will be spared from the updating process.

Figure 28.1 illustrates an example of how this logic was applied, using segregation by common motion to manipulate the degree to which masks and targets were represented as single or separate objects. The figure shows only one of the many items in each display. The participant's task was to find the one dark circle-with-a-gap among the many light circles-with-gaps and to report the direction in which that target circle's gap pointed (up, down, left, right). Each trial began, however, with placeholders that were circles surrounded by four dots. There were no gaps in the placeholder circles, and they were all the identical shade of gray. Therefore there was no indication as to which item would eventually be the target, and no information about the identity (i.e., gap direction) of any of the circles. The displays did, however, allow for the establishment of initial object representations for the circles and dot masks. Before displaying the target information, a short movie was shown. In the *segregated* condition, the movie was of the circles and masks jiggling around their center point independently. In the *conjoined* condition, the circles and masks jiggled around their center point together as a unit. The movie ended in both conditions, with the circles and masks recentered within each pair, and the trial then unfolded as a usual dot-masking

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

trial. Thus, the presentation of the target information was identical in the two conditions; they differed only in the object history with which the viewer approached that information. Consistent with the object-updating hypothesis, that history influenced dot masking, such that masking was eliminated in the *segregated*, but not in the *conjoined,* condition. Our interpretation is that by establishing an object representation for the target distinct from the mask early in processing, the target was protected from updating, and therefore from four-dot masking. Several different strategies of facilitating the establishment of separate object representations for the target and mask early in processing render the same effect (Moore & Lleras 2005).

Together, these studies provide a dissociation between the effects of image location and object representation in four-dot masking that clearly implicates the object as the vehicle of interference. In particular, as long as the mask is represented as the same object as the target, interference can occur even when the mask never appears at the location of the target (Lleras & Moore 2003). At the same time, if the mask is represented as a different object from the target, the interference does not occur even when the mask lingers at the location of the target (Moore & Lleras 2005).

28.3 Object updating and the flash-lag illusion

The success we had in understanding backward masking in terms of object updating prompted us to examine several other visual phenomena in which errors that participants make are consistent with the effects of having lost earlier information about an object in favor of updated information. We refer to this class of errors as those of *perceptual lag*, because the incorrect responses given by participants are not random guesses. Rather, they can be linked directly to a competition between features that were presented first for some object versus features that were presented more recently, with the errors showing a strong bias for the features presented most recently.

An interesting phenomenon of this kind is the *flash-lag illusion* that involves a misperception of the spatial relations between a moving object and a briefly flashed one. This illusion is of particular interest to us because it seems to illustrate the object-updating process in one of the most basic of perceptual experiences, the seen position of an object in space while it is undergoing motion.

A typical version of the illusion is illustrated in Fig. 28.2. A disc travels in a circular path and at some point in its journey a square is flashed briefly. If the disc continues on its path, past the frame in which it was aligned on a radial axis with the square, then the moving disc tends to be seen as spatially “ahead” of the briefly flashed stationary square, even though when the square appeared, the objects were aligned (Mackay 1958; Nijhawan 1994). But if instead of continuing on its motion path, the moving disc is erased immediately after it is aligned with the square, then the disc and the square are seen accurately as aligned (Whitney & Murakami 1998; Eagleman & Sejnowski 2000; Whitney et al. 2000).

Moore and Enns (2004) examined whether the different perceptions associated with these two conditions, which they referred to as *continued-motion* versus *stopped-motion*, could

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

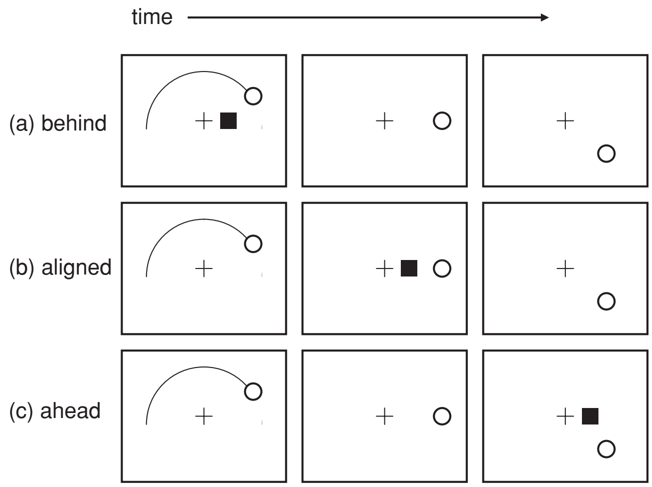

Fig. 28.2 Illustration of the displays used in Moore and Enns (2004). A moving disk (circle) moved around the circumference of an imaginary circle. At some point in this journey a flash occurred (square), and it was the participant's task to indicate whether the moving circle was (a) behind, (b) aligned, or (c) ahead of the square at the moment it flashed.

be understood in terms of object updating. (Other authors in this volume refer to these conditions as "continuous cycle" and "flash-terminated," respectively). Our guiding hypothesis was that the difference in perception between these two conditions hinges directly on the object-updating process. When the disc moves to a new position following the flash in the continued-motion condition, the new position information replaces that acquired at the time of the flash. In contrast, when there is no new position following the flash in the stopped-motion condition, there is no new position information to update the previous information, and so the alignment of the two objects is perceived accurately. Demonstrations of the conditions tested in this study can be found online: www.psych.ubc.ca/∼ennslab/research.

Moore and Enns (2004) provided direct evidence for this interpretation of the flash-lag illusion in three steps. In a first phase of the study we confirmed that the continued- and stopped-motion conditions indeed led to a very different set of perceptual reports. A large flash-lag illusion was recorded in the continued-motion condition and no illusion was observed in the stopped-motion condition. In the second phase, changes were made to the displays such that on some trials participants saw the moving disc as undergoing an abrupt and large change in either size or color at the moment of the flash. These changes were intended to disrupt the normal process of object continuity. Our reasoning was that if an object was seen to change in a radical way then a new object representation would have to be formed for it after the change occurred. Consistent with our prediction, under these conditions the flash-lag illusion was no longer experienced by participants, even though the motion was as continuous as it was before. The change in size or color of the moving disc was enough to disrupt the experience of perceptual continuity for that object.

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

In phase three of the study we tested a unique prediction of the object-updating account of the flash-lag effect. If the disruption in perceptual continuity caused by the large changes in size or color actually leads to the formation of a new object representation, then we should be able to find direct evidence for this new representation in the form of a change in the total number of objects perceived in the no-change (standard flash-lag) versus change (disrupted size or color) conditions. Specifically, two discs should be visible at the moment of the flash in the changed condition but not in the unchanged condition. This was tested by asking participants to report on the number of discs that were visible at the time of the flash. The possibility of illusory reports of more than one disc was anchored to reality by including trials in which there actually were two discs presented. The results showed that the change in size or color of the moving disc led to the perception of "double discs," with one corresponding to the original moving disc and the other corresponding to the suddenly changed disc. Note that this is consistent with the flash-lag illusion actually being a form of backward masking, as described in the previous section, where the spatial position of the disc seen after the flashed square is automatically and irretrievably updated so that the actual position at the time of the flash is lost. In the case of the large change in size or color, the updated spatial position with regard to the original disc is not assigned to the suddenly changed disc. Thus, the changed disc is spared from the normal process of object updating and its spatial position can be perceived accurately.

In conclusion to this section, we want to note that our object-updating account is not at odds with previous accounts of flash-lag illusion, which have focused either on visual integration over a temporal window (Eagleman & Sejnowski 2000; Krekelberg & Lappe 2000), visual extrapolation into the future (Nijhawan 1994, 2002; Schlag & Schlag-Rey 2002), or the relative speed of neural signals from moving and static stimuli (Whitney & Murakami 1998; Whitney et al. 2000; Whitney 2002). Those theories have all offered various mechanisms to account for the illusory percept and yet they have remained silent on the level of representation that is involved in the illusion. As such, we propose our object-updating account of the illusion to add a critical level of detail to the mechanisms involved in each of these theories. The object-updating account contributes to an understanding of the illusion at a level of analysis that has so far been ignored, namely, the critical role of perceptual objects in flash-lag sequences.

28.4 Object updating and response priming

So far we have discussed how object updating influences the representations that form the basis of our conscious awareness. However, we do not only use vision to help us identify objects, we also use vision to help us interact with those objects through actions. That is, we sometimes grasp objects, catch them, navigate around them, or simply point to them. There is now a large body of evidence that visual information is processed by two distinct and somewhat independent neurological subsystems in the primate cortex: a ventral pathway that is specialized for the conscious perception of objects and scenes, and a dorsal pathway that is specialized for visually guided action (Milner & Goodale 1995; Goodale &

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

Humphreys 1998). From this perspective, our previous findings on object updating can be characterized as vision solely for conscious perception. We therefore began to ask whether object updating is also relevant for the representations that inform the subsystem responsible for visually guided action.

Masked priming is a behavioral tool used by researchers to better understand how visual processing influences motor responses. The method is simple, consisting of the presentation of three shapes in rapid succession: a prime, followed by a mask, which is itself followed by a target. The participant is instructed to respond to the target as rapidly as possible without making too many errors, usually by making a left or a right key press. The purpose of the prime is to provide advance information about the upcoming target and therefore the response that will be required. The typical finding is that when the prime specifies the same response as the target (compatible trial), then response time to the target is faster and more accurate than when the prime specifies the opposite response from the target (incompatible trial). This occurs even under conditions in which the prime shape is not visible. This result is taken as evidence that the motor system starts preparing its response even before conscious perception occurs, and the magnitude of this preparation is indexed by the difference between incompatible and compatible responses.

The role of the mask in traditional studies of masked priming has been to limit the strength or duration of the perceptual representation of the prime. This is done by varying the period of time that the prime may be processed prior to the appearance of the mask and/or by varying the intensity of the mask (the number and contrast of its contours along with its duration). In other words, masking is used as a tool of convenience to limit the exposure of the visual system to the prime shape. However, this intuitive and widely accepted use of masked priming fails to acknowledge that mask shapes do not merely decrease the visibility of a prime shape. In fact, as we have seen in previous sections, masks can fundamentally alter the way in which a prime is processed.

Lleras and Enns (2004) provided a striking demonstration of just how much the choice of a mask can alter the processes under investigation in masked priming. We began this study when we learned about the so-called negative compatibility effect (NCE), first described by Eimer & Schlaghecken (1998). As suggested by the word *negative* in the name, the NCE is a masked priming effect in which *compatible* primes actually lead to slower and less accurate responses than *incompatible* primes. This counterintuitive finding has been studied extensively, and several theories have been put forward, but what struck us was that the masks used in previous studies shared a commonality: the mask was composed of the same visual features that were used to create the prime and target shapes. For example, when the prime and target were each comprised of arrows (pointing either left or right), the mask consisted of a single image obtained by superimposing the same left- and right-pointing arrows. When these masks followed the primes, thereby reducing their visibility, negative priming was obtained. However, when these masks were omitted, and the primes were perfectly visible, then the positive priming effect obtained in most previous studies of masked priming was observed (Eimer & Schlaghecken 1998; Klapp & Hinkley 2002).

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

To be sure, the shared similarity between the primes and these masks suggested that these masks should be very effective in their reduction of the participant's ability to respond to the primes when specifically asked to do so, which they were. But we suspected that these masks might be doing something else besides simply limiting the visibility of the prime shapes.

Our past experience with the concepts of object updating led us to test a very different hypothesis regarding the critical ingredient for the NCE (Lleras & Enns 2004). It was prompted by our observation that the display sequences used in the NCE were similar to those in which we had previously studied object updating in masking and the flash-lag effect. Specifically, prime and mask shapes were presented close together in both time (intervals between separate shapes were on the order of 50–150 ms) and space (shapes were presented at the same or in nearby locations). Moreover, the shared features among the prime-mask-target shapes might readily induce the visual system to interpret these sequences not as discrete and formally unrelated events (as they were interpreted by the experimenters) but as instantiations of the same object changing rapidly over time. If so, even though the mask shape might be formally neutral, in terms of the information it provided with regard to the upcoming target and its required response, the mask might be informative to a visual-motor system that was rapidly incorporating new information about an emerging object representation.

To illustrate our thinking, consider the case of a right-pointing arrow prime followed by a mask consisting of superimposed left and right arrows, and this mask itself being followed by a right-pointing target. When the right-pointing prime first appears, the visual system may begin to form an object representation (let's call it P), such that the attributes of the prime will be encoded and linked to P. If some of these prime attributes are strongly associated with motor responses, even in the very early stages of their formation, this representation will get linked to its associated response and give rise to corresponding motor preparation. In this case, preparation for a right response will begin. When the mask shape appears, especially if it shares task-relevant features with the prime, it is likely to be interpreted as a new and updated instantiation of P. Therefore, the mask will not receive its own object representation but will instead be incorporated into the existing representation of P that has already been initiated. This means that task-relevant information now detected in the mask that was not already present in P will be added to it. In this example, the new features correspond to a left-pointing arrow, the right-pointing features having already been encoded.

This kind of thinking led to our hypothesis that in the NCE, priming is determined by the most recently detected set of response-relevant features in a representation that has been updated as part of the ongoing task of the participant to respond rapidly to the target object. We tested this hypothesis in a series of experiments by comparing the priming effects obtained from masks comprised of either task-relevant or task-irrelevant visual features. Fig. 28.3(a) shows a typical sequence of trial events and Fig. 28.3(b) shows a summary of the priming effects observed under the masking conditions we tested. As can readily

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

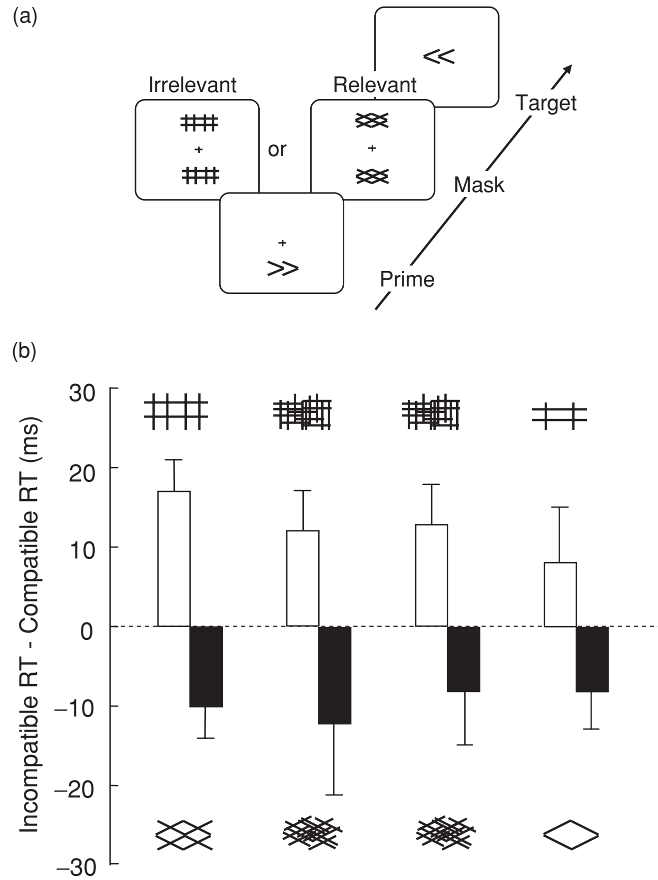

Fig. 28.3 (a) A typical sequence of trial events in an experiment on the negative compatibility effect. (b) Mean priming effects (incompatible minus compatible response time) as a function of various masks inserted between the prime and the target arrows in Lleras and Enns (2004). The specific masks used are shown above and below each data point. Positive priming effects (white bars) occurred for masks that had no features relevant to the target discrimination task. Negative priming effects (black bars) occurred when masks contained features that were relevant to the target discrimination task.

be seen, masks consisting of task-relevant features invariably produced negative priming effects, replicating previous studies of the NCE, whereas irrelevant masks systematically yielded positive priming effects, as expected based on our views of object updating.

These results directly refute theories of masked priming in which the critical ingredient is the strength of the representation corresponding to the prime, which is usually assessed by measuring the visibility of the prime shapes. In such theories, all masks are equal provided that they reduce the visibility of the prime sufficiently to activate the unconscious processes associated with the subsequent target. However, this was not what we found. Rather, we

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

found that completely opposite priming effects could be obtained at equal levels of high- or low-prime visibility. What determined the direction of the priming in both cases was whether the mask contained task-relevant (negative priming) or task-irrelevant features (positive priming).

The alternative theory we proposed to account for the NCE was based on our interpretation of backward masking and the flash-lag, namely, that the prime and mask may be interpreted by the visual system as a change in the same object over time. As a result, the prime–mask sequence is susceptible to the normal processes of object updating, and a single representation may represent both stimuli as one object changing over time. According to this interpretation, although the observed priming effect may superficially appear to be negative with respect to the prime, it actually corresponds to a positive priming effect induced by the novel task-relevant features in the mask. In sum, the observed priming effect reflects the influence not of the prime alone, but of the prime–mask bundle, on response selection. On irrelevant-mask trials, the most recent set of task-relevant features encoded prior to target onset are those present in the prime because the mask has none of these features. In this case then, the observed priming effect really does reflect the positive influence of the prime on response selection to the target.

Lleras and Enns (2004) also showed that this object-updating account could be applied to masks that did not contain exact replicas of the primes but also to randomly oriented line masks that merely contain similar features in spatially noncorresponding locations to those in the target. Once again, this is consistent with an object-based account of updating, in that what is updated is a representation of a group of features and not merely an updating at the level of specific contours or surface features.

In summary, these experiments help emphasize that the behavioral consequences of masking should always be examined directly rather than simply being assumed. They also demonstrate that object updating is a powerful concept that can be applied not only to the understanding of conscious perception but also to unconscious processes that lead to motor responses. Finally, they show that the concept of object updating leads to testable predictions and that it can help explain some behavioral phenomena with a simpler set of assumptions than those needed when masking is used as a tool to merely reduce visibility.

28.5 Object updating and the perceptual asynchrony illusion

When participants view a display in which all objects alternate for equal durations between two colors and two directions of motion, the apparent coincidence of a specific color and motion does not always match their physical coincidence. This is the *perceptual asynchrony illusion* (PAI), and we have recently begun to examine the possible role played by object updating in the illusion (Oriet & Enns, under review).

In the seminal study of the PAI (Moutoussis & Zeki 1997a), participants viewed numerous squares that each alternated between red and green while at the same time alternating between moving up and down. Participants were asked to report the color of the squares while they were moving upward. Different feature durations were tested (e.g., 250 ms

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

and 350 ms) along with different phase relationships between the changes in color versus motion direction (e.g., simultaneous or 0 deg changes versus changes that were 90, 180, or 270 deg out of phase). The duration of each feature was not nearly as influential in the color reports of participants as was temporal asynchrony. When the change in color preceded the change in motion (e.g., upward moving squares were initially red and then turned green), the reported color was the one presented during the last portion of the motion period (e.g., upward squares were reported as green). In fact, changes in motion direction had to precede changes in color by 80 ms before changes in the two features were reported as coincident.

The PAI has attracted considerable theoretical interest because it has been claimed as a way to index the modularity of consciousness (Moutoussis & Zeki 1997a, 1997b; Zeki & Bartels 1998) or at least the relative time required for different types of visual properties to become accessible for conscious report (Arnold et al. 2001; Arnold & Clifford 2002; Nishida & Johnston 2002). In contrast, our interest in the PAI was piqued when we saw that almost no consideration had been given to the possibilities that (a) the illusion was influenced by the need to switch attention from one feature to another, or (b) the processes of object updating may be playing a role in this illusion as well. The fact that both of these possibilities had been overlooked was surprising to us because the PAI bears at least superficial resemblance to the other visual phenomena we have reviewed, such as backward masking, the flash-lag effect, and masked priming. In each of these effects, the focus of attention plays an important role and each of them involves errors of perceptual lag.

To be more specific, in each of the effects we have already discussed (i.e., masking, flash-lag, masked priming) there is always a *defining feature* that participants must become aware of first, before they are required to turn their attention to the *report feature*. For backward masking, the defining feature is the detection of a visible mask shape, which is the cue to try to report the immediately preceding target shape; in the flash-lag the defining feature is the flashed object, which is the cue to report the spatial position of the moving object. In the case of the PAI, participants must first see a defining direction of motion in a square before they are able to report on the color of the same square. If this similarity between tasks is more than superficial, then factors that influence the speed or difficulty of this switch in attention from the defining to the report feature should also influence the PAI. We therefore thought it was worth checking to see whether this illusion was influenced by the tendencies of the visual system to be somewhat sluggish in switching attention from one visual feature to another, and in the dynamic updating of object representations as new information becomes available.

In a first experiment, Oriet and Enns (under review) tested whether (1) perceptual lag errors are greater when features change asynchronously rather than synchronously, and (2) errors in either of these types of display are influenced by the difficulty of feature detection. Displays consisted of a series of 16–20 moving checkerboards that varied in both color and motion direction. The participant's task was to detect the single red checkerboard in the series (defining feature = color red) and to report the direction of motion taken by the red checkerboard (report feature = one of four motion directions). Having only a single target (red checkerboard) in the series allowed us to determine whether errors in reports

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

of motion direction reflected a perceptual lag, in which case errors would come from the motion directions that occurred *after* the target color, or whether they merely reflected a noisy signal, in which case errors should distribute randomly from values occurring before and after the target color.

The results did not support the hypothesis that slow attention plays the most important role in the PAI. To be sure, lag errors were more likely when displays changed more rapidly than when they changed slowly. This confirms that when participants err in their report of the features in a rapid sequence, they are not simply guessing but are responding based on features they see later in the sequence. However, the most important finding in this experiment was that lag errors occurred more frequently for asynchronous than for synchronous displays at the same display rates. If the time required to switch attention from defining to report features was the primary limiting factor in the illusion, then these two types of displays should show similar lag errors at the same rates of display change. This means that the PAI arises from factors that go well beyond a simple sluggishness in being able to shift attention from one feature to another.

The importance of temporal synchrony for feature binding has been documented previously. Perceptual objects can be defined solely by the common behavior of their parts in the temporal domain (Usher & Donnelly 1998; Lee & Blake 1999; Sekuler & Bennett 2001). The unusual circumstance that confronts the visual system in the PAI is that features are beginning and ending out of temporal step with one another. This likely 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. And while this is going on, the scene continues to change, so the system becomes vulnerable to the feature updating and overwriting that is normally beneficial as one views what first might look like a bird, but then gets reinterpreted as an airplane or even Superman.

To explore an object-updating interpretation of the PAI, Enns and Oriet (under review) conducted three more experiments. In one, they reasoned that if the PAI involves a fundamental ambiguity about the temporal characteristics of features that otherwise appear to be present in the same object, then it should be possible to reverse the direction of the illusion by switching the role of defining and report features. In other words, the PAI should reverse when participants report the color of the upward moving squares versus the motion direction of the red squares. When the roles of defining and report features were reversed in this way, for two different groups of participants viewing the same display sequences, the main finding was that perceptual lag errors were linked to the report feature, not to whether the feature was color or motion.

In another experiment, it was reasoned that the principles of object updating should apply equally well when the changing features are different sensory dimensions (e.g., color and motion direction, as in many previous studies) as when they are from the same sensory dimension (e.g., both are color features). They also tested whether the illusion was stronger when the features that needed to be linked in time were associated with the same object versus when they were associated with separate objects. They reasoned that the object-updating processes that lead to perceptual lag errors would be stronger when the changing

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

features came from the same object, as has already been observed in the flash-lag illusion (Moore & Enns 2004). Both of these hypotheses were confirmed.

In a last experiment, Enns and Oriet (under review) tested whether the illusion would vary with the perceptual grouping of features based on spatial proximity. According to the object-updating hypothesis, the PAI should occur with greater likelihood when the changing features occur in the same location in space (promoting object updating) versus when they occur in different locations (preventing object updating). This hypothesis was also confirmed.

Taken together, we interpret the results of the Enns and Oriet (under review) study as strong support for the idea that the PAI is a consequence of the ubiquitous perceptual updating process that occurs when objects are viewed in the context of dynamic and changing visual scenes. We believe that under normal circumstances, the temporal onset and offset of feature values are concurrent when they derive from the same object. The dark color and characteristic motion pattern of a bird both begin and end with the appearance and disappearance of the bird from view.

28.6 Conclusion

We began this chapter by asking how the visual brain creates a sense of order out of the seeming chaos of the visual input, so that it can arrive at the perception of a stable world, yet one in which objects may also reasonably change their characteristics over time. We proposed one very general solution to this problem based on the idea that visual representations are mediated at the level of perceived objects, rather than at the level of image features or at even higher conceptual levels of representation. We reviewed work demonstrating the consequences of object-mediated updating across a wide range of visual phenomena, including backward masking, the flash-lag illusion, unconscious response priming, and the perceptual asynchrony illusion. The ease with which a single theoretical construct – object updating – can help account for such a wide range of phenomena gives us confidence in the account. At the same time, we think it points to the generality of the process. Object updating is effective in providing stability for perception in the face of highly unstable sensory input and it appears to be ubiquitous in the visual processing of dynamically changing scenes.

Although evidence for object updating is robust in the laboratory studies we have described, we must hasten to add that the conditions in these studies were tailored to give insight into the emerging or online development of perceptual representations. The conditions were specifically designed to show that visual information obtained later in time could influence the emerging representation of an object. From this it should not be concluded that all perceptual representations formed from brief glimpses are necessarily susceptible to these forces. That is, we are not implying that there are no stable representations or that all perceptions are vulnerable to the updating process. Indeed, an important next goal in our research is to better understand what underlies the difference between representations that are vulnerable to updating versus those that are immune. We suspect

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

that visual attention will play a key role in this distinction, and that when we understand its role we will be better positioned to link current understanding of visual perception with that of visual short-term memory.

Acknowledgments

JTE was funded by a Discovery Grant from NSERC (Canada), AL by NSF (Awards 0527361 and 0309998), and CMM by NIH (MH067793).

References

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J Exp Psychol Gen* **129**: 481–507.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science* **287**: 2036–2038.

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: Behavioral and electrophysiological evidence. *J Exp Psychol Hum Percept Perform* **24**: 1737–1747.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Res* **44**: 1321–1331.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychol Sci* **8**: 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends Cogn Sci* **4**: 345–352.

Goodale, M., & Humphreys, G. K. (1998). The objects of action and perception. *Cognition* **67**: 181–207.

Henderson, J. M., & Ferreira, F. (2004). *The Interface of Language, Vision, and Action.* New York: Psychology Press.

Jiang, Y. H., & Chun, M. M. (2001). Asymmetric object substitution masking. *J Exp Psychol Hum Percept Perform* **27**(4): 895–918.

Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: Unconscious inhibition influences reaction time and response selection. *J Exp Psychol Gen* **131**: 255–269.

Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Res* **40**: 201–215.

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. *J Exp Psychol Gen* **133**: 475–493.

Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: Using apparent motion to isolate the object-level component of object substitution masking. *J Exp Psychol Hum Percept Perform* **29**: 106–120.

MacKay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature* **181**: 507–508.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action.* London: Oxford University Press.

P1: JZP Trim: 174mm × 247mm Top: 0.581in Gutter: 0.747in
CUUK929-28 cuuk929/Nijhawan ISBN: 978 0 521 86318 6 October 24, 2009 0:41

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychol Sci* **15**: 866–871.

Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *J Exp Psychol Hum Percept Perform* **31**: 1171–1180.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc Lond B Biol Sci* **264**: 393–399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proc R Soc Lond B Biol Sci* **264**: 1407–1414.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature* **370**: 256–257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci* **6**: 387–393.

Oriet, C., & Enns, J. T. (under review). The perceptual asynchrony illusion: Object updating and unbinding take time. *J Exp Psychol Hum Percept Perform*.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: Delays and localization errors in the visual system. *Nat Rev Neurosci* **3**: 191–200.

Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychol Sci* **12**: 437–444.

Tsotsos, J. (1990). Analyzing vision at the complexity level, *Behav Brain Sci* **13**: 423–445.

Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature* **394**: 179–182.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *J Exp Psychol Hum Percept Perform* **27**: 92–114.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends Cogn Sci* **6**: 211–216.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nat Neurosci* **1**: 656–657.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Res* **40**: 137–149.

Zeki, S., & Bartels, A. (1998). The asynchrony of consciousness. *Proc R Soc Lond B Biol Sci* **265**: 1583–1585.