BEHAVIORAL AND BRAIN SCIENCES (2008) **31**, 179–239
*Printed in the United States of America*
doi: 10.1017/S0140525X08003804

# Visual prediction: Psychophysics and neurophysiology of compensation for time delays

**Romi Nijhawan**
*Department of Psychology, University of Sussex, Falmer, East Sussex, BN1 9QH, United Kingdom*
**romin@sussex.ac.uk**
**http://www.sussex.ac.uk/psychology/profile116415.html**

**Abstract:** A necessary consequence of the nature of neural transmission systems is that as change in the physical state of a time-varying event takes place, delays produce error between the instantaneous registered state and the external state. Another source of delay is the transmission of internal motor commands to muscles and the inertia of the musculoskeletal system. How does the central nervous system compensate for these pervasive delays? Although it has been argued that delay compensation occurs late in the motor planning stages, even the earliest visual processes, such as phototransduction, contribute significantly to delays. I argue that compensation is not an exclusive property of the motor system, but rather, is a pervasive feature of the central nervous system (CNS) organization. Although the motor planning system may contain a highly flexible compensation mechanism, accounting not just for delays but also variability in delays (e.g., those resulting from variations in luminance contrast, internal body temperature, muscle fatigue, etc.), visual mechanisms also contribute to compensation. Previous suggestions of this notion of "visual prediction" led to a lively debate producing re-examination of previous arguments, new analyses, and review of the experiments presented here. Understanding visual prediction will inform our theories of sensory processes and visual perception, and will impact our notion of visual awareness.

**Keywords:** biased competition; compensation; coordinate transformation; feedforward control; flash-lag; internal models; lateral interactions; neural representation; reference frames; time delays; visual awareness

## 1. Introduction

### 1.1. Time delays in the nervous system

Time delays are intrinsic to all neural processes. Helmholtz, an eminent physicist and neurophysiologist of the nineteenth century, was among the first scientists to provide a clear measure of the speed of signal transmission within the nervous system. Using a nerve–muscle preparation, he electrically stimulated the motor nerve at two different points and noted the time of contraction of the connected muscle. The distance between the stimulated points divided by the time difference with which the muscle responded, gave the speed of neural transmission along the given section of the motor nerve. Before this seminal experiment in 1850, many well-known scientists including Johannes Müller had speculated that transmission of neural signals would be too fast to allow experimental measurement. However, to the surprise of the scientific community, the experiment revealed not only that the speed of neural conduction was measurable, but also that it was an order of magnitude slower than the speed of sound through air! In this article, I focus on visual delays, the problem of measurement of visual delays, and the effect these delays have on neural representations of change – such as those that result from visual motion – and on perception and behavior. In addition, the focus is on potential mechanisms that might compensate for visual delays.

Visual processing occurs in a series of hierarchical steps involving photoreceptors, retinal bipolar cells, ganglion cells, the lateral geniculate nucleus (LGN), the primary visual cortex (V1), and beyond. Neural delays have been extensively investigated at various levels within the visual system. For example, in response to retinal stimulation, significant neural delays (>10 msec) have been measured within the retina and at the optic nerve (Dreher et al. 1976; Kaplan & Shapley 1982; Ratliff & Hartline 1959). Schiller and Malpeli (1978) electrically stimulated axons of the optic nerve at the optic chiasm and measured the delay (≈2–3 msec) in the response of cells in the magno- and parvo-cellular layers of the LGN. Many

ROMI NIJHAWAN is a Reader (Associate Professor) in Psychology at the University of Sussex, England. He received his Ph.D. in 1990 from Rutgers University–New Brunswick. During his post-doctorate at the University of California–Berkeley, Nijhawan "re-discovered" what he termed as the *flash-lag effect* and subsequently published a brief paper on it (in *Nature*, 1994). This paper provocatively introduced the notion of prediction in sensory pathways, giving rise both to active empirical research on the flash-lag phenomenon, and to a spirited debate regarding the existence and nature of visual prediction. Since then, Nijhawan has published multiple papers in the area of sensory prediction.

© 2008 Cambridge University Press 0140-525X/08 $40.00



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

studies on the macaque have recorded the delay with which neurons in the primary visual cortex (V1) respond to retinal stimulation (Raiguel et al. 1989; Maunsell & Gibson 1992; Schmolesky et al. 1998). An estimate based on a large database yields an average delay of approximately 72 msec of V1 neurons (Lamme & Roelfsema 2000). However, there is usually a range of delays with which neurons in a given area of the cortex respond. Different neurons in the macaque inferotemporal cortex, for example, respond to complex visual information with delays ranging from 100 to 200 msec (Nakamura et al. 1994).

### 1.2. Neuronal and behavioral delays for discrete stimuli

Discreteness of stimulation is key to defining time delays for both neural responses and in behavioral tasks. For visual neurons, time delays are typically measured in response to light flashes, light onsets/offsets, or short electrical pulses applied to some point in the visual pathway. Neural delay is defined as the time interval between the discrete change in stimulation and change in neural activity at the target site. The onset time of a discrete stimulus is easily determined. However, the definition of neural delays becomes more complex for time-varying stimuli that are a continuous function of time. For such stimuli, neural delay can only be defined with respect to an instantaneous value of the stimulus.

Discreteness of stimulation is central also to defining behavioral delays. The paradigm that formed the cornerstone of experimental psychology in the latter half of the nineteenth century with the work of Helmholtz, Donders, Wundt, and Cattell, involved measurement of simple reaction times (Meyer et al. 1988). In simple reaction-time tasks, the participant produces a prespecified overt response – for example a button press – in response to a discrete suprathreshold stimulus such as a flash of light, a sound burst, or a tactile stimulus. The stimulus–response interval for the light stimulus is 200–250 msec, whereas for sound or touch it is about 150 msec. The minimum latency for a voluntary learned motor response appears to be around 100–120 msec (Woodworth & Schlosberg 1954, p 9). Simple reaction time is directly related to input–output delays of single neurons and of neural chains (see, e.g., DiCarlo & Maunsell 2005). Neural delays vary substantially across cell types and modalities. For example, the transduction of mechanical stimulation by the mechanoreceptors of the touch system, and the response of hair cells to mechanical movement of the cochlear fluid caused by sound, is much quicker than the speed with which light is transduced by the photoreceptors. The fact that reaction time to a light flash is longer than that to a sound burst, for example, is thought to be a direct consequence of the faster processing of auditory neurons. Likewise, a response requiring a longer neural pathway will be slower. Thus, reflex behaviors such as the knee jerk can be produced in about 40 msec via a shorter pathway involving only the spinal cord, whereas behaviors such as in "choice reaction time" experiments, which involve additional cortical processing, are decidedly slower. Additional variables that impact behavioral delays are: the size of the neurons involved, whether the neurons have myelinated axons or not, the number of synapses between the peripheral receptors and the central nervous system (CNS), the types of the intervening synapses, the strength of stimulation, and to an extent certain qualitative aspects of the stimulus.

Several investigators have directly studied the relationship between neuronal latency and reaction time. In measurements of reaction time, presentation of a discrete stimulus starts a time-counter, which the subject's overt motor response stops. In measurements of neuronal latency, however, discerning the neural event used to stop the time-counter can be more challenging. In one study, researchers recorded from single neurons in the monkey motor cortex while the animal performed a learned flexion/extension arm movement in response to a visual, somatosensory, or auditory stimulus (Lamarre et al. 1983). They reported a constant delay between the change in the firing rate of the motor cortical neuron and arm movement, irrespective of the stochastic and modality contingent variation in reaction times. Whether change in a neuron's spike rate following the stimulus, on a single trial, can be used in the measurement of neuronal latency (i.e., to stop the time-counter) is debatable. Individual spikes provide a sparse sample of the assumed underlying rate function of the neuron. Thus, even if one were to assume a step change in the underlying rate function triggered by a discrete stimulus, the variance in spike times dictates averaging over many trials to determine the precise time of the "true" change in the firing rate triggered by sensory stimulation (DiCarlo & Maunsell 2005).

### 1.3. Neuronal and behavioral delays for continuous stimuli

Animals encounter both discrete and continuous environmental events. At one extreme are discrete stimuli resulting from unexpected events in nature (Walls 1942). At the opposite extreme are stationary stimuli, for example, nearby objects such as trees, that can be just as behaviorally relevant as changing stimuli, but for which the issue of neural delays arises only when the animal itself moves (see further on). Between the two extremes of the discrete–continuous continuum, there are a multitude of time-varying stimuli unfolding as smooth functions with respect to time. Consider the role of neural delays in a situation where a potential prey confronts a predator who has appeared in a surprise encounter. In relation to the speed at which changes occur in the prey's retinal images, the central transmission of retinal information can be extremely slow. In other words, a significant environmental event, such as the predator adopting the attack posture, can take a significant fraction of a second to be communicated to and be processed by the parts of prey's CNS that control the escape response.

Properties of neurons allowing for higher transmission speeds confer an enormous advantage to animals, and adaptations related to faster conduction of neural information are well known. For example, the myelination of nerve fibers, which is not found in invertebrates, allows for much faster conduction speeds in vertebrate nervous systems. However, slowness of neural response does not necessarily mean a disadvantage to the animal. In fact, some slower processes appear to be linked to the animals producing more adaptive behaviors. First, there is the well-known speed–accuracy tradeoff, where



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

increased speed of actions necessarily leads to lower accuracy (Woodworth 1899; Meyer et al. 1988). Second, although visual response is slower than touch, it does not necessarily lead to a disadvantage for the animals, as vision allows for objects to be detected at a distance, so animals are able to plan behaviors (Sarnat & Netsky 1981). Within limits, animals are able to plan their behavior when conditions are stationary or are changing smoothly, but they must rely directly on sensory input when there is a discrete change in conditions for which there are no priors.

A temperature drop over time and an object changing position over time are common time-varying stimuli. A closer look at neural delays in the context of such stimuli reveals some complexities and a puzzle. At the stimulus end, neural delays in the processing of a continuously unfolding stimulus can be meaningfully defined only in relation to an instantaneous value of a continuous variable (IVCV; e.g., 0° temperature). The IVCV at a given instant of time is what would, so to speak, start the time-counter. As discussed earlier, the question of determining the neural event that may be used to stop the time-counter is challenging enough for discrete events, but an even more difficult problem arises in discerning the neural event that should stop the time-counter for a continuous stimulus.

Suppose a scientist wants to determine the latency with which a spiking neuron responds to a continuous stimulus. Delay in the neural response can only be measured in relation to event $IVCVt_0$ (IVCV at selected time $t_0$). Typically, a neuron tuned to a particular variable will respond to not just one value of the variable but to a range of values. Thus, the neuron will respond not just to $IVCVt_0$ but also to IVCVs in the neighborhood of $IVCVt_0$. Consider a hypothetical V1 neuron with narrow orientation selectivity. Suppose further that this neuron responds with a latency of 70 msec and produces a maximum spike rate to vertical line orientation (0°). If a line rotates continuously at speed $\omega$, and is presented within the receptive field of this neuron, then the neuron's response will vary as a function of line orientation. Narrow orientation selectivity means that this cell will produce its strongest response when the line is vertical, and a somewhat weaker response to the line orientation that deviates from the vertical by a small angle ($\omega \bullet 70$ msec). For a continuously rotating line, neural delay may be defined, for example, between the two following events: (1) the line reaching the vertical orientation, and (2) the moment the neuron's spike rate reaches maximum. In this experiment, the signal to stop the time-counter would be "spike rate to line orientation 0°" – "spike rate to line orientation 0° – ($\omega \bullet 70$ msec)," where the negative sign means line orientation before the line reached vertical. Because spikes are collected in time bins, detecting such a signal can be challenging (Bialek et al. 1991).

However, even if it were in principle possible to pick up the neural signal for $IVCVt_0$ in order to determine latency, there is a more basic issue. The method described here would require the following assumption to be true: Neural response to $IVCVt_0$ (in the preceding example the vertical line orientation) is the same whether the stimulus is presented in isolation, or as part of continuous stimulus presentation. This assumption may be valid for "early" neurons in the vision pathway. However, such an assumption is almost certain to be invalid for neurons as one moves higher up in the visual hierarchy. This is simply because neurons in the "later" parts of the visual pathway are stimulated not just by direct sensory input, but also by ongoing perceptual and motor processes (Treue 2003).

The question of where a neuron is located in the visual hierarchy and how its response properties change has been revisited recently. One can define a neuron as being "stimulation centered" or "perception centered," depending on its response properties. Recent neurophysiological studies suggest that the activity of neurons earlier in the visual pathway corresponds more with sensory stimulation, whereas behavior of later neurons seems to correspond more with perception (Treue 2003; Williams et al. 2003). Depending on whether a given neuron is stimulation centered or perception centered, one can or cannot define its sensory input delays in response to continuous stimuli. For example, a perception-centered neuron's response may be related to what the animal anticipates the stimulus would do in the near future, rather than what the stimulus is doing at a given moment (Eskandar & Assad 1999). Hence, in such cases there is the potential of relating a neural signal to the wrong IVCV – which suggests that if and only if the $IVCVt_0$ directly drives a neuron's response, may one meaningfully define neural delays for that stimulus event.

Interaction of animals with continuous stimuli, such as moving objects, reveals a puzzle: Behaviors directed at moving objects seem subject to no appreciable delays. Among the innumerable examples, this point is brought forth clearly in the following three examples:

1. In vision-based predator–prey interactions, there is often extreme pressure on each animal to produce behavior vital to its survival. On the basis of neural and behavioral delays, as revealed in experiments employing discrete stimulation, the prediction would be that actions should always lag behind targets in proportion to the magnitude of the delay, irrespective of whether the stimulus is discrete or continuous. The lags are expected to be large when the rate of change of the external stimulus, for example, the speed of a moving object, is large. The expected lags for smoothly changing stimuli are, however, never seen in typical predator–prey interactions.

2. Scientists have observed behavior of toads catching woodlice in extremely low levels of illumination. Close to the detection threshold of the toad's retinal ganglion cells there are one to two photoisomerizations per 100 rods per second. At such low levels of light intensity, the ganglion cells sum signals from 400–1000 rods over a period of 1.0–2.5 seconds, so vision necessarily becomes slow. Aho et al. (1993) noted that when a worm-like stimulus was presented to a ganglion cell's receptive field, the cell response occurred about 3 seconds after the presentation of the stimulus (Aho et al. 1993). Nonetheless, even under these extremely limited visibility conditions a toad could still produce an accurate tongue-snap response to catch a moving worm!

3. The final example concerns the highly practiced behavior of humans in relation to fast-moving objects. In the game of cricket, a fast bowler can bowl at speeds of more than 90 mph. At these speeds, the batsman has a very short time to decide how to hit the ball, and an even shorter time in which to decide how to get his head



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

(say) out of the line of the ball in case the ball is a bouncer heading directly towards his head. The required temporal accuracy could be as small as 5 msec (Tresilian 1993) or even smaller (Land & McLeod 2000; Regan 1992). As delays in the output of photoreceptors can be appreciably longer than 5 msec, common observations in high-speed ball games immediately bring up the issue of compensation for neural delays.

## 2. Compensation of neural delays

The delay between the input and output of a black box – which is a widely used concept in engineering and systems neurobiology (DeYoe & Van Essen 1988) – is known as phase-lag. During animal interactions phase-lags can be large, and appropriate compensation is necessary for adaptive behavior. According to Ghez and Krakauer (2000, p. 657): "If [phase] lags are long and external conditions change rapidly, specific feedback corrections may not be appropriate by the time they are implemented." One type of mechanism that could compensate for phase-lags, called feedforward control, relies on information that the nervous system acquired before the actual arrival of external input from the sensors (Ghez & Krakauer 2000).

It is perhaps not a coincidence that Helmholtz was also a pioneer in recognizing the need for compensation of delays and proposing feedforward control. Feedforward control is most clearly seen in situations in which sensory stimulation results as a consequence of movements produced by the animal itself (reafference). A famous example by Helmholtz is the "canceling" of the effects of retinal image motion during voluntary eye movements (Jeannerod et al. 1979); when one voluntarily moves one's eyes, the visual world, which shifts on the retina, is not seen as moving. The notion of feedforward control suggests that a "comparator" receiving afferent signals from the retina due to image motion also receives a corollary of the motor command to the eye muscles (von Holst & Mittelstaedt 1950; Sperry 1950). The copy of the motor command "cancels" the retinal input signals, resulting in perceptual stability of the visual scene. This view is supported by the fact that when a copy of the motor command to eye muscles is absent, such as during involuntary eye movements (e.g., due to externally forced movement of the eyeball), the visual environment does appear to shift.

Predictive or anticipatory control, which are concepts related to feedforward control, have been demonstrated at the level of single cells, in modules serving sensory-motor functions (e.g., the cerebellum), and in psychophysical tasks in humans. Single neurons in the parietal cortex (lateral intraparietal area) have been shown to respond in a predictive manner to changes in visual input caused by voluntary saccades (Duhamel et al. 1992). Prediction has also been shown in psychophysical experiments investigating the grip force required to lift weights (Johansson & Westling 1988), in a pole-balancing task (Mehta & Schaal 2002), and in weight unloading (Diedrichsen et al. 2003). The cerebellum has been identified as a possible sensory-motor structure contributing to predictive control (Kawato et al. 1987; Wolpert et al. 1998). However, predictive mechanisms are unlikely to be localized in any one sub-system of the CNS. For example, in humans, damage to the cerebellum leaves predictive responses in the weight-unloading task intact (Diedrichsen et al. 2003), which suggests that prediction is a distributed property of animal nervous systems.

The notion that sensory processes per se may be predictive is not widely recognized. In fact, previous suggestions of sensory prediction (Nijhawan 1994) have led to controversy (Cavanagh 1997; Krekelberg & Lappe 2001; Nijhawan 2002; Schlag & Schlag-Rey 2002). The position that sensory processes are predictive should not be surprising for the following reasons. First, there is debate as to whether the function of parietal neurons, a cortical area in which predictive processes appear to be prevalent, is sensory or motor. Second, sensory processes are known to contribute to prediction in sensory-motor tasks, such as in the canceling of retinal image motion signals during voluntary eye movements and reduction in the latency of parietal neurons (Duhamel et al. 1992). Therefore, in the very least, sensory processes participate in prediction. Finally, it is rather unparsimonious to suggest that although prediction is a general property of the nervous system found in many different areas, sensory areas are excluded. This would be particularly awkward for vision, given that more than 50% of the primate neocortex is involved with visual processes (Kandel & Wurtz 2000, p. 497). Indeed, the concept of the "internal model" of the motor system, which serves a predictive function, may be generalized to the domain of perception (Kawato 1999). The goal of this article is to place the notion of sensory prediction in general, and visual prediction in particular, on firmer ground by addressing the existing controversies, re-analyzing experimental results, and outlining the wide-ranging issues that consequently arise.

## 3. The problem: Topography of the early visual system, neural delays, and spatial lags for motion

For humans, moving objects can be targets for perception, orienting responses, interception, or avoidance. The concept of prediction as it exists in the literature is a somewhat limited point of view resulting from focus on motor or sensory-motor processes at the expense of visual processes. A majority of the literature suggests that predictive processes are one step removed from sensory processes. These processes are located in the parietal and frontal cortical areas, perhaps even beyond the sensory-motor integration areas, or in sub-cortical structures such as the cerebellum, with the immediate goal of producing movement. One corollary of the viewpoint that there is no visual compensation for delays is: Perceptual state of a changing visual stimulus should trail behind the actual current state of the stimulus. Changing patterns of illumination, especially those resulting from movement, stimulate the visual system strongly and frequently elicit a behavioral response from the animal (Barlow 1961a). However the quintessential stimulus of visual motion leads to the following conundrum. The retina is mapped topographically on to many maps of the brain, and particularly well known are topographic maps found in V1 (Felleman & Van Essen 1991; Tootell et al. 1982). As an object moves over the retina, the representation of the object (wave of neural activity generated by the object)



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

shifts over the cortical surface containing the topographic maps. Because of neural delays, the object's representation in the cortical map should lag behind the object's retinal representation (Figure 1a). As our perception is based in part on neural activity in cortical maps, this renders a perceptual lag in the position of a moving object in relation to its actual position (Figure 1b).

The recent statements concerning the perceived lag of moving objects have lead to a vigorous debate (Cavanagh 1997; Gegenfurtner 1999; Krekelberg & Lappe 2001; Nijhawan 1994; 2002; Schlag & Schlag-Rey 2002). What might be the magnitude of the lag? At a simplistic level the lag should equal the product of visual input delay and object velocity. However, in order to precisely measure input delay in perception one would have to first make several assumptions. The most problematic assumption is that there is some specific brain site where neural processes culminate in perception. It has been possible to measure delays between two neural sites (Maunsell & Gibson 1992; Raiguelet al. 1989; Schmolesky et al. 1998) and to show that these delays differ for different pathways (Schiller & Malpeli 1978). However, there are arguments that there is no specific brain site where the stream of visual processing ends to yield perception (Dennett & Kinsbourne 1992).

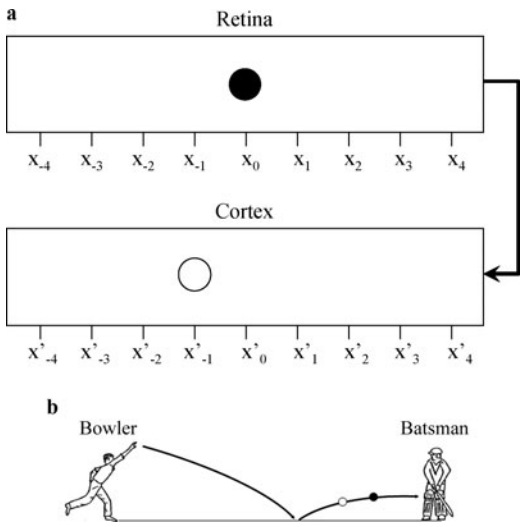



Figure 1. (**a**) The retina is shown as a one-dimensional coordinate system mapped topographically onto the cortex. For simplicity, the thalamic stage is not included. A moving ball's image travels from left to right at constant velocity. It is assumed that the neural delay between the retina and the cortex is equal to the time the ball takes in moving from retinal coordinate $x_{-1}$ to $x_0$. At the instant depicted, the cortical coordinate of the ball, showing greatest activity is $x'_{-1}$, which corresponds to retinal coordinate $x_{-1}$, while the ball's actual retinal coordinate at the instant depicted is $x_0$. (**b**) The batsman in cricket views a ball heading toward him after bouncing some distance in front of him. The batsman's perception is assumed to depend on the cortical activity triggered by the moving ball. At the instant depicted, the physical position of the ball (filled circle) is ahead of the ball's perceived position (unfilled circle). The lag is proportional to the neural delay and object velocity, which for a cricket ball can be more than 90 mph. For an assumed delay of 100 msec, the difference between the ball's perceived position and its actual position should be 13.2 ft. Adapted from Land and McLeod (2000).

Despite the existing controversies, however, limits can be imposed on the "where" and "when" of perception. In response to retinal stimulation significant neural delays have been measured within the retina and the optic nerve (Dreher et al. 1976; Kaplan & Shapley 1982; Ratliff & Hartline 1959). It is generally believed that neural activity at the level of photoreceptors is not sufficient to yield visual awareness. It follows that if the photoreceptor delay is 10 msec and object velocity is 10 m/sec then the lag in perceived position of the moving object should be at least 10 cm. The same logic holds for measurements at different points of the optic nerve, the LGN, and so on; in fact, it may be argued that observers are not directly aware of neural activity even in V1 (primary visual cortex) (Crick & Koch 1995; He et al. 1996) which involves much longer delays. Therefore, the perceptual lag postulate may be re-stated as: Neural delays should cause the perceived position of a moving object to lag behind its physical position by at least a distance $v\Delta t_{max}$, where $v$ is object velocity and $\Delta t_{max}$ is the maximum cumulative delay in the vision pathway until a point in the pathway where neural activity can be said to *not* yield visual perception.

Because of the slowness of vision, a large chunk of the delay during visual-motor interactions is due to visual processes per se. Which component(s) of the animal nervous system compensate for visual delays? Although late compensatory mechanisms are frequently invoked, an interesting possibility is that of contribution from the visual system itself (Ramachandran & Anstis 1990; De Valois & De Valois 1991). This possibility is related to a conjecture by David Marr, who, in his famous treatise on vision, considered the possibility of visual compensation in diving birds that feed on fish first seen from above the water's surface (Marr 1982). Because of the water's refractive properties, the fish's position is optically shifted when it is viewed from the air, and the bird's nervous system must correct for this optical distortion. Although a correction is not present in humans, who continue to perceive submerged objects as shifted in relation to their true positions, Marr speculated that the nervous system of birds whose survival depends on correctly localizing the fish might visually correct for the optical distortion. Thus, the fish's true position would be given "directly" by the bird's visual system, as opposed to through an "inferential," non-visual process. For humans, the case of moving objects (Fig. 1b) presents an analogous situation where the basis of error and impetus for correction is neural rather than optical.

## 4. The phenomenon: Concatenation of continuous and discrete visual events

Neurophysiological and psychophysical experiments reveal delays for discrete stimulus events. In a seeming contradiction, animal behaviors directed at continuous stimuli (e.g., moving objects) reveal virtually no delay. Does a similar contradiction exist for visual processes? In other words, are visual responses to discrete stimuli delayed, but visual responses to continuous stimuli not? Were this the case, the concatenation of discrete and continuous visual events in the same display should reveal an anomaly. An experimental paradigm, known as the



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

flash-lag effect, combines a discrete stimulus (a light flash) and a continuous stimulus (a moving object). Indeed, when observers view a display in which the moving item is spatially and temporally aligned with the flashed item (Fig. 2a), the flashed item is seen as spatially lagging behind the moving item (Fig. 2b) (Hazelhoff & Wiersma 1924; Metzger 1932; Mateeff & Hohnsbein 1988; Nijhawan 1992).

There is a parallel between the notion of visual prediction emerging from the flash-lag anomaly and the notion of motor prediction appearing on the basis of behavioral contradiction (Nijhawan 1994) – in both cases there is a discrete event and a continuous event. Visual prediction suggests that visually guided behaviors, such as hitting a fast-moving ball, require the observer to have information concerning the ball's physical position before contact. If there were no visual prediction, then for a 100 msec delay, the observer would see a ball traveling at 90 mph to be about 13.2 ft behind its actual instantaneous position (see Fig. 1b). Consistent with Marr's conjecture that visual mechanisms of diving birds compensate for optical displacement of submerged objects, the visual compensation account of the flash-lag effect (Nijhawan 1994) suggests that the expected lag in the perceived position of moving objects (resulting from neural delays) is compensated for by visual mechanisms. In contrast, discrete events such as flashes have no priors, and their perception can only be based on delayed sensory input. (There is probably no neural process that could overcome the delay in the registration of a flash [van de Grind 2002]). This results in the flash appearing in a position lagging behind the perceived (spatially extrapolated) position of the moving object. The goal of visual prediction is to use priors contained in the unfolding visual stimulus to create a perceived state of the world that matches, as far as possible, the actual state of the world.

## 5. The flash-lag debate

There is intense debate as to whether the flash-lag effect does indeed provide prima facie evidence of a predictive process in vision, and more generally as to whether prediction in vision could serve any function. The controversy is multifaceted. First, the results of the flash-initiated and flash-terminated conditions of the flash-lag effect (Eagleman & Sejnowski 2000; Khurana & Nijhawan 1995; Khurana et al. 2000; Nijhawan 1992), both of which involve unpredictable changes to the moving stimulus, render results that appear contradictory to the notion of visual prediction. The second basis for the controversy is logical. How can prediction, which is typically considered to be a high-level or cognitive phenomenon (see, e.g., Thier & Ilg 2005), be part and parcel of perception – predominantly a result of processes in the feedforward vision pathway? On the other hand, how can visual percepts be influenced by higher-level (e.g., parietal) neurons known to be involved in prediction? The third source that has fueled the debate is related to doubts over the functional significance of visual prediction. There is unequivocal evidence and acceptance that "late" motor or sensory-motor processes serve a predictive role. These late processes are certainly capable of compensating for all the delays in the sensorimotor loop, including those incurred by visual processes per se. So what could be the functional significance of prediction in vision? Each of these challenges is now addressed in turn.

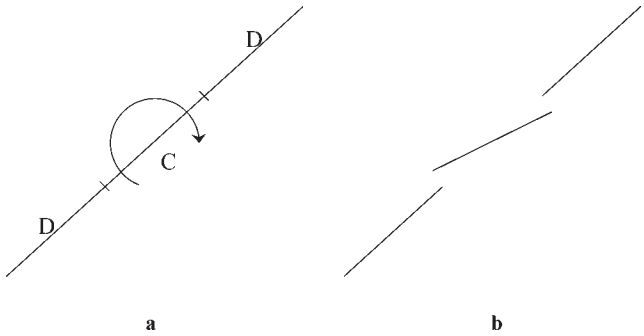



Figure 2. (**a**) In a dark room, a single physical rod, made of three segments, rotates clockwise in the direction of the arrow. The segment labeled "C" is illuminated with a continuous light source, while the discrete segments labeled "D" are illuminated with a brief flash produced by a stroboscope. (**b**) The percept of the observers. Adapted from Nijhawan (1992).

### 5.1. The standard view and its revision suggested by the flash-lag effect

I attribute to the "standard view," which represents one side of the debate, the following statement: A moving object should be visible in a position that it occupied in the recent past. Figure 3 presents this standard view in space–time plots. One of the two plots in Figure 3 represents an object moving at constant velocity $v$, and a flash presented at position $x_0$ at time $t_0$. This plot represents the physical state of affairs. The second plot represents how a neuron "sees" these events. This plot, called the ds-error line, results from the standard view of neural delays in the processing of physical events.

If the scenario depicted in Figure 3 were true, then there would be no flash-lag effect (e.g., note that on the intersection of the ds-error line and the $x = x_0$ line of Figure 3 the moving and flashed objects coincide). Since the flash-lag effect occurs, a revision of the view presented in Figure 3 is in order. Figure 4 shows the revised picture.

The reduced ds-error line in Figure 4 corresponds to perception. On the generalized visual prediction view (see below), there are multiple reduced ds-error lines, some of which will not be relevant for perception but only for visually guided action.

### 5.2. Experiments with unpredictable motion

In the flash-terminated condition, a smoothly moving object disappears unpredictably, simultaneously with the flash. In the flash-initiated condition, an unpredictable onset of the moving object occurs simultaneously with the flash, following which the motion is smooth. The results of these conditions (Nijhawan 1992) are as follows: When the flash-lag display "was separated into two temporal intervals ... [o]ne interval consisted only of events before the flash (past-interval) and the other interval only of events after the flash (future-interval). The observers reported no [flash-lag] effect in the past-interval but did so in the future-interval. The magnitude of the



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

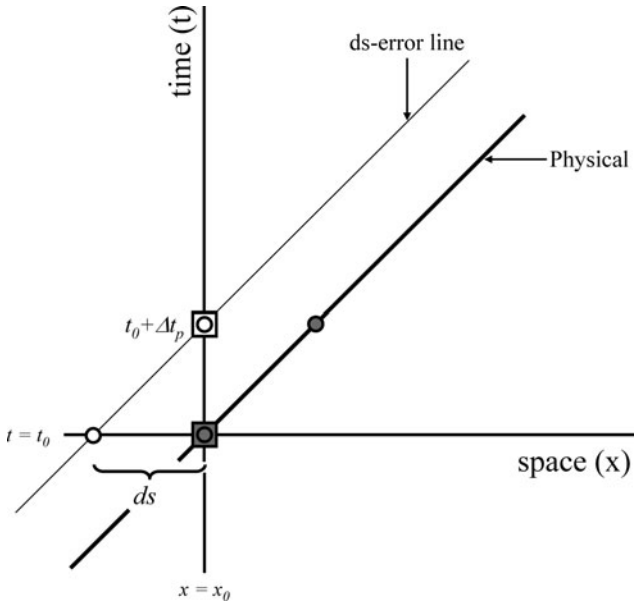



Figure 3. Representation of the "standard" view with space–time plots of an object moving at constant velocity $v$ (thick line) and how a neuron "sees" the object with some delay (thin line). A brief flash (filled gray square) presented in position $x_0$ at time $t_0$ is "seen" (outline square) by the neuron in position $x_0$ at time $t_0 + \Delta t_p$, where $\Delta t_p$ is input delay to perception. Two events, the arrival-of-the-moving-object in position $x_0$ (filled gray circle on the $x = x_0$ line) and the "seeing" of this event by the neuron (outline circle on the $x = x_0$ line) occur at different times because of neural latency between the event and its perception. At a given time (say $t_0$), the physical position of the moving object spatially leads (filled gray circle on $t = t_0$ line) the position in which the neurons "see" the object (outline circle on $t = t_0$ line). The spatial gap between the physical position and the neurally represented position are referred to as ds-error. The standard view suggests that the perceived object travels on the ds-error line. Adapted from Krekelberg and Lappe (2001).

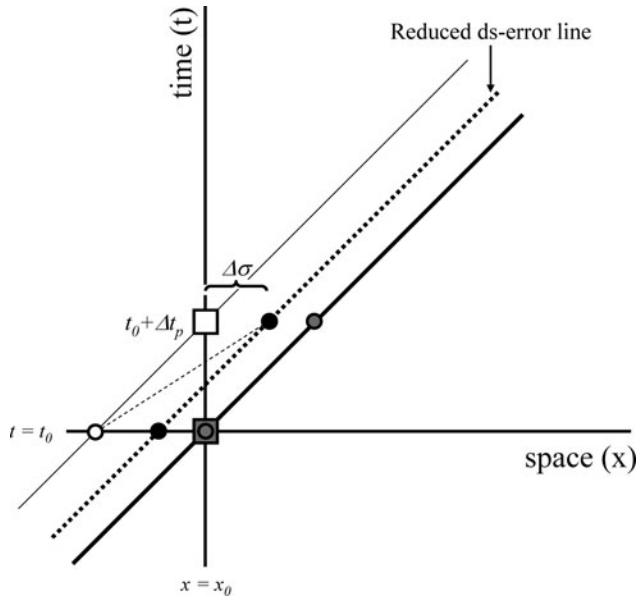



Figure 4. A revision of the standard view is forced by the flash-lag effect. This figure shows a new trajectory, the "reduced ds-error" line, which is a line parallel to the ds-error (and the physical) line, passing through the point in which the moving item is seen in the flash-lag effect. The distance along the $x$ (space)-axis between the reduced ds-error line and the ds-error is $\Delta\sigma$; where $\Delta\sigma$ is the experimentally determined flash-lag effect (see, e.g., Nijhawan 1994). The standard view needs to be revised, as on this view, the moving object is seen on the ds-error line throughout, and in particular, at $t_0$ – the time of the flash (open circle). The flash-lag effect shows that the moving object is seen ahead of the flash by distance $\Delta\sigma$ at $t_0 + \Delta t_p$ (filled black circle). Thus, for both the standard view to be true and the flash-lag effect to occur, the moving object would have to appear to speed up (which corresponds to a segment of a space–time trajectory, shown as the thin broken line of different slope). Because the moving object is assumed to travel at constant velocity, the standard view is untenable.

[flash-lag] effect for the future-interval was as strong as that for the past + future interval."

**5.2.1. Flash-initiated condition.** In the spatial extrapolation view, the perceived position of the moving object is based on a type of "guess" as to the object's actual position. All guesses are a product of prior information. The crucial feature of the flash-initiated display is that the physical time (and position) at which the moving object appears is unpredictable. Prima facie, one would expect no flash-lag in the flash-initiated display. Yet, the findings are that the flash-lag effect in the flash-initiated condition is undiminished (Khurana & Nijhawan 1995). This result has been obtained with real (analog) motion and extremely brief (100 μs) flashes (Nijhawan 1992), and with digital displays in which the moving and flashed objects are aligned in the first video frame for 10 or more milliseconds (Alais & Burr 2003; Brenner & Smeets 2000; Eagleman & Sejnowski 2000; Khurana et al. 2000; Sheth et al. 2000).

Since the visual prediction account is strictly spatial, it might seem that it is necessary to assume a shift in the coordinates of the moving object occurring in a direction parallel to the time axis (see Fig. 4), which would necessitate infinite speed of neural processing. Excluding the sub-millisecond photochemical responses to bright flashes (Wald 1968), which are unlikely to contribute to extrapolation, no processes in the visual system are fast enough. However, this reasoning is flawed, as the flash-lag effect is a relative judgment that only requires that the time taken for the forward shift in coordinates of the moving object is small in relation to the time taken for the processing of the flash. Therefore, the all-important questions for visual prediction are: When are the neural processes responsible for motion extrapolation initiated, and how long do they take to be completed?

On the spatial extrapolation account, perception of the moving object in the flash-initiated condition is subject to an input delay ($\Delta t_p$) following its abrupt onset, which is the same as the delay for the flash. Say motion onset occurs at time $t_0$ (simultaneously with the flash); the moving object and the flash are both first perceived at time $t_0 + \Delta t_p$. Further, suppose that retinal coordinate $x_0$ is represented by coordinate $x_0'$ on a cortical retinotopic map. If the abrupt onset of the moving object occurred in retinal position $x_0$, and there were no motion extrapolation, then the neural activity as a result of this event will occur in cortical position $x_0'$ after a delay of ≈100 msec (assuming 100 msec latency). How much



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

additional time over and above the ≈100 msec delay is required to generate a neural representation of the object that is shifted ahead of $x'_0$ by a distance corresponding to $\Delta\sigma$ (see reduced ds-error line in Figure 4)? For spatial extrapolation to be the correct account of the flash-initiated result, and for this mechanism to be behaviorally relevant (e.g., for producing quick actions), the additional time must be small.

The following analysis estimates the additional time for extrapolation to be less than 2% of the "baseline" delay of ≈100 msec between stimulus presentation and its perception (De Valois & De Valois 1991). This analysis is based on a prominent property of the vertebrate visual system that it consists of two types of pathways (Dowling 1979; Tessier-Lavigne 2000): the vertical pathway and the horizontal pathway (Fig. 5). In the early visual system, the vertical pathway corresponds to information flow from photoreceptors to bipolar cells to ganglion cells to LGN cells, and so on. The horizontal pathway in the retina includes information flow from one retinal site to a neighboring one via horizontal cells, the amacrine cells, and tangential extensions of dendrites and axon terminals. Analogous processes are present in the cortex (Bringuier et al. 1999). In evaluating the consequences of neural delays and compensation, it is best to treat these two pathways independently. Delays in signals traveling along the vertical pathway are the putative source of lagging retinotopic coordinates for visual motion. In contrast, the notion of visual prediction is concerned primarily with horizontal processes, which transmit neural information between two neighboring retinotopic sites (Barlow 1953; 1981; Berry et al. 1999).

It is premature to commit to a specific neural mechanism that would cause a "forward" shift in the coordinates for moving objects. This shift could be based on the interaction of signals between the magnocellular and parvocellular systems (Barlow 1981) (see Fig. 6), on a spatial reorganization of receptive fields or of retinotopic maps, or on some other yet more specialized mechanisms (see further on). Nonetheless, any viable predictive mechanism would involve two time-consuming processes: "obtaining a motion sample" and "spatial extrapolation." How soon after motion onset can the sampling of motion and the spatial extrapolation processes be completed to produce the required forward shift ($\Delta\sigma$)? At the outset it must be recognized that the correct estimate of the time will be based on the neural representation of $\Delta\sigma$ (i.e., retinotopic distance corresponding to $\Delta\sigma$) and the time taken by horizontal processes to cover this distance (see, e.g., Anderson & Van Essen 1987; Berry et al. 1999). In particular, a correct time estimate will be based on neither the time the moving object takes to travel the physical distance corresponding to $\Delta\sigma$, nor on physical or perceived velocity of the object; rather, it will be based on neural distances and speed of neural signals (see further on).

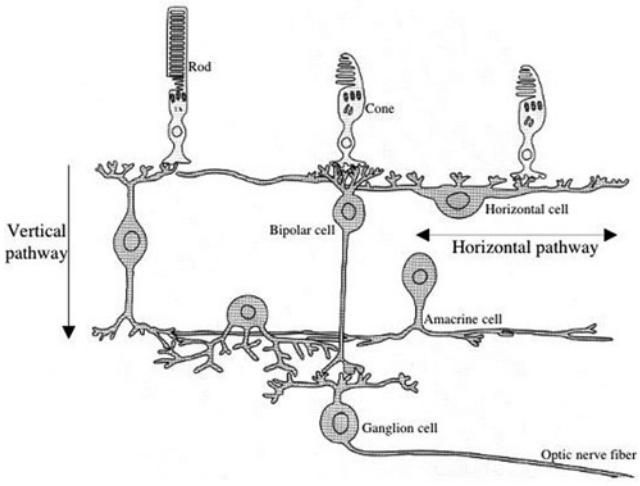



Figure 5. The pathways in the retina depicting the vertical (distal–proximal) and the horizontal pathways of information flow. Adapted from Dowling (1979) and Tessier-Lavigne (2000).

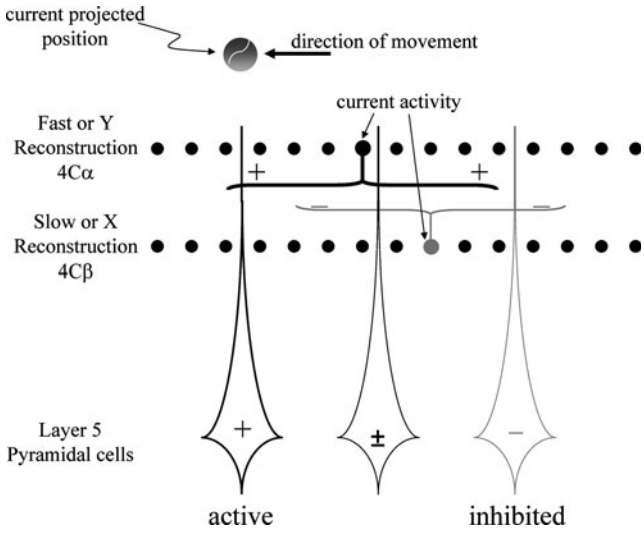



Figure 6. A hypothetical model based on a network with excitatory and inhibitory interactions. This network involves the layer 5 pyramidal cells in V1. The vertical processes of the pyramidal cells are the apical dendrites extending into layer 4. These dendrites receive horizontal connections from both the faster transient (Y) cells and the slower sustained (X) cells located in different sub-layers of layer 4. A moving ball's current retinal position is shown at the top of the figure. The resulting neural representations in layer 4 X and Y maps are "misaligned" with respect to the retinal representation and with respect to each other (due to different speeds of the sustained and transient channels). The excitatory and inhibitory horizontal connections cause the leftmost pyramidal cell to be excited, the rightmost pyramidal cell to be inhibited, and the middle cell's activity to remain unchanged (because it receives both excitatory and inhibitory inputs). Thus, the leftmost pyramidal cell behaves as if it is representing the "current" position of the ball. Adapted form Barlow (1981).

***5.2.1.1. Motion sampling.*** In the flash-initiated display, the sudden motion onset causes neurons at early levels of the visual pathway, for example, a linear set of photoreceptors, to respond before the response of corresponding thalamic or cortical neurons. A flash likewise stimulates photoreceptors before the thalamus or the cortex. However, there is a crucial difference between the motion onset stimulus and the flashed stimulus. At ordinary speeds, the motion-onset stimulus will undergo a significant shift across several photoreceptors even prior to the completion of the early-retinal-potential phase triggered by the flash (Kirschfeld 1983), which is



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

well before the impact of the flash on the reduction of neurotransmitter (glutamate) released by the photoreceptors. Thus, the moving object will have stimulated an array of photoreceptors even before the output of the first stimulated photoreceptor in the array.

If the onset of the object moving at velocity $v$ occurs in position $x_0$, the finding in the flash-initiated condition is that, at the instant the flash is perceived, the moving object is seen in position corresponding to $x_0 + \Delta\sigma$. In the primate visual pathway, a directionally selective response is first observed in V1 (Hubel & Wiesel 1962). Input from two neighboring presynaptic neurons is sufficient to trigger a directionally selective response in a cortical simple cell (De Valois & Cottaris 1998). In humans, successive stimulation by a moving object of two neighboring foveal receptors separated by 2.8 μm will occur in a very small time window; for example, within 0.9 msec for an object traveling at 10 deg/s (Fein & Szuts 1982). Hence, it may be argued that for objects moving at 10 deg/s, early neurons will have acquired an adequate motion sample following motion onset in just under one millisecond. This estimate is in approximate agreement with previous estimates (Anderson et al. 1990; Westheimer & McKee 1977). Note that this does not imply that a full response to motion (e.g., a directional response) occurs in less than 1 msec, but simply that 0.9 msec after motion onset, there is motion information in the pathway capable (later) of a directionally selective neural response. This information could feed either directly into a motion extrapolation mechanism or indirectly, after the computation of a directional response.

***5.2.1.2. Spatial extrapolation.*** The second "spatial extrapolation" stage involves speed of communication of neural signals from one retinotopic site to a neighboring one. A flash-lag effect of magnitude 0.5 degrees of visual angle corresponds approximately to 0.15 mm of retinal surface. Frequently cited speeds of signals traveling along neural pathways in the nervous system range from 0.2 to 120 m/sec. Lateral communication between two neighboring retinal locations separated by 0.15 mm, via neural signals traveling even at the slowest speed in this range of 0.2 m/sec, would occur within 0.75 msec. A similar scenario holds for other retinotopic maps, for example, a cortical map, after accounting for cortical magnification.

According to the estimate[1] provided here, for an object moving at 10 deg/sec, the required input sample and the resulting extrapolated output will together take only 1.65 msec. This time is less than 2% of the 100 msec baseline processing time (De Valois & De Valois 1991). Because the required time for spatial extrapolation is so small, this represents an extremely efficient mechanism that could start and finish virtually anytime during the required baseline (≈100 msec) delay after motion onset (Khurana & Nijhawan 1995) (see Fig. 7).

To further clarify, the suggestion here is not that very early mechanisms accomplish extrapolation, or that within 1.65 msec after motion onset extrapolated output is already available. It is highly unlikely that extrapolated output is available before the hyperpolarization response of the photoreceptors, which has a latency of more than 10 msec. Rather, the suggestion is that the time taken for an extrapolated

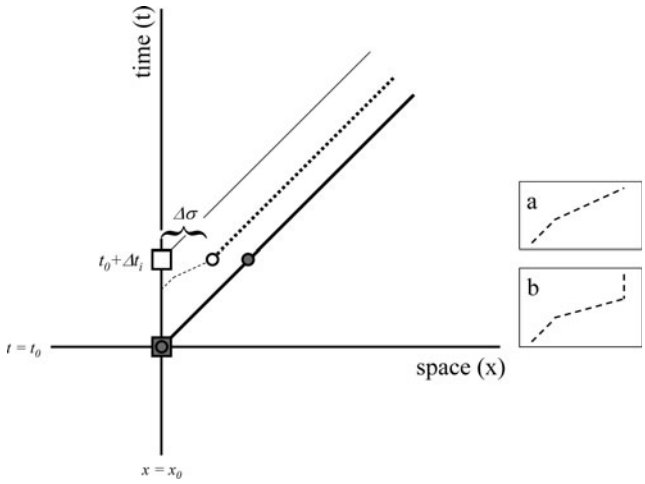



Figure 7. On the spatial extrapolation view, the latency for the moving object and the flash, in the flash-initiated condition, is $\Delta t_i$ (input delay). The figure shows how this model applies when the motion trajectory before the flash is missing, as in the flash-initiated condition. The reduced ds-error line stops before intersecting the $x = x_0$ line as the flash-lag effect occurs in the flash-initiated condition. The extrapolation model brings the moving object to the correct position on the reduced ds-error line by two processes: *sampling of motion* and *spatial extrapolation*. These two processes correspond to the thin broken line consisting of two segments of different slopes, enlarged in inset "a". The first segment, parallel to the space–time plot of physical motion, depicts a motion sample taken by neighboring neurons in a retinotopic map. The height at which this segment intersects the $x = x_0$ line is arbitrary and will depend on whether an early or a late sampling is assumed. The second segment of greater slope represents the travel of neural signals along the horizontal visual pathway, which can occur at higher speed. Inset "b" is made of three segments of different slopes. It depicts a situation of "over-extrapolation" that temporarily brings the extrapolated object closer to the thick continuous (physical) line. The object is brought back to the reduced ds-error line, as neural processes receiving the over-extrapolated input are presumed not capable of compensating for delays (see vertical segment in inset "b", and see text).

output is only a small fraction of the time longer than the baseline delay for an "un-extrapolated" output. The added segment of horizontal pathway responsible for the lateral shift in coordinates can be interjected at either an "early" (e.g., at the level of the retinal ganglion cells) or a "late" cortical level. Two further points are worth emphasizing. First, the time required in the generation of an extrapolated representation is so short that visual mechanisms could "over-extrapolate" and place the neural representation of the moving object farther ahead of the reduced ds-error line (see Fig. 4) with little computational time cost. There is already some evidence for over-extrapolation (Berry et al. 1999; Khurana & Nijhawan 1995; and see further on in the target article). Over-extrapolation could be useful, for example, if time-consuming visual processes further downstream, to which the extrapolation mechanisms report, cannot compensate for delays (see Fig. 7, inset "b"). Second, on the present model, further spatial shift than that depicted in Figure 7, for further reduction or even complete elimination of the ds-error, is possible. Such an outcome would be impossible with the differential latency view (Purushothaman et al. 1998; Whitney &



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Murakami 1998), as this would require motion to be processed with close to zero latency.

**5.2.2. Flash-terminated condition.** The foregoing discussion leaves little doubt that an undiminished flash-lag effect in the flash-initiated condition is not incompatible with the proposal of visual prediction. I now address another major empirical challenge for visual prediction resulting from the flash-terminated condition, in which the moving object disappears simultaneously with the flash. At the outset, spatial extrapolation should continue beyond the time of unpredictable disappearance of the object. Yet these displays produce no flash-lag effect.

On the spatial extrapolation view, vision contributes to compensation for neural delays whenever a moving object is viewed; the flash has only a role of a spatiotemporal reference. So, the correct space–time trajectory for the moving object is the reduced ds-error line. However, why then does the moving object not appear to overshoot its disappearance point in this condition (Fig. 8)?

It is commonly assumed that the full-motion condition of the flash-lag display is equal to the flash-terminated condition plus the flash-initiated condition. Although this relationship is correct for describing methodology, neural responses to the full-motion condition are *not* equal to the neural responses to the flash-terminated condition plus neural responses to the flash-initiated condition. For example, because of the visual transients generated by the offset of the moving object, the moving object in the flash-terminated condition stimulates the visual system in a dramatically different manner than in the full-motion condition. The transient generates a strong retinal neural representation that opposes the previously set-up cortical representation supporting spatial extrapolation. It is proposed that the newer representation suppresses the expected "predictive-overshoot" of the moving object and acts like a "correction-for-spatial-extrapolation."

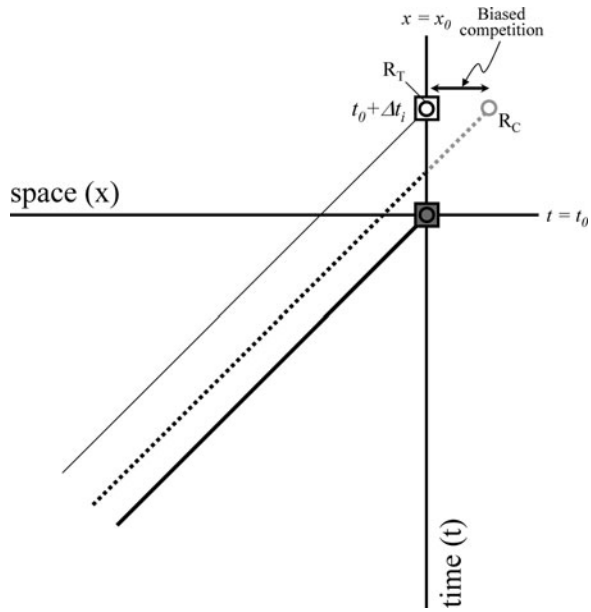



Figure 8. In this space–time plot, the moving object stops at position $x_0$ at time $t_0$, simultaneously with the presentation of the flash presented in position $x_0$. The fully delayed ds-error line (from the standard view) stops at the $x_0$ line, which is consistent with the result of "no effect" in the flash-terminated condition. However, spatial extrapolation can also explain this result. On this view, the cortical representation of the moving object ($R_C$), which is only partially maintained by retinal input, corresponds to the reduced ds-error (thick broken) line. Once the physical object is switched off, $R_C$ quickly decays (shown by lighter thick broken line) as retinal input stops. The switching off of the moving object triggers transient signals, which create a strong thalamic representation ($R_T$). Because of biased competition, and because $R_T$ is newer input, $R_T$ wins and dominates $R_C$ and the percept, overwhelming any predictive-overshoot of the moving object.

The above proposal is inspired by the notion of "biased competition" in which two neural representations compete for the control of cell activity (Desimone 1998; Desimone & Duncan 1995). The assumptions of biased competition are: that competing representations typically suppress each other; that competition can occur between two representations at a given level, or between low-level (e.g., retinal) and high-level (e.g., cortical) representations; and that the representation that wins the competition could be "stronger" for one or more of several reasons, including its recency, novelty, and behavioral significance to the animal (Desimone 1998). A related notion has been invoked recently to explain the perceptual phenomenon of visual masking (Keysers & Perrett 2002). On the spatial extrapolation view, after some initial motion input, the neural activity representing the moving object is less dependent on (external) retinal input, and is supported more by partially "autonomous" (internal) neural representations. These highly processed cortical representations feed into representations that control limb movements and typically control actions directed at moving objects. These representations could, however, produce their own localization errors if things in the world abruptly changed. Nonetheless, as these representations are based on prediction, new retinal input due to unpredictable change is given a higher relative weight by the visual system. In addition, greater weighting is assigned to retinal input, as stimulus offsets strongly stimulate the retina (Macknik et al. 2000). It is suggested that localization errors (predictive-overshoots) are suppressed by the strong, transient-fed, retinal representation generated by motion offset. Thus, for example, in the competition between a high-level cortical representation ($R_C$) and a low-level thalamic representation ($R_T$), $R_T$ dominates (Fig. 8). This leads to a "correction-for-extrapolation" via feedforward competition (Desimone 1998; Keysers & Perrett 2002). (In section 9.1 of the target article, I discuss interactions between non-competing representations.)

The hallmark of competition models is that, when strengths of the competing representations are changed, the representation that dominates will change (Keysers & Perrett 2002). Indeed, there are psychophysical results that produce the overshoot of moving objects predicted by the extrapolation model (Freyd 1987; Fu et al. 2001; Kanai et al. 2004; Whitaker et al. 1998). Interestingly, in one study an overshoot of the moving object was not found in human psychophysics, but was revealed in the monkey visual area V4 (Sundberg et al. 2006). The authors conjectured that processing at higher cortical areas beyond V4 must be engaged to fully account for the lack of overshoot in human psychophysics. I suggest an



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

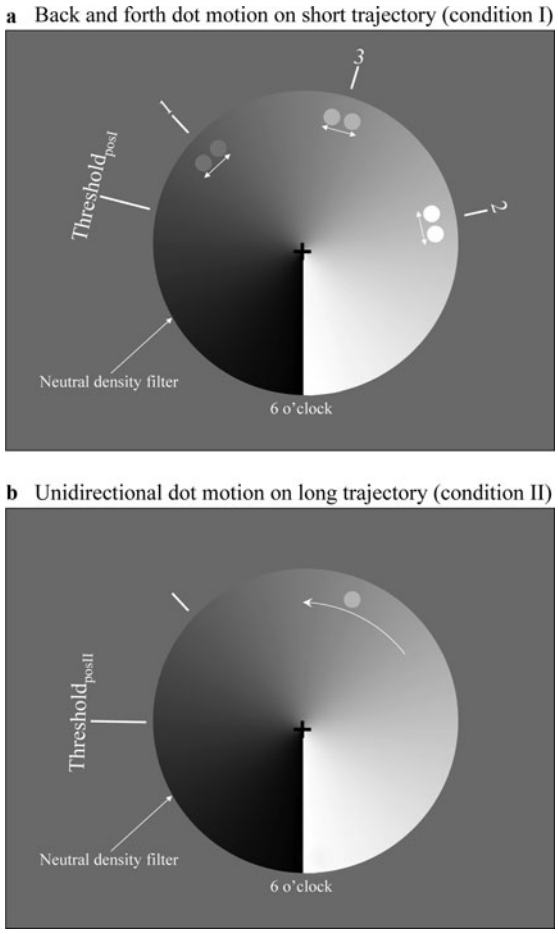



Figure 9. (**a**) The gray background is used only for illustration; the actual experiment was performed with a homogenous dark background. A neutral density filter, which decreased in transmission from 100% (at the 6 o'clock position) to 0.01% in the counterclockwise direction, was used. Thus, a moving stimulus (a white dot of fixed luminance) behind the filter appeared to vary in luminance; it was always visible in the 6 o'clock position and never visible past about the 9 o'clock position. In condition I, on different trials, the dot moved back and forth on a short trajectory behind the filter. The average position of the short trajectory dot was changed from trial to trial (figure depicts a trial sequence of trials 1, 2, and 3). A reference line (white radial line segment of shorter length next to the filter's edge) was presented on each trial. Participants said "yes" or "no" depending on whether they saw the dot at the position of the reference line or not. Detection thresholds were determined ($Threshold_{posI}$, schematically depicted by the longer line segment). At $Threshold_{posI}$, "yes" and "no" responses were equally likely. (**b**) In condition II, the same white dot moved on a long trajectory from the 6 o'clock position into the filter's denser part until it disappeared, and again around the 9 o'clock position. A reference line was presented (white radial line segment of shorter length). For this task, participants said "ahead" or "behind" depending on whether they perceived the dot as disappearing before it reached the reference line or after it had passed it. From trial to trial, the position of the reference line was changed to determine the disappearance threshold ($Threshold_{posII}$, schematically depicted by the longer segment of radial line). At $Threshold_{posII}$, "ahead" and "behind" responses were equally likely. Our main finding was that $(Threshold_{posII} - Threshold_{posI})/v = 175$ msec, where $v$ is dot velocity. Adapted from Maus and Nijhawan (2006).

alternative point of view: Lower-level representations, possibly thalamic, compete with higher-level representations such as found in V4. This competition leads to the suppression of extrapolated V4 output.

Recently, Maus and Nijhawan (2006) empirically tested the afore-discussed "correction-for-extrapolation" hypothesis (Fig. 9). In the flash-terminated condition, the disappearance of the moving object is an all-or-none event. We weakened the retinal representation (and its thalamic counterpart $R_T$) by replacing the all-or-none event by a luminance ramp. The correction-for-extrapolation hypothesis predicts that in the absence of a strong $R_T$ competing with $R_C$ (cortical representation), $R_C$ supporting the moving dot's extrapolated position will be revealed in perception. Indeed, this was the case. The experiments (Maus & Nijhawan 2006) revealed a forward extrapolation of $v \times 175$ msec, where $v$ is moving object velocity.

The main empirical challenges for spatial extrapolation were based on the flash-initiated and flash-terminated conditions. As has been shown, an unabated flash-lag effect in the flash-initiated cycle is not inconsistent with visual prediction, as neural representations supporting extrapolation can be set up quickly in relation to the input delay for motion. In fact, the extrapolation times may be so small that when required, some processes could "over-extrapolate." Finally, the "missing predictive-overshoot" in the flash-terminated condition can be accommodated by the biased competition model (Desimone 1998), which suggests a correction for spatial extrapolation caused by a strong retinal input due to transients. This new analysis, based on empirical results, reinforces the general proposal that when possible, the visual system attempts to minimize localization errors. Both spatial extrapolation and the correction for it achieve this goal.

## 6. Further challenges to the notion of visual prediction

### 6.1. Luminance contrast dependent variation in the flash-lag effect

It is known that luminance contrast can influence visual delay, which in turn can influence the perceptual localization and shape of moving objects, as has been shown by several phenomena. The Pulfrich pendulum effect occurs when the observer views a moving pendulum bob with one eye covered by a neutral density filter. As a result of reduced contrast in one eye, the observer perceives the motion of the swinging bob in a plane as movement along a curved path in depth (Burr & Ross 1979; Morgan & Thompson 1975). This finding can be explained in terms of longer visual processing delays for the eye with the filter, which receives lower image contrast relative to the other eye. The effect of variations in luminance contrast on perceptual distortions in the shape of moving objects has also been shown (Roufs 1963; Williams & Lit 1983; Zanker et al. 2001). In addition, the effect of luminance contrast on localization of moving objects has been studied using the flash-lag effect (Purushothaman et al. 1998). Purushothaman et al. (1998) found that value of $\Delta\sigma/v$ changed from 20 to 70 msec as the luminance of the moving object was increased (with luminance of flash held constant) by 1 log unit above detection



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

threshold. Zanker et al. (2001) found a luminance dependent shape distortion for moving objects corresponding to 3 msec. In order for visual compensation to be effective, it must be flexible, accounting not just for the visual delay but also for variations in it. Since the flash-lag magnitude varies with luminance (Purushothaman et al. 1998), this has been taken to imply that the visual system is not able to take account of the variations in visual delays. This has led to the suggestion that visual processes do not contribute to compensation for neural delays (Purushothaman et al. 1998; Zanker et al. 2001).

This reasoning, however, is questionable. Consider a related question: How are actions such as catching or hitting a ball affected by a change in a ball's luminance contrast? Depending on the degree to which luminance contrast varies, performance will either be affected or not (Anstis et al. 2000). Suppose the change in luminance contrast is large enough that the catching ability of a player is affected; that is, at luminance contrast level 1 the player is able to catch the ball, but at luminance contrast level 2, he or she is unable to catch the ball. If we follow the reasoning that luminance contrast–dependent modulation of the flash-lag effect implies that the visual system does not compensate, then the luminance contrast–dependent modulation in behavior would suggest that the nervous system (as a whole) does not compensate for visual delays. This is clearly mistaken reasoning because it is all but certain that the nervous system (as a whole) cannot function without compensation for neural delays. Hence, variation in performance, measured either visually (e.g., via the flash-lag effect) or through action, cannot be taken to indicate that there is no compensation for visual delays; variation in performance simply means that the output of the system changes when conditions change. These considerations lead to what I call *the fundamental assumption of compensation for visual delays*,[2] which states: "In the absence of mechanisms compensating for visual delays, many behaviors in otherwise healthy animals would be disrupted."

A further question, however, arises. If there is a visual compensation mechanism, then what might be the range of environmental conditions (and/or conditions internal to the animal) within which the mechanism produces an invariant output? All known neural mechanisms have limitations. Consider the well-known size constancy mechanism. This mechanism functions well (i.e., produces an invariant size estimate) only within a limited range of distances – imperfect size constancy for distances outside the range does not mean that size constancy does not occur at all. What might be the range of luminance contrasts for the moving object over which spatial extrapolation might produce an invariant output? It is premature to give a definitive answer to this question. In the study of size constancy, zero distance is the natural "anchor point." One of the difficulties is determining what might be the anchor point for luminance contrast. There are three possibilities: (a) the stimulus in an "Off" state; (b) the observer's absolute detection threshold; or (c) some "average" photopic luminance level in which the observer's visual system evolved.

It is well known that in sports, for example, degradation of performance occurs outside certain limits of ambient illumination; play is sometimes stopped because of "bad light." From failures of performance meeting task criteria (e.g., if one is forced to play in bad light), however, one cannot conclude that compensation for visual delays does not occur when light is "good" – this would violate the fundamental assumption of compensation for visual delays – nor can one conclude that breakdown of performance reflects no compensation for visual delays; behavior can break down because of partial, or inappropriate, compensation. The only conclusion one can draw from behavior not meeting task criteria is that compensation is not appropriate for the total (sensory + motor) delay in the system.

In producing goal-directed behavior, the flexibility of the animal's nervous system as a whole would be expected to be higher than the flexibility of the visual system per se. This is simply because the activity of single neurons located relatively late in the sensory-motor stream, for example, in the spinal cord, can be influenced by many more neurons, than can the activity of neurons located early in the visual pathway. Indeed, spinal motor neurons could receive input from more than two thousand other neurons (Poritsky 1969). This suggests that if the visual system contributes to compensation for neural delays, then this contribution may be less flexible in accommodating for luminance contrast variations than the animal's nervous system as a whole. Therefore, for different visual delays, compensation could be adjusted dynamically between visual and motor processes. For example, for low luminance contrast, where performance still meets task criteria, the compensation could be carried out more by the motor than by the visual processes. This suggestion is consistent with the results of Purushothaman et al. (1998) and Zanker et al. (2001), and reconciles the contrast-dependent variation in the flash-lag effect with the notion of visual prediction.

### 6.2. Logical challenges to visual prediction

Compensation for all neural delays can be carried out by late components of the sensorimotor pathways. As suggested by Purushothaman et al. (1998), "accurate visually guided motor actions are likely to depend on motor instead of perceptual compensation" (p. 424). There is unequivocal evidence for compensation mechanisms that are located within the non-visual parts of the nervous system (Lacquaniti & Maioli 1989). The view challenging the role of visual prediction and assigning a compensatory role only to non-visual mechanisms seems parsimonious. Hence, it has been suggested that visual prediction is unusual in that, on this proposal, the visual system somehow attempts to compensate for its own input delays. In fact, Schlag and Schlag-Rey write: "Clearly, this is one of the most daring proposals of a top-down hypothesis. Here, 'top–down' means that somehow the brain is attempting to correct its own input from the senses" (Schlag & Schlag-Rey 2002, p 197).

On the *fundamental assumption of compensation for visual delays*, the delays must be compensated somewhere within the nervous system. For concreteness, suppose that the visual delay for a moving ball between the photoreceptors and the LGN is 50 msec, and the total delay (for which the nervous system must compensate) involved in catching the ball is 200 msec. If the 50 msec delay is not compensated, and if the required precision in performing the catch is greater than 50 msec, which is common in even moderately demanding tasks (Tresilian 1993), then



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

the animal's behavior will not meet the task criteria. If, on the other hand, the animal is able to produce behavior that meets the task criteria, then it follows that if compensation for 50 msec (visual delay) is not carried out in the vision pathway, then it must be carried out in later pathways before neural stimulation of muscles. If one assumes only non-visual compensation, then one must also assume that the non-visual part of the nervous system has information concerning the 50 msec delay, and that the nervous system as whole is attempting to correct for the visual delay. Thus, the criticism that seemed specific to visual prediction in terms of "correcting its input from the senses" (Schlag & Schlag-Rey 2002) is not specific at all, and would apply to any and in fact all predictive mechanisms within the CNS.

Furthermore, there are clear examples in which the CNS does monitor its own input delays. Variables such as muscle fatigue impact the force a muscle can generate in response to a given neural signal, so the CNS must monitor such variables in order to produce a desired trajectory of limb movement (Bullock 2003). Therefore, mechanisms compensating for delays in the motor system require information not only about delays resulting from the sluggish response of muscles to neural stimulation, but also about variations in these delays caused by varying degree of muscle fatigue. Helmholtz's efference copy notion suggests another example. In order to maintain a motionless visual scene during voluntary eye movement, the efference copy of the motor command and the visual feedback must be available for comparison within some time window. As change in luminance contrast changes the retinal input delays, the fact that we perceive visual scenes as remaining stable under a wide range of luminance contrasts already supports the view that the visual system is able to monitor its input delays.

**6.2.1. Prismatic adaptation and visual prediction.** A related challenge is posed by adaptation experiments involving optical distortions (Harris 1963). Humans can adapt to optical displacements produced by wedge prisms. Upon first wearing prism goggles, the wearer will incorrectly reach for objects, but later he or she will adapt and reach accurately. Such results might suggest that errors in visual localization, whether resulting from optical or neural factors, can be dealt with by late sensory-motor mechanisms, and that compensation for delays in vision per se is unnecessary for interceptive or reaching actions.

There are, however three issues that need further attention. First, there is debate as to whether the adaptation observed in displaced/inverted vision experiments is visual or motor (Harris 1980). To the degree that adaptation is at the level of vision, the notion of visual prediction would be supported. Second, adaptation to inverted vision has been tested after days or, at the most, weeks of the subject's visual experience while wearing prism goggles. Although it is not known what the state of the visual adaptation would be if the experience with displaced vision was lifelong, the expectation is that adaptation would be more complete with longer experience. Following his famous experiments, Stratton reported that the difficulty of seeing things as being upright with upright retinal images (as opposed to inverted retinal images) seemed entirely due to the "resistance offered by long-established previous experience" (Stratton 1896). There are many reports, including those from Stratton in the 1890s and Kohler in the 1950s and 1960s, indicating that with time, things appeared more "normal" and, therefore, that there was an adaptation at the perceptual level (Harris 1980). Clearly, with time, the observer's interaction with visual objects improved, so over time the correlation between normalcy in appearance and improved interaction with things suggests that perception does impact action. I argue (see further on) that the improved level of perceptual adaptation feeds into more accurate action, and more accurate action feeds (via reafference) into improved perception. A lifetime of experience with neural delays should no doubt render the visual adjustment complete, and so in this case, action may be influenced strongly by perception. Finally, there are clear examples of some animal species that do not adapt at all to displaced vision (Hess 1956). Such animals under displaced vision conditions would starve to death even though food was nearby. Therefore, if one were to follow the reasoning that ready adjustment in humans to optical distortion suggests late compensation, then, by the same token, a lack of adjustment would imply compensation at the early stage of visual processing. Because in evolution older systems are modified, not discarded, the lack of adaptation in some older systems suggests that some form of visual compensation, perhaps less flexible than motor compensation, may also exist in more recent systems such as humans.

## 7. Functions and mechanisms of spatial extrapolation

We have so far considered the effect of extrapolation on perception. This is indeed appropriate, as the one empirical basis of the current viewpoint is psychophysical. Although extrapolation carried out by visual mechanisms could impact behavior directly with no perceptual consequences, here I consider only processes that can and do reveal themselves perceptually. Therefore, the discussion here will keep in view the human flash-lag results. Multiple retina-, eye-, head-, and body-centered coordinate systems that encode locations of objects for various functions exist within the CNS. One encounters the issue of neural delays and compensation not just when these coordinate frames are static, but also when they are moving, as, for example, when the eyes or the head track a moving object. Although this situation can be incorporated within the general framework presented here, detailed analysis of neural delays for moving reference frames is beyond the scope of this article – the reader is directed to previous empirical work on moving reference frames (Brenner et al. 2001; Nijhawan 2001; Schlag & Schlag-Rey 2002). A forthcoming article considers the consequences of delay compensation when the whole animal moves in relation to static objects (Changizi et al., in press).

The important function of compensation for neural delays is likely to be carried out by multiple spatial extrapolation mechanisms at different levels of processing; the requirements for spatial extrapolation of coordinates are very simple, and mechanisms that could potentially accomplish extrapolation are present at virtually all levels



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

of neural processing. However, the immediate goal of the mechanisms at different levels may not be exactly the same. In addition, these mechanisms may be variants of one underlying process or could be significantly different from each other. I now consider these mechanisms.

There is a natural link between mechanisms of spatial extrapolation and their function. For example, a mechanism serving a high-speed response may be different from a mechanism for which speed is less relevant. The simplest function of visual mechanisms compensating for delays would be to remove errors in localization and support rapid behaviors such as orienting toward moving objects. Mechanisms serving this function might be located early in the visual pathway. Intermediate-level mechanisms also may serve the function of perception of objects and of planning of more complex actions that involve coordination of gross and fine bodily movements over a longer time frame. Finally, late mechanisms of visual prediction may be related to visual feedback to the animal about successes or failures of actions that have already been performed. The early-to-late hierarchy mimics the literal-to-abstract representation of space in the early to late representations in parietal and frontal cortical areas (Andersen et al. 1993). The distinction between early versus late compensation mechanisms may also be understood in terms of the requirements of speed, accuracy, and flexibility of actions; early mechanisms may feed into fast relatively stereotyped actions, whereas late mechanisms might feed into actions in which variability in sensory and motor delays is likely to be high, or in which information about delays in the feedforward vision pathways is less reliable.

The suggestion that compensation mechanisms may be ubiquitous within the CNS is not without support. There exist a great variety of potential mechanisms capable of spatial extrapolation (see, e.g., Baldo & Caticha 2005). The required lateral shifts in coordinates could be based on shifts in peaks of population response already shown in the retina (Berry et al. 1999) and in V1 (Jancke et al. 1999); shifts in receptive fields, or on the basis of lateral spread of neural facilitation (Erlhagen 2003; Grzywacz & Amthor 1993; Kirschfeld & Kammer 1999; Sillito et al. 1994), and other forms of distortion of retinotopic representations (Sundberg et al. 2006). Such mechanisms are known to exist not just in the visual system, but also in the auditory system (Witten et al. 2006). Other, more specialized, computational schemes, such as "shifter circuits," can also produce a horizontal shift in neural array activity on the basis of motion information (Anderson & Van Essen 1987). Finally, there are mechanisms located in higher cortical areas or the cerebellum that integrate information from multiple sources. These mechanisms fall in the general category of internal models (Kawato 1999).

### 7.1. Prediction in feedforward vision pathways serves perception for online control of behavior

The goal of visual prediction cannot be solely to inform perception. Rather, visual prediction must impact the behavior of animals and ultimately contribute to a survival advantage. Barlow has argued that significant recoding of sensory information occurs within the retina itself (Barlow 1979), and that retinal processes per se contribute to transformations of information in preparation for use during behavior. Indeed, mechanisms for spatial extrapolation have been discovered in the retina (Berry et al. 1999) and in other early stages of visual processing (Barlow 1953). Berry et al. (1999) used isolated retinas of the salamander and the rabbit in their studies. A moving or a flashed bar was projected on the retina, and spiking activity from a population of ganglion cells was recorded simultaneously. The cells responded with a latency of about 50 msec to the flash. However, the peak of the population cell response to the moving bar appeared close to the leading edge of the bar, instead of trailing behind the bar's center by 50 msec. (Note that full compensation requires the population response to be aligned with the bar's center. The forward shift of the population response in relation to the bar's center is noteworthy, and will be further considered later.) This extrapolation of the response of the ganglion cells may be explained in terms of spatio-temporal filtering, biphasic response of cells, and contrast gain control (Berry et al. 1999). These mechanisms are not unique either to lower species or to the early parts of the vision pathway; analogous mechanisms exist at higher levels of processing in primates (Shapley & Victor 1978), which may explain the flash-lag results in humans.

If one assumes that online behavior is controlled by perception of events (Davidson 1970), then in humans the simplest possible function of extrapolation in the feedforward visual pathway, say, between the photoreceptors and V1, would be to aid interaction with moving objects. On this view, extrapolation reduces latency-based spatial error in the perceived position of moving objects such as tennis balls, which in turn aids successful interceptive actions such as hitting tennis balls. This possibility, however, cannot be entirely correct, as humans can perform goal-directed actions without awareness of the position of the visual objects (Goodale et al. 1986). Thus, it cannot be that the perception of the true position of objects is necessary for action. The general issue of whether perception is necessary for action has bearing on this point.

The vision pathway consists of two parallel streams. The magnocellular stream consists of larger, faster-conducting neurons; these neurons appear to be specialized for processing spatial and motion information. The parvocellular stream consists of smaller, slower-conducting neurons. These neurons seem more specialized for processing of object and color information. The functional specialization of the two streams, which starts with parvo- and magnoretinal ganglion cells, continues beyond V1 as the dorsal and ventral pathways (Mishkin & Ungerleider 1983). One point of view argues for substantial functional independence between the two streams captured by the dichotomy of "vision for perception" versus "vision for action" (Goodale & Milner 1992). The suggestion is that the dorsal stream provides visual information to action systems, whereas the ventral stream provides visual information for perception (Goodale & Milner 1992). Support for this position comes from findings in humans that damage to the dorsal stream produces a deficit (optic ataxia) in the action system but not the perception system, whereas damage to the ventral pathway impairs perception and recognition of objects but spares the ability to produce correct actions.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

If perception and action were based on visual processes that were completely isolated in independent pathways, then one might be forced to come to the conclusion that perception does not affect action systems, and so compensation for delays at the perceptual level cannot contribute to adaptive behavior. There are, however, virtually no researchers who believe there to be a complete segregation of visual processes for perception versus action. For example, Goodale and Milner (1992) have argued that the visual phenomenon of size constancy plays a role in the scaling of grasp size. Therefore, it may not be unreasonable to expect that the delay-compensated perceived position of moving objects, which is a type of constancy phenomenon (see further on), does play a role in interceptive actions.

### *7.2. Internal model for visual prediction*

Much of the flash-lag debate has centered on the interpretation of the effect given in section 7.1; that is, spatial extrapolation is caused by visual mechanisms in the feedforward vision pathway serving perception. I now develop a generalized concept of visual prediction that assumes that spatial extrapolation is *not* based on mechanisms in the stream that directly feeds perception, that is, the parvocellular stream. A modular approach is adopted, with multiple mechanisms located at different levels within the CNS. The picture that emerges is consistent with the notion of internal models that have been useful in describing both motor and sensory systems. In outlining the internal model for visual prediction (IMVP), I begin by describing the various processes and component mechanisms that make up the IMVP.

**7.2.1. Prediction in feedforward vision pathways serves online control of behavior: Impact of "motor maps" on perception.** There are psychophysical (Nijhawan 1997) and neurophysiological (Sundberg et al. 2006) reports of color-specific effects that are easiest to explain in terms of spatial extrapolation within the ventral stream. Therefore, mechanisms for extrapolation may be located in both vision streams. However, if one assumes that the ventral stream is mainly responsible for visual awareness (Goodale & Milner 1992), then an extrapolation mechanism cannot be located only in the ventral stream. This is because goal-directed actions can be performed without awareness of the stimulus, as revealed by the pathological phenomenon of blindsight (Stoerig & Cowey 1997; Weiskrantz 1996) and in normal observers (Milner & Goodale 1995). For simplicity, and for the purposes of addressing the debate directly, I assume that visual extrapolation mechanisms are located only in the sensory pathways feeding motor processes for control of behavior. Cortical and subcortical areas receiving sensory input that function mainly to control behaviors such as orienting and reaching towards, or withdrawing from moving objects, are common in the CNS.

***7.2.1.1. Compensation for lag of moving objects in motor-oriented reference frames.*** Figure 1 depicted the effect of visual delays on a wave of neural activity in a retina-centered cortical frame and on human perception. In order to execute fast interceptive actions, the CNS requires not just spatial but also temporal information. I first consider neural transformations that provide action systems with spatial information. Visual input is initially encoded in a retina-centered frame. Interaction with objects requires that the location of objects be transformed from a sensory-based representation into a muscle-centered representation. Intermediate eye- and head-centered maps contribute to such transformations. In the lateral intraparietal area and the parietal reach region of the cortex, there are maps for controlling eye movements and reaching movements, respectively (Andersen & Buneo 2002). The primary function of such maps is to specify the position of stimuli for action (Taira et al. 1990).

In addition to cortical maps, there are many subcortical retina-centered frames, for example, in the primate superior colliculus (SC), which serve mainly a motor function (Cynader & Berman 1972). Findings of various studies suggest that the SC represents a "motor map" of locations for the guidance of orienting responses such as saccades and head movements (DuBois & Cohen 2000; C. Lee et al. 1988; Schiller 1984; Sparks & Jay 1986; Sparks et al. 1990). Many SC cells respond to moving visual stimuli (Dean et al. 1989). An uncompensated delay in the retina to the SC pathway will, for example, cause a saccadic target in motion to lag in the retina-centered SC frame. This lag, in turn, would affect the direction, amplitude, and velocity of saccades, impeding the animal's ability to rapidly orient toward a moving target that might be of imminent interest or danger to the animal. This suggests a strong basis for compensation of visual delays in the retina–SC pathway. Indeed, mechanisms for extrapolation have recently (Witten et al. 2006) been uncovered in the owl's optic tectum (homolog of the mammalian superior colliculus).

There is also a wealth of data related to how the CNS obtains timing information for potential action. In order to hit a fast-moving tennis ball (say), the player's racket must arrive in the correct position at the correct time; so in addition to position information, the action systems of the hitter require precise timing information. One related line of research is based on the time-to-collision (TTC) paradigm (Bootsma & Oudejans 1993; D. N. Lee 1976; D. N. Lee & Reddish 1981; Regan 1992; Tresilian 1993; Wagner 1982). Gibson proposed that information for acting on the world is contained in the stimulus array in the form of invariants, and can be directly used by the animal without recourse to internal representations (Gibson 1961). D. N. Lee (1976) applied Gibson's approach to situations in which animals encounter moving objects. For an approaching object, TTC can be computed on the basis of the ratio of retinal image size divided by the rate of expansion of the retinal image. Recent evidence suggests that computation of TTC is considerably more complex than previously expected (Tresilian 1999), however, the basic premise of much of the TTC research is that the CNS uses visual information directly to predict the time of collision. In particular, this point of view rarely considers internal representations and perception as necessary for action. The evidence that optical information is used in the computation of TTC in lower species, such as pigeons (Wang & Frost 1992), gannets (D. N. Lee & Reddish 1981), and houseflies (Wagner 1982), underscores this point.

Speeded actions directed at moving targets depend on precise spatial information. The TTC approach does not



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

explicitly address where or how visual delays, which would contribute to spatial errors, are compensated. I argue that a corollary of the TTC approach is that visual mechanisms compensate for visual delays. Consider, for example, experiments revealing the ability of humans to estimate the future time of arrival at a predetermined position of visual objects traveling at constant speed. Observers are accurate to within 10 msec in estimating arrival times of objects at positions that are more than 2000 msec in the future (Bootsma & Oudejans 1993). As the visual processing delay is much longer than 10 msec, it is clear that in order to perform such tasks, this delay must be compensated. In addition, visual processing delay is much smaller than 2000 msec. Thus, the mechanisms that purportedly predict a moving object's position 2000 msec into the future can certainly predict the object's position over a much shorter time, corresponding to a visual delay of, for example, 100 msec. Therefore, it is implicit in the TTC approach that the CNS uses optical information to compensate for visual delays. This compensation, however, may be temporal rather than spatial in nature. Thus, the compensation may not involve an explicit spatially extrapolated representation.

Frost and colleagues (Sun & Frost 1998; Wang & Frost 1992) have found evidence of neurons that compute TTC in the nucleus rotundus of pigeons (the homolog of the pulvinar in the mammalian thalamus), a major midbrain area receiving input from the optic tectum. Interestingly, these researchers did not find evidence of neurons that compute TTC in the optic tectum (Wang & Frost 1992). Since mechanisms for spatial extrapolation have been found in the optic tectum (Witten et al. 2006), this suggests the interesting possibility that neurons in the optic tectum spatially extrapolate sensory input, and that this extrapolated output is sent to nucleus rotundus for computations of timing.

***7.2.1.2. Prediction minimizes the interference of non-veridical percepts with online action.*** The present hypothesis suggests that compensation for visual delays is not carried out in the feedforward ventral pathway serving perception. What, then, is the function of extrapolation in perception revealed by the flash-lag, and how does it occur? I suggest that extrapolation in perception occurs in two steps. The first step is extrapolation carried out by feedforward visual processes for online control of action. This extrapolation is not directly revealed in perception. The second step is the communication of the extrapolated visual data to ventral processes at a later stage of the vision pathway. Thus, the perceptual consequences of extrapolation are there due to crosstalk between the dorsal and the ventral pathways at a late stage (Khurana et al. 2006).

The function of this mechanism is to remove conflict between object positions available to perception and action systems. Note that the lack of necessity of correct perceptual information for online action (as noted above) is one thing, but the perceptual system delivering non-veridical position information is quite another. This point is brought out quite revealingly when one reconsiders the effect of delays in the ventral stream on the perception of moving objects. In the absence of compensation, the perceived position of moving objects would be non-veridical (Fig. 1b). Thus, in light of everyday observations of successful interceptive actions, a suggestion that there is no perceptual compensation for delays would imply not just that perceptual compensation is unnecessary for online actions, but also that online actions are immune to non-veridical percepts of moving objects. Contrary to the expectation that perception may be irrelevant for action, non-veridical percepts (e.g., geometric illusions) have been shown to influence actions (Franz et al. 2000). Therefore, a behavioral advantage would be gained for visual mechanisms that remove errors in perceptual localization of moving objects.

**7.2.2. Prediction resulting from processes beyond feedforward vision pathways: Extrapolated perception caused by feedback from motor processes.** Early sensory mechanisms and extensions of these mechanisms found in later areas of the primate visual systems might explain the flash-lag effect in humans (Berry et al. 1999; Witten et al. 2006). There is a distinct possibility, however, that the extrapolation mechanisms revealed in the early processes in lower species feed only action systems. Thus, the connection between early mechanisms and human perception may be less straightforward than previously suggested (Gegenfurtner 1999).

The currently predominant view is that prediction exists mainly in the motor components of the nervous system (Wolpert & Flanagan 2001). In consonance with this view, I assume that extrapolation mechanisms exist in neither the action- nor the perception-oriented feedforward vision pathways, but are located in later brain areas serving action: Motor mechanisms receive delayed sensory data consisting a spatial error, extrapolate, and send the compensated information to the muscles. Motor extrapolation is carried out during actual interactions, or planning of interactions, or during imagined interactions (mental rehearsal). Extrapolation in perception then results either because of predictive motor processes interacting with visual processes, or because of a deep-seated unity between motor and perceptual processes (see further on). The interaction could be supported by feedback from motor areas to the visual areas, analogous to the efference copy notion of feedforward control, or by communication between motor and visual neurons within a given area of the CNS. On the feedback view, the extrapolated motor output is sent not just to the muscles, but a copy of this output is also fed back to visual cortical areas influencing perception. A strong case for the impact of feedback from higher cortical areas on perception has been made (Lamme & Roelfsema 2000). In addition, it is known that in humans, frontal cortical areas can modulate neural activity in visual areas (Sterzer & Kleinschmidt 2007). The function of the interaction between motor and visual processes is to provide the animal with visual feedback resulting from actual actions directed at moving objects. Without this mechanism, there would be an offset between the time–space of the performed action and the time–space of the visual feedback resulting from the action.

Similarities between motor extrapolation and its counterpart in perception may be understood along a different line. Many influential researchers have suggested that despite apparent differences, perception and action are two sides of the same coin (see, e.g., Sperry 1952). Evidence for this position comes from multiple sources. For example, the perception and production of vocal sounds



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

appear to be served by the same underlying neural structures (Liberman & Mattingly 1985; Williams & Nottebohm 1985); experiments requiring human subjects to both produce actions and view events relatable to those actions (Müsseler & Prinz 1996; Parkinson & Khurana 2007; Stürmer et al. 2000) have found similarities in results across the two domains; and neurons in many parts of the CNS respond both when the animal performs an action and also when the animal views an event relatable to that action. In one study, Port et al. (2001) recorded neural activity in the monkey motor cortex either while the animal acted to catch a moving object, or when it simply viewed the same object without performing an overt action. A significant proportion of investigated neurons responded both when the monkey performed the action and when the monkey simply looked at the moving object.

Many neuroscientists (Arbib 1972; Rizzolatti et al. 1997; Sperry 1952) have attempted to diffuse the sharp boundary between the sensory and the motor systems on theoretical grounds. During evolution, there is no "jump" that can be discerned between the nervous systems of primitive species, whose primary function is to produce overt responses conducive to survival, and of higher species, in which perceptual-cognitive systems seem to have become disengaged from action systems (Sperry 1952). On the single system view, one should not be surprised to find similar mechanisms performing similar functions in apparently different parts of the CNS. Parsimony suggests that if there are mechanisms compensating for delays within the motor system, then similar mechanisms are also likely to be present in the visual system, and vice versa (Nijhawan & Kirschfeld 2003).

It is worth noting that the flash-lag effect in the flash-initiated condition is better explained in terms of visual mechanisms in the feedforward vision pathways. Nonetheless, the late mechanism described here could modulate the extrapolated output of feedforward visual mechanisms. As discussed in section 5.2.1, the time required for extrapolation in the feedforward vision pathway is vanishingly small, so a visual mechanism could "over-extrapolate." There is, in fact, evidence that the feedforward pathway does over-extrapolate. First, the flash-lag effect in the flash-initiated condition is larger in magnitude than in the complete motion condition (Khurana & Nijhawan 1995). Second, Berry et al. (1999) found the population ganglion-cell response to be shifted further forward of the middle of the moving bar. (Note that precise compensation for delays would require the population response to be aligned with the center of the moving bar.) It is unlikely that early mechanisms could precisely compensate for the neural delays or deliver a reliable enough output for the animal to act on. Thus, over-extrapolation may be a result of "rough and ready" computation provided by the early visual pathway. For the purposes of action, the over-extrapolated output would then require modulation, which would be undertaken by the late mechanism described here.

**7.2.3. Forward models.** There are many types of internal models. Internal models of the motor system are "neural mechanisms that can mimic the input/output characteristics, or their inverses, of the motor apparatus" (Kawato 1999, p. 718). There is also evidence for internal models within sensory systems; one such model within the vestibular system functions to represent a stable orientation of the head with respect to the world (Merfeld et al. 1999).

There are two varieties of internal models: *forward models* predict the next state of the motor apparatus given the current state and motor command, and *inverse models* infer the motor command that caused a change in state on the basis of feedback from sensory data (e.g., vision and proprioception). Forward models are necessary when delays in the sensory motor loops are significant. A generalized internal model that combines prediction and sensory feedback to estimate the current state of the motor apparatus is known as the "internal forward model" (Wolpert et al. 1995). In a study testing the existence of this model, participants moved their unseen arm in the presence of a null, an assistive, or a resistive force. When required to estimate their arm position during the initial stages of movement, participants overestimated the distance by which their arm had moved (Wolpert et al. 1995). The overestimation error reduced during the later parts of movement, suggesting the engagement of a forward model when sensory feedback from the limb is relatively weak at the beginning of movement. (An alternative account of this result is considered later.)

Control and feedback are essential features of the internal models of the motor system. Muscles and limbs may be considered part and parcel of the internal model for the motor apparatus (Kawato 1999), as the cortex controls the movement of limbs, and the limbs in turn send sensory input to the cortex. An IMVP cannot have visual objects as its components. This is because visual objects are "external" entities that cannot be controlled by the internal model. Likewise, the IMVP cannot include the retina as its component. (There does not seem to be any evidence that feedback from higher brain areas can descend down to the retina.) However, such a model can include any or all components of the visual system up to and including the LGN. The LGN receives a massive number of descending fibers from the cortex, and is the lowest known level of the visual system at which cortical feedback can modulate ongoing neural activity (Cudeiro & Sillito 2006). Descending fibers affect the ongoing thalamocortical transfer of sensory information and may, indeed, serve a predictive function (Engel et al. 2001; Sillito et al. 1994).

I suggest that the concept of the internal model for the motor apparatus may be naturally extended to incorporate visual prediction, with the qualification that in IMVP sensory input has a different role than it does in the motor model. The suggestion here is that neural representations in the LGN are both stimulus-driven *sensory entities* (based on input from the retina) and *controlled entities* (receiving descending cortical input), just as the neural limb representations (not the limbs per se) are in the motor models. What distinguishes IMVP from the motor model is the retina-to-LGN segment, which cannot be controlled. In the motor model, there is no strictly sensory segment.

The present IMVP model has another unique feature. In the IMVP, extrapolation can occur within the feedforward vision pathway, which may provide a rough and ready over-extrapolated output (Berry et al. 1999; Khurana & Nijhawan 1995). This suggests an alternative account of Wolpert et al.'s (1995) results in which participants overestimated the distance by which their arm had



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

moved during the initial stages of movement. This result may be due to over-extrapolation in the proprioceptive input pathway, and not due to a greater engagement of a high-level forward model during the initial stages of movement. The overestimation is revealed during the early stages of movement because the late modulatory process that keeps over-extrapolation in check has as yet not been engaged.

### 7.3. Support for IMVP

Without doubt, the motor system contributes to compensation for delays during action (Lacquaniti & Maioli 1989). It follows that if the flash-lag has something to do with compensation for neural delays, then one should certainly observe a flash-lag effect for action. Such an effect has indeed been observed. If during voluntary limb movement a flash is presented in alignment with the moving, invisible, limb, then the observer sees the flash in a position lagging the registered position of the invisible limb (Nijhawan & Kirschfeld 2003; and see further on in the target article). One may attempt to explain this "motor flash-lag" result in terms of shorter latency for limb proprioception relative to the latency for the flash. However, the average motor flash-lag corresponds to $\Delta\sigma/v = 128.5$ msec. Because the input delay in the perception of the flash itself is likely to be in this time range, this explanation would have to assume an impossible, close-to-zero latency for limb proprioception.

Two suggestions follow from the motor flash-lag result. First, the finding of similar effects in perception and action reinforces previous suggestions that the sensory and the motor systems have an underlying unity (Arbib 1972; Rizzolatti et al. 1997; Sperry 1952). Second, there is little doubt that the two flash-lag effects, one measured in relation to a moving external object and the other in relation to the moving invisible limb under the observer's voluntary control, belong to the same category of phenomena. Hence, it is probable that the two effects have a common origin. Compensation for neural delays via feedforward (Georgopoulos 1986) or anticipatory motor control (Ghez & Krakauer 2000) that readily explains motor flash-lag, has the advantage of parsimony: the two effects, the standard flash-lag and the motor flash-lag, can be explained with one hypothesis.

We employed the flash-lag effect to test for the existence of an IMVP (Rojas-Anya et al. 2005). A previous study (Batista et al. 1999) supporting the existence of internal models found that visually guided reach plans in posterior parietal cortex are represented in an eye-centered frame; neural response was modulated by gaze direction irrespective of retinal position of visual reach targets in this study. Using the flash-lag effect, in our study we asked whether the instantaneous position of a moving item is represented in an eye-centered frame (Rojas-Anya et al. 2005). The moving item was either a visual target or the observer's invisible limb under voluntary control. During movement, a flash was presented in alignment with the moving item. The observers' gaze direction was manipulated such that at the time of the flash, the moving item (visual target or observer's limb) either approached the fovea or receded away from it. We found that in both cases, a much stronger flash-lag effect occurred when the item approached the fovea than when the item receded away from the fovea. The asymmetric flash-lag for visual motion is consistent with previous results with visual stimuli (Mateeff & Hohnsbein 1988). However, the same asymmetry for the voluntarily moved limb (Rojas-Anya et al. 2005), in the absence of any retinal image motion, shows two things. First, consistent with previous observations (Batista et al. 1999), the nervous system uses an eye-centered frame for representing location of moving limbs. Second, the similarity of gaze-contingent modulation of the flash-lag effect across the modalities suggests that the nervous system uses a common rule for localizing moving items. These results support the suggestion that the CNS uses an internal model to perform the important task of localizing both moving objects and moving limbs.

On a final note, it is likely that there is reciprocal influence between visual and motor processes – thus, motor processes not only inform but also receive feedback from visual processes, and vice versa. As Helmholtz argued (see Warren & Warren 1968, p. 112), the interactions between motor and visual representations may be trained by experiences such as the stimulation of the observer's own retinas caused by voluntary movements of his or her own limbs. The suggested interaction between voluntary movements and vision leading to common localization codes for visual objects and limbs would require cross-modal plasticity, for which there is ample evidence (Goel et al. 2006; Held & Freedman 1963).

## 8. Visual prediction: New, or another constancy?

Although the proposal of visual prediction has been deemed to be novel by some, this concept is in consonance with various constancy phenomena. Consider, for example, size constancy, which minimizes variation in the perceived size of objects despite variation in their retinal image size. Visual prediction may be thought of as a process that achieves "lag-constancy for moving objects." Just as the distance between the observer's eyes and objects is variable, so is the velocity of moving objects. Let an "early" and a "late" neural visual representation of a moving ball be denoted by $R_A$ and $R_B$, respectively, with the goal of the visual compensation being to reduce discrepancy between $R_A$ and $R_B$. Suppose further that $v_n$ and $v_N$ are two different velocities of the ball, with $v_N > v_n$. If there were no visual compensation, then the difference between $R_A$ and $R_B$ for $v_N$ will be greater than the difference between $R_A$ and $R_B$ for $v_n$. Furthermore, in the absence of any visual compensation, the difference between $R_A$ and $R_B$ would be expected to increase linearly with object velocity. Visual compensation for delays would achieve lag constancy for moving objects by minimizing the variation of the difference between $R_A$ and $R_B$ for different velocities. This suggestion is analogous to suggestions concerning motion deblurring (Burr 1980; Burr & Morgan 1997), in which the visual system removes motion-smear that would otherwise result from persistence of neural activity. Due to motion deblurring, a moving object appears relatively constant in shape, a process termed "sharpness constancy" (Ramachandran et al. 1974). In the absence of sharpness constancy, the same moving object, traveling at different speeds, would appear more or less elongated (smeared).



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

## 9. Implications for neural prediction

Until now, the approach developed here has focused on visual processes leading to perception or on visual processes feeding into action systems, which nonetheless do ultimately manifest themselves perceptually. However, the current approach can be easily adapted for visual spatial extrapolation processes that exclusively supply controllers of end-effectors with no concomitant influence on perception. The current approach can, furthermore, be applied to the investigation of prediction in the CNS more generally – an approach that may be termed *neural prediction*. Such an approach would complement other approaches (Duhamel et al. 1992; Ghez & Krakauer 2000; Kawato 1999; Wolpert & Flanagan 2001) investigating prediction phenomena in the CNS. The general implication drawn from the present approach is that the function of predictive mechanisms, located anywhere in the CNS, is to coordinate neural activity between different parts of the CNS.

### 9.1. Coordination of neural representations

It is suggested that a basic goal of the CNS is coordination of neural activity across various levels within the sensory and the motor systems. During normal behaviors under sensory guidance, the animal's actions depend on many neural representations existing at multiple levels in different modalities. Information must be exchanged between representations, therefore necessitating the translation of information (Andersen et al. 1993). For example, during visually guided action, information must be exchanged between visual brain areas representing position of objects in the environment and somatosensory/motor brain areas representing the position of the animal's effectors. Typically the distances that separate the representations are large, so the information exchange between the representations is time consuming. Neural prediction suggests that the goal of any type of compensation is to produce coordination between the time-varying neural representations that exist in parts of the CNS separated by significant delays.

Suppose two neural representations, $R_A$ and $R_B$, exist in two different parts of the CNS. As discussed in section 5.2.2, when there are inconsistencies, these representations might compete to control cell activity. For non-competing representations, there are three types of potential errors between $R_A$ and $R_B$, resulting from neural delays: (1) The moving coordinates of an item in one representation (say, $R_A$) may lead relative to the coordinates of the same item in another representation (say, $R_B$); for example, $R_A$ and $R_B$ could both be representing the visual position of a moving ball (say), with one being a retinal representation and the other a representation in LGN. In another situation, $R_A$ could be located in the visual and $R_B$ in the motor system, with $R_B$ representing the position of the ball in torso-centered coordinates. In yet another situation, $R_A$ could belong to one item (e.g., the ball) and $R_B$ to a different item (e.g., the subject's arm). Which representations are created, and which representations interact during behavior, will depend on the specific behavior. In general, the distances separating the representations – and consequently the neural delays – will vary. The task of coordination would be to minimize the differences between $R_A$ and $R_B$. (2) Different modalities process information at different speeds. Coordination would reduce the potential error between two representations (e.g., vision and touch), created by one object (e.g., an insect moving quickly on the observer's arm). (3) The neural latencies for different stimuli within a given modality vary depending on various properties (e.g., stimulus intensity). In the absence of a coordinating influence, for moving objects, the "rule" that connects the object location in retinal coordinates (say) to the object location in head-centered coordinates further downstream, would vary as a function of the contrast level of the stimulus. In this case, the coordinating influence would act to minimize the required modification of the rule in different situations.

## 10. Summary and conclusions

It is clear that predictive phenomena could not have evolved without input to the CNS from continuous time-varying stimuli such as those resulting from movement of objects, or if the world consisted only of discrete aperiodic events and unchanging objects. It is also clear that predictive responses seen in animals and neurons could not exist without compensation of the significant time delays that are intrinsic to all neural processes. Previous researchers have asked how the CNS reconstructs the continuous time-varying stimulus from discrete spike trains (Bialek et al. 1991). However, neurons located far from the receptor epithelia face the additional challenge of reconstructing the current state of the time-varying stimulus. This current state of the stimulus may be necessary data for decision processes, the outcome of which would directly determine whether the animal acts or withholds action, or which action the animal selects. In addition, the current state of changing stimuli may be necessary for providing the animal feedback about actions it has performed.

The reconstruction of the current state of the stimulus could be undertaken by a predictive process within the visual system; however, at the outset, prediction appears to be a high-level phenomenon. Hence, previous proposals of visual prediction led to wide debate. On the one hand, the debate is intertwined with the flash-lag findings based on unpredictable motion, and, on the other, with logical challenges for the notion of visual prediction. New analysis and data reveal that empirical results with unpredictable motion, and in particular the flash-terminated and flash-initiated conditions, are not inconsistent with visual prediction. In fact, the analysis shows that the times involved by the putative extrapolation mechanism are so small that the mechanisms may be more efficient than previously anticipated. In terms of logic, the fuel for the debate is the assumption that late (non-visual) processes compensate for all the delays during sensory-motor interaction. The corollary of this viewpoint is that the motor system receives spatially lagged visual input. Visual prediction suggests that the lag in the sensory data is compensated by visual spatial mechanisms in the feedforward vision pathways or by interactions of vision processes with late neural process linked to action.

Interception of moving targets is thought to depend on the motor system's ability to produce an output with features that match properties of the moving target. For example, when we attempt to catch a ball in flight, our motor system attempts to match the instantaneous



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

position of our moving hand with the position of the moving ball. However, an opposite position that is not often entertained is that the goal of the visual system is to deliver to the animal visual information that is suitably shaped for action directed at a moving ball. Thus, during a catch, the visual system modifies the perceived position of the ball so that it matches the position of the moving hand.[3] On this view, the goal of the visual system is to generate an output that has shared features with motor processes: in particular, predictions (Wilson & Knoblich 2005).

The proposal is that visual representations that receive feedback from higher cortical areas are susceptible to modification. Thus, these visual representations are controlled entities, just like neural limb representations. The descending visual signals cannot, of course, activate otherwise silent neurons, which is presumably only possible on the basis of stimulus-driven retinal input (Hupe et al. 1998). But the descending signals can, nonetheless, affect ongoing activity in many areas (e.g., the thalamus) and produce anticipatory spatial adjustments (Sillito et al. 1994).

Although my somewhat limited goal was to evaluate the feasibility of visual prediction, during the course of this endeavor it seems that prediction may be far more pervasive in the CNS than originally expected. The novel approach developed here may be easily adapted to investigate predictive phenomena in the CNS more generally. Visual prediction has a strong logical basis and seems consonant with other visual phenomena such as the various constancies and motion deblurring, as well as theoretical constructs such as neural competition. Prediction may be a multi-level, multi-modal phenomenon found in both sensory and motor systems. Furthermore, prediction may result from computations carried out by single neurons, or neural networks, or both. This general approach to the study of prediction suggests possibilities that could unify research from single cells to cognition.

ACKNOWLEDGMENTS

I would like to thank Beena Khurana for her continuous involvement with this work. I would also like to thank Raje Nijhawan, Mark Changizi, Wim van de Grind, Mike Beaton, Rich Ivry, Shin Shimojo, Kuno Kirschfeld, Gerrit Maus, and Raghbir Singh for stimulating discussions, and Barbara Finlay, Marcus Baldo and five referees for insightful feedback on an earlier draft of the article.

NOTES

**1.** Since signals within the retina itself are transmitted via graded potentials, as opposed to action potentials, the time taken for lateral interactions across 0.15 mm of retina could take significantly shorter time than the above estimate (Bruce Bridgeman 2005, personal communication).

**2.** There could be many reasons for behaviors normally displayed by animals to break down. Breakdown in behavior is known to occur, for example, in nervous systems in which disease has affected neural processing delays. For example, Multiple Sclerosis occurs because of demyelination, which affects neural transmission delays. Commonly observed behavior in healthy animals leads to the following assumption: In the absence of mechanisms compensating for neural delays, many behaviors in otherwise healthy animals would be disrupted. An analogous assumption holds for visual delays. The fundamental assumption of compensation for visual delays states that in the absence of mechanisms compensating for visual delays, many behaviors in otherwise healthy animals would be disrupted. Note that this last statement makes no assumption about whether visual or non-visual mechanisms compensate for visual delays.

**3.** One significant difference between the treatments of the visual position of a moving object and the sensed position of a moving limb is that we appear to have no conscious control over the position of the moving object, while we do have conscious control over the position of our limb. However, we are aware of only some internal representations that allow us to predict the future states of our limbs from current states during movement (Blakemore et al. 2002). Thus, many representations that allow for prediction of moving visual objects and of limbs during movement are not available to awareness.

# Open Peer Commentary

## Shifting attention to the flash-lag effect

doi: 10.1017/S0140525X08003816

Marcus Vinícius C. Baldo[a] and Stanley A. Klein[b]

[a] *"Roberto Vieira" Laboratory of Sensory Physiology, Department of Physiology and Biophysics, Institute of Biomedical Sciences, University of São Paulo, São Paulo, SP 05508-900, Brazil;* [b] *Visual Processing Laboratory, School of Optometry, University of California at Berkeley, Berkeley, CA 94720-2020.*

**baldo@usp.br** **sklein@berkeley.edu**

**http://www.fisio.icb.usp.br/~vinicius/index.html**

**Abstract:** An attention shift from a stationary to a changing object has to occur in feature – rather than physical – space, in order to bind these stimuli into a unitary percept. This time-consuming shift leads to the perception of a changing stimulus further ahead along its trajectory. This attentional framework is able to accommodate the flash-lag effect in its multiple empirical manifestations.

The flash-lag effect (FLE) is an empirical fact. Understanding the effect, however, requires us to build conceptual models that can range from the biophysical description of neuronal interactions to global cognitive schemes. Whichever model may be conceived, it should explain the FLE and related findings, rely on sound physiological grounds, adopt the smallest set of assumptions, and predict novel phenomena.

Motion extrapolation, as originally proposed by Romi Nijhawan (1994), nicely explains the flash-lag phenomenon in its purest form. However, in order to accommodate that explanation with discordant empirical findings, Nijhawan, in addition to splitting perceptual extrapolation in two successive steps, now resorts to at least three different putative mechanisms: (1) the role of lateral communications between neighboring retinotopic locations, (2) the impact of visual transients generated by offsets, and (3) a dynamical interplay between motor and visual processes. Therefore, Nijhawan's effort to fit motion extrapolation into challenging empirical data has compelled him to adopt a piecewise mosaic of physiological functions. Also, even though Nijhawan does a good job in weaving the explanatory content of his account, he leaves aside a deeper exploration of its predictive potential.

Both empirical and theoretical evidences point to the ability of basic sensory operations to generate the FLE (Baldo & Caticha 2005; Berry et al. 1999; Erlhagen 2003). In fact, a simple neural network model built upon a small set of physiologically grounded assumptions was able to replicate the standard FLE (under several conditions), as well as its dependence on stimuli luminance, priming, trajectory, and predictability (Baldo & Caticha 2005). This model also helps us evaluate the relative



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

underlying involvement of spatial sources (lateral interactions between neighboring receptive fields) and temporal sources (time delays required to activate the neurons in the network or to broadcast signals among them).

However, despite their physiological appeal, simplistic models are still unable to encompass the phenomenological richness portrayed by this illusion, such as has been observed in the chromatic (Nijhawan 1997; Sheth et al. 2000), spatial frequency (Sheth et al. 2000), and auditory (Alais & Burr 2003) domains. This wide-ranging manifestation of the FLE would call for the contribution of higher, more integrative, underlying mechanisms. Over the last few years, we have been refining a conceptual framework in which attention plays a role in contributing to the FLE. As pointed out by Nijhawan himself, "distances that separate the representations are large, so the information exchange between the representations is time consuming" (target article, sect. 9.1, para. 1). Actually, the first conceptual alternative challenging the motion extrapolation account was based on the time consumed by a unidirectional shift of (spatial) attention from the flashed location to the moving location (Baldo & Klein 1995). Meanwhile, whereas the role of attention in causing the FLE remained elusive, the modulation of the FLE by attentional factors was gradually being established (Baldo & Namba 2002; Baldo et al. 2002; Chappell et al. 2006; Namba & Baldo 2004; Sarich et al. 2007).

More recently, we have advanced our attentional framework by proposing that the FLE could originate from the time needed for attention to bind the flash and moving percept into a unitary whole (Baldo & Klein, in press). Starting with the detection of an abrupt event (a stationary flash or a beep, for instance), a shift of attention from a stationary object to a changing ("moving") object has to occur in feature rather than physical space in order to bind them together into a unitary percept. This object-based attentional shift would require some time to be carried out, regardless of any spatial separation between both visual stimuli, and would naturally lead to a percept consisting of a changing stimulus further ahead along its "trajectory" (equivalently, we can also consider a non-directional attentional spread over the object, in feature space, which starts when an abrupt-onset stimulus is presented – though not necessarily from where it is presented).

It is easy to see that this feature-based "attentional" explanation is not in conflict with the findings concerning the flash-initiated and flash-terminated conditions, as claimed before in relation to the purely spatial attention shift (Alais & Burr 2003; Khurana & Nijhawan 1995): Whereas the time required for the attentional binding will lead to the FLE in the former condition, no FLE will be observed in the latter (in opposition to motion extrapolation's predictions), since the moving stimulus never reaches a position beyond that where it disappeared or stopped.

The relationship between delays required to bind a spatially extended object and time-consuming shifts (or spread) of attention has been extensively reported (Chappell et al. 2006; Enns & Oriet 2004; Houtkamp et al. 2003; Kanai et al. 2004; Roelfsema et al. 2000; Sarich et al. 2007). Because the scheme we propose involves an attentional binding in feature space, our account unifies empirical findings observed in a broad class of visual features (such as position, color, luminance, and spatial frequency), as well as in crossmodal phenomena (Alais & Burr 2003). It is worth noting that only cognitive-oriented models are presently able to capture the entire perceptual spectrum revealed by the FLE. Nijhawan's motion extrapolation and the attentional binding account we presently offer are the main conceptual frameworks embodying this class of integrative models. Our proposal, however, is favored by a more parsimonious set of assumptions, a deeper physiological root, and a greater predictive power.

Eventually, cognitive models are to be brought down to the wiring structure and functional operations of neuronal lattices. It is conceivable that lateral interactions giving rise to spatio-temporal facilitations, as evidenced by simple networks (Baldo & Caticha 2005) and also assumed by Nijhawan, could be further elaborated and extended to perceptual processing other than visual. A desirable connection between the fundamental rules of neural functioning and the large-scale expression of cognitive integration could thus be envisaged.

In conclusion, it is encouraging to recognize that the decade-long dispute over the underpinnings of the FLE has already contributed toward refining, discarding, or amalgamating fragmented and contradictory conceptions. Whatever the final consensus turns out to be, Nijhawan's legacy to the current debate on visual perception will endure.

ACKNOWLEDGMENTS

This work has been partially supported by Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico.

## Anticipation requires adaptation

doi: 10.1017/S0140525X08003828

Christian Balkenius and Peter Gärdenfors

*Department of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden.*

**christian.balkenius@lucs.lu.se** **http://www.lucs.lu.se**
**peter.gardenfors@lucs.lu.se** **http://www.lucs.lu.se**

**Abstract:** To successfully interact with a dynamic world, our actions must be guided by a continuously changing anticipated future. Such anticipations must be tuned to the processing delays in the nervous system as well as to the slowness of the body, something that requires constant adaptation of the predictive mechanisms, which in turn require that sensory information be processed at different time-scales.

The target article presents an interesting analysis of visual prediction, and we fully agree with Nijhawan that predictions need to be made at different levels of the sensory-motor processing. However, Nijhawan's article severely underestimates the complexity of anticipation in the sensory-motor system. In particular, it ignores the important question of how an anticipatory system can tune its predictions to internal and external delays as a result of experience. Consider the example of catching a ball. It is not sufficient that "the visual system modifies the perceived position of the ball so that it matches the position of the moving hand," as Nijhawan suggests (sect. 10, para. 3). Instead, such a task involves at least the following components that can be divided into a visual pursuit and a catch component. Even to just visually focus on the ball, its trajectory needs to be anticipated (Balkenius & Johansson 2007). Since visual processing and the eye movements following it are not instantaneous, it is necessary to predict where the ball is right now to be able to fixate it while it is moving. We call this the "anticipated now" because any sensory code that is synchronous with an external state must be based on anticipation.

The predictions resulting in the anticipated now may or may not be correct, but there is no way for the organism to correct these predictions until at a later time, when the true sensory input becomes available. At this time, it is possible to adapt the earlier predictions to the actual sensory input, something that requires that the earlier anticipation, as well as the sensory information used for it, should still be available. This implies that at every moment, the sensory input is used both to anticipate the future and to adapt earlier predictions, but because of processing delays, it cannot be used to code for the current state of the external world. Similarly, eye movements cannot be based on the anticipated now, but must be controlled by the anticipated future. Looking at a moving object therefore requires that the



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

organism simultaneously maintain sensory information at five different time-scales: the current sensory input, the anticipated now and future, and previous predictions of the now and the future. By combining information at the different time-scales in an appropriate way, it is possible both to change the currently anticipated now and to make future predictions more accurate. The most appropriate way to model the combination of time-scales is presumably to use feed-forward models from control theory (Grush 2004; Kawato 1999). The predictions can be generated with different Kalman filters. The Kalman gains of the scales can be weighted in different ways, depending on the experience of the organism.

Assuming the gaze system is correctly tuned, via some feed-forward mechanism, the temporal unfolding of the ongoing interaction with the visual target contains the information needed to predict the location of the ball in the future; but the task for the hand is not to move to any arbitrary point along the predicted trajectory of the ball. Instead, the sensory-motor system must direct the hand to the location where the ball will be once the motor command to reach that location has been executed. This introduces an additional type of complexity, since the time in the future when the hand will catch the ball depends on properties of the arm and hand as well as on the ball. Although this is strictly also true for eye movements, the physical lag of the system becomes more critical for arm movements.

The properties of the flash-lag effect become perfectly sensible within this framework. Because unexpected events cannot become part of the anticipated now until after a processing delay, they will be perceived as lagging any predictable event. Moreover, since point events are perfect predictors of themselves, an adaptive system will learn to let them replace any prediction based on prior information.

In the flash-terminated condition, the flash becomes part of the anticipated now at the same time as the detection of the disappearance of the target. The flash and the disappearance of the target should therefore be perceived as simultaneous and occurring at the physical location where these two events actually take place. Once this information is received, the best prediction is that the moving object disappeared where it was at the time of the flash. This does not mean that the movement of the object is not extrapolated; it only suggests that that extrapolation is replaced by better information when it is available. Although it is possible that biased competition plays a role in this process, it may be sufficient to assume that a system that adapts its predictions to actual information will learn to behave in this way.

It is clear that the number of different time-scales that are necessary in the brain is much larger than the few described here. Given that substantial delays influence all processing in the brain, it appears necessary to compensate for these by predictive mechanisms at all levels. This suggests that predictive abilities should be necessary at the level of individual neurons – or at least local circuits. The ongoing dynamical interaction among different parts of the brain is not too different from the brain's interaction with the external world.

ACKNOWLEDGMENT

This work was funded in part by the European Union project MindRaces, FP6-511931.

## Visuomotor delay in interceptive actions

doi: 10.1017/S0140525X0800383X

Nicolas Benguigui,[a] Robin Baurès,[a] and Cyrille Le Runigo[b]

[a]*Laboratory Motor Control and Perception (EA 4042), University Paris-Sud, 91405 Orsay Cedex, France;* [b]*Laboratory Efficience et Deficience Motrices, University Montpellier 1, 34090 Montpellier, France.*

**nicolas.benguigui@u-psud.fr**
**http://www.cress.u-psud.fr/monde/benguigui.htm**
**Robin.baures@u-psud.fr**
**http://www.cress.u-psud.fr/monde/baures.htm**
**cyrille_lerunigo@yahoo.fr**
**http://www.edm.univ-montp1.fr/affich_cv.php?id=95**

**Abstract:** Neural delays, which are generally defined as visuomotor delays in interceptive actions, must be compensated to enable accurate timing in movement. Visuomotor delays can depend on the kind of task, the use of information, and the skill of the performer. The compensation for such delays does not necessarily require prediction or representation but can be made by an attunement of some parameters in what is called a *law of control*.

In this target article, Nijhawan proposes that neural delays in perception must be compensated by visual predictions and internal models. He emphasizes that this compensation is particularly important in interceptive actions and that this issue remains to be addressed. We agree, but would like to point out that there are empirical data and an alternative hypothesis which do not depend on the use of internal models.

In interceptive actions, researchers generally refer to a *visuomotor delay* (VMD) to define the time period between the pick-up of information and its use in producing an adjustment in movement (e.g., Tresilian 1993). It has been shown that VMD duration depends on the task (Benguigui et al. 2003). Lee et al. (1983) calculated VMDs that ranged from 50 to 135 msec in a ball striking task. Whiting et al. (1970) showed, in a ball-catching experiment, that performance did not degrade when the occlusion of the final part of the trajectory was equal or inferior to 100 msec. This period of time was interpreted as a VMD during which no information was used for catching. Bootsma and van Wieringen (1990), in a table-tennis task, observed that the variability was minimal at about 100 msec before contact. This phase of minimal variability was described as the end of the control of the action, reflecting the duration of VMD.

Some researchers have shown that VMD could be longer than 100 msec. McLeod (1987) reported that expert cricket batsmen, confronted with unexpected changes in the ball trajectory at the bounce, needed at least 190 msec to adapt their swing to the new trajectory. Benguigui et al. (2003) demonstrated in a task consisting of intercepting a moving stimulus with a thrown projectile, that movements were initiated around 200 msec after a specific source of information had reached a critical value. This delay was interpreted as a VMD between the pick-up of the information and the initiation of movement.

The duration of VMD appears to depend on the use of information. It can be as short as 100 msec when information is used in a continuous mode, but can reach values near 200 msec when information is used in a discrete mode, such as the beginning of the movement or some important correction of that movement.

One can assume that (1) a short VMD and (2) an accurate compensation of it are essential for accurate timing. First, because information is continuously changing during the approach of a moving object, a reduction in VMD allows later pick-up of increasingly accurate information as contact draws near. The later the information pick-up, the more accurate the interceptive action will be (Lobjois et al. 2006). Le Runigo et al. (2005) showed that the time required to produce an adaptation in an interceptive movement after an unexpected change in the moving object's velocity was shorter in expert tennis players (162 msec) than in novices (221 msec). They also showed that the reduction of VMD was highly correlated to timing accuracy, suggesting that a reduced VMD provides the opportunity to improve on-line regulations and to adapt these regulations at later stages before contact.

Second, the movement itself must compensate for VMD under the constraint of an incompressible delay. Moreover, the difficulty in compensating for VMD would increase as VMD increases, as is the case in older adults (Bennett & Castiello



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

1995). Lobjois et al. (2006) showed that this increase in VMD explained the lateness of elderly people in a coincidence-timing task. Interestingly, elderly people who had a regular sport activity (e.g., tennis) and who had an increase in VMD that was similar to that of sedentary elderly people were not late in their responses. These results suggest that sport practice also allows a better compensation for the age-related VMD increase.

Although the problem of compensation for VMD has not yet been extensively addressed, some suggestions deserve consideration. When an interceptive movement is very short and controlled in a ballistic mode (movement time: MT $\leq$ 150 msec; Tresilian et al. 2004), the actors have to estimate temporally the ensemble of the visuomotor sequence, including VMD and the MT, and detect the instant at which time-to-collision (TTC) becomes equal to the duration of this sequence (Benguigui et al. 2003).

Analogous mechanisms could be also involved when action must be controlled with continuous regulation. In contrast to Nijhawan's claim (sect. 7.2.1.1), the visual system is not an accurate predictor of moving objects when the last part of the trajectory is occluded beyond a duration of 200 msec (or of a VMD). Numerous studies using prediction motion tasks consisting of estimating the TTC of a moving object after the occlusion of the final part of the trajectory have shown that accuracy and variability in timing responses are dramatically affected by occlusions longer than 200 msec (e.g., Benguigui et al., in press). This means that for MTs above 200 msec, on-line regulation of movement is indispensable for dealing with both unpredictable and predictable[1] changes in the object trajectory.

Once again, however, on-line regulations require compensation for VMD. This compensation process can be understood according to the calibration principle laid out by ecological psychologists (Jacobs & Michaels 2006). As motor control is based on laws corresponding to functional relationships between information and movement (Warren 1988), the calibration process consists in attuning some parameters of the law of control in order to adapt the movement to the constraints of a specific task. Regarding VMD compensation, laws of control include some parameters that do not correspond to a representation of VMD as is suggested by Nijhawan. Instead, these parameters have to be set according to the task-dependent effect of VMD in the interceptive timing. This process could explain why athletes who master a specific set of skills (exploiting a specific law of control) generally need a few trials of preparation to optimize their efficiency when beginning a new session of practice. This process could also correspond to an issue of learning that remains to be explored.

ACKNOWLEDGMENTS

We thank Michael P. Broderick and Isabelle A. Siegler for feedback on earlier draft of this commentary.

NOTE

**1.** For instance, even if the acceleration of a moving object is detectable by the visual system and theoretically predictable, results have shown that the perceptuo-motor system is unable to make predictions in extrapolating time-to-collision (TTC) that take into account the variation of velocity (Benguigui et al. 2003).

## Flash-lag: Prediction or emergent property of directional selectivity mechanisms?

doi: 10.1017/S0140525X08003841

Julia Berzhanskaya

*Allen Institute for Brain Science, Seattle, WA 98103.*

**juliab@alleninstitute.org**

**Abstract:** 3D FORMOTION, a unified cortical model of motion integration and segmentation, explains how brain mechanisms of form and motion processing interact to generate coherent percepts of object motion from spatially distributed and ambiguous visual information. The same cortical circuits reproduce motion-induced distortion of position maps, including both flash-lag and flash-drag effects.

Perceived position of objects is often distorted in the presence of motion. In addition to the flash-lag effect discussed in the target article, other examples include flash-drag effects (Whitney & Cavanagh 2000) and illusory boundary distortion by motion (Anderson & Barth 1999; De Valois & De Valois 1991; Ramachandran & Anstis 1990; Weiss & Adelson 2000). The same neural circuits that carry out directional selectivity, motion integration and segmentation computations are suggested to underlie phenomena of distortion of positional maps (Berzhanskaya et al. 2004).

The model 3D FORMOTION (Berzhanskaya et al. 2007) is based on formotion (Grossberg et al. 2001) and employs feedforward and feedback circuits involving areas V1, V2, medial temporal (MT), and medial superior temporal (MST) to solve both the motion aperture and correspondence problems (as well as motion capture, barber-pole illusion, plaid motion, and integration of object motion across apertures). It incorporates depth-selective input from V2 (FACADE) to MT (Formotion) to simulate the separation of ambiguous boundaries in depth (chopsticks illusion) and motion transparency.

***1. Description of 3D FORMOTION features relevant to the distortion of positional maps.*** The formotion model consists of a few stages (see Fig. 1). The first stage in the motion system consists of directionally insensitive transient cells that respond briefly to a change in the image luminance, irrespective of the direction of movement. The second stage is the directionally sensitive layer. Directional selectivity results from a combination of gradient processing (Reichardt 1961; van Santen & Sperling 1985) and an asymmetric inhibition mechanism (Jagadeesh et al. 1997; Livingstone 1998). Further, short-range filters accumulate motion signals in a certain direction. Finally, spatial competition weakens ambiguous signals from line interiors and also amplifies feature-tracking signals. The spatial competition kernel has an asymmetric shape with its excitation offset from inhibition in the direction of motion. On the basis of anatomical connections and directional opponency, these spatial/directional competition units are allocated to the layer 4B of V1. As will be demonstrated later, asymmetric inhibition at this motion processing stage may contribute to the flash-lag effect.

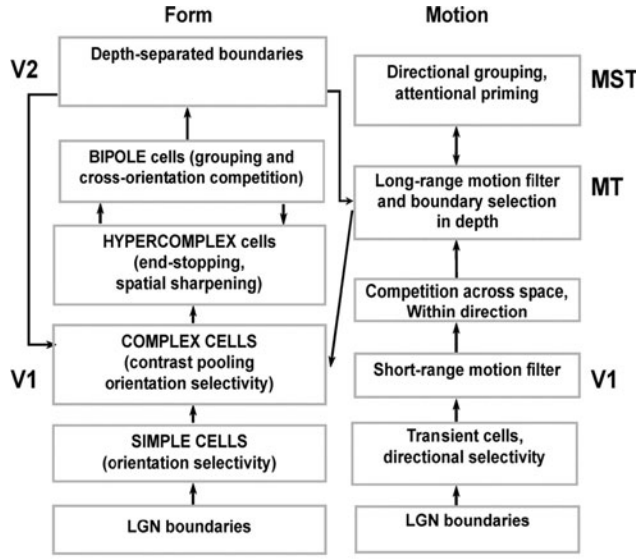



Figure 1 (Berzhanskaya). Schematic view of 3D FORMOTION.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

The signal from the layer 4B of V1 is projected to MT. Long-range anisotropic receptive fields in MT accumulate evidence of motion in the direction of their elongation. The final stage of the model is a "winner-take-all" interaction between MT center-surround and wide-field motion cells (in MT or MST). The MT-MST network determines a single motion direction at each position and may incorporate attentional modulation. Unambiguous feature-tracking signals propagate to ambiguous motion regions via MST-to-MT feedback, and anisotropic receptive fields allow motion integration across occluders. Thus, only at the MT-MST level does the model solve both the aperture and motion capture problems.

The FACADE boundary computation system includes stages of contrast-dependent orientation-selective cells (simple cells in V1), contrast-independent orientation-selective cells (complex cells in V1 and V2), hypercomplex cells that support both orientation competition and spatial competition, and finally, bipole grouping cells (Grossberg & Mingolla 1985). In the current model, the hypercomplex cell layer is the site of projection from the motion stream to the form stream, carried out by MT-V1 feedback. This projection plays a role in assigning depth order to the object boundaries in ambiguous displays. The same projection, at the MT-V1 (simulated) level and possibly also at the MT-V2 level, is suggested to influence object boundaries in the flash-drag effect.

***2. Motion distortion of boundary positions: modeling insights into neuronal mechanisms.*** In the flash-lag effect (Nijhawan 1994), a flashed object is perceived to be lagging behind a moving one. 3D FORMOTION suggests that when a spatial/directional competition filter in V1 is activated by a moving stimulus, the trailing surround suppression shifts the peak of population activation of a flashed stimulus in the direction opposite to the direction of motion (Berzhanskaya et al. 2004); see Figure 2. A similar phenomenon may exist at levels higher than V1. However, based on the size of the shift (a few degrees of visual angle) and on the limited spatial scope, one can suggest that the main mechanisms are localized in V1.

In an opposite effect, flash-drag (Whitney & Cavanagh 2000), a flashed stimulus is shifted in the direction of motion. It can be explained by the feedback from MT to V1 (implemented in 3D FORMOTION) or MT to V2. In these experiments, a motion stimulus occupies a larger area (and presumably activates larger MT/MST fields more efficiently), and the effect, a positive peak shift in the direction of motion, can be observed at larger spatial distances (see Fig. 3).

Another category of visual displays, gratings with ambiguous (Gaussian envelope) boundaries (DeValois & DeValois 1991; Whitney et al. 2003) and illusory contours formed by random dots (Ramachandran & Anstis 1990), demonstrates that an ambiguity (low contrast, brief, or illusory) in the boundary presentation is a key to a shifted percept. Although not specifically analyzed here, differences in latencies resulting from weaker and slower-developing boundary signals may play a role at the later stages (a possibility discussed in the target article). Feedback projection from large, motion-sensitive MT receptive fields in 3D FORMOTION to lower areas such as V1 might be interpreted as carrying a predictive signal (also addressed in the target article). However, in our application of 3D FORMOTION to flash-lag and flash-drag phenomena, first presented at the Visual Science Society Conference in 2004, we were mainly interested in gaining insight into the neuronal mechanisms of these effects.

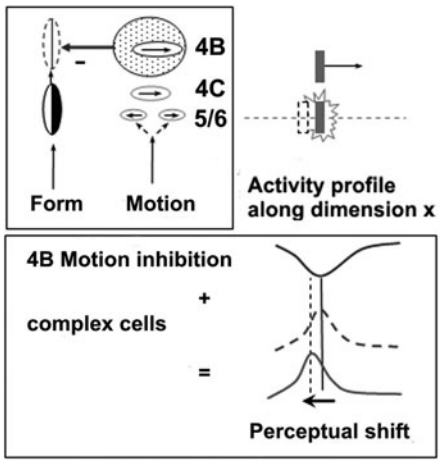



Figure 2 (Berzhanskaya). Illustration of motion inhibition effect on flashed stimulus position.

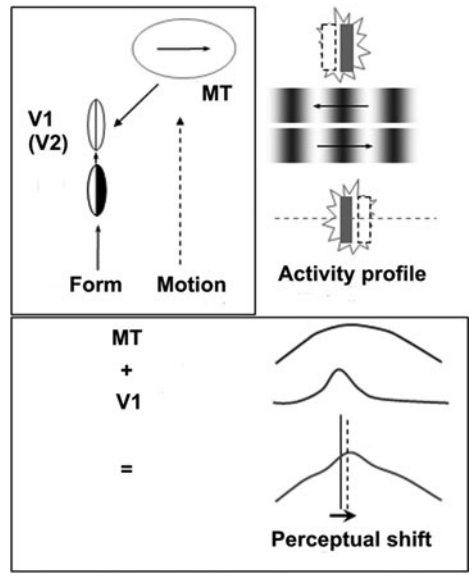



Figure 3 (Berzhanskaya). Illustration of MT contribution to flash-drag.

Our model makes multiple predictions: (1) It predicts that the flash-lag effect is a consequence of directional selectivity mechanisms. This effect would exist not only in V1 or later visual areas, but also in any area, sensory or otherwise, in which directionality of the stimulus can be processed. This was confirmed by flash-lag demonstrations in multiple domains (Alais & Burr 2003; Cai & Schlag et al. 2001; Sheth et al. 2000). (2) Flash-lag is a function of local motion processing and is only weakly affected by a global motion signal. An elegant demonstration by Anstis (2007), in which the flash-lag is influenced by a local motion and not by (an opposite) global motion of the chopsticks, provides some support to this idea. (3) A disruption of inhibition, and therefore of directionality mechanisms, near the flashed object should cancel the flash-lag effect. (4) A disruption of feedback from MT to V1 [e.g., by transcranial magnetic stimulation (TMS)] will interfere with the flash-drag effect (McGraw et al. 2004). In agreement with the model's predictions, neurophysiological evidence of interactions between form and motion streams that may contribute to object mislocalization has been presented in both single-neuron recordings (Fu et al. 2004; Sundberg et al. 2006) and fMRI (functional magnetic resonance imaging) (Whitney et al. 2003).

In summary, motion distortion of position maps can arise as an emergent property of directionality mechanisms, and, for the visual system, via interactions between the form and motion streams. An integrative approach, including analysis of multiple position distortion phenomena, will provide an insight into possible neuronal mechanisms.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

ACKNOWLEDGMENTS
I thank Professors S. Grossberg and E. Mingolla for their support and guidance during development of 3D-FORMOTION.

## The trade-off between speed and complexity

doi: 10.1017/S0140525X08003853

Mark Andrew Changizi
*Cognitive Science, Rensselaer Polytechnic Institute, Troy, NY 12180.*
**changizi@rpi.edu**
**http://www.changizi.com**

**Abstract:** The hypothesis that there has been selection pressure for mechanisms which enable us to perceive the present tends to be conflated with the hypothesis that there has been selection pressure for mechanisms that compensate for inevitable neural delay. The relationship between the two is more subtle, because increases in neural delay can be advantageous for building more useful perceptions.

Proponents of the hypothesis that the brain has mechanisms for perceiving the present (i.e., mechanisms designed to generate a perception at time $t$ that is representative of the scene at time $t$) typically say that the advantage is that it helps overcome inevitable neural delays. That is, "latency compensation" and "perceiving the present" have gone hand in hand. I, too, have made this equation in my own articles on how these ideas may be integral to a unified account of illusions (Changizi 2001; Changizi & Widders 2002; Changizi et al., in press; although see Changizi 2003, pp. 75–77). The implicit assumption can often seem to be that natural selection has attempted to minimize neural delays – by shortening wires, speeding up signal propagation, and using rapid computational algorithms for generating a visual percept – and whatever latency between retina and perception is left is handed over to the compensation mechanisms to deal with. Although this is an open possibility, the hypothesis that we perceive the present is not committed to this possibility; it is only committed to the idea that perceptions belong to the present. What is left open is how long the delay is, and whether it is all "inevitable" or whether the delay may be much longer than it would be if selection for short processing times trumped all other selection pressures.

Consider computer software as an analogy. Computer processing speed has risen by many orders of magnitude over the course of the last 20 years, but you may have noticed that many of your programs still take considerable time to start up. Computer designers know how long a wait we are willing to endure, and use that time to carry out fancier computations. That is, when faster computers arrive, computer designers do not appear to be saying, "Now we can compute the same old things nearly instantaneously!" Instead, they seem to be saying, "Now think about how much more we can compute while the user waits!"

Just as computer software delay is a consequence of a trade-off between shorter delay and more complex computations, our perceptual delay is a trade-off between shorter delay and fancier visual computations. For example, if evolution can find a new clever trick for extrapolating farther out into the future – say from 30 msec to 120 msec – then it could utilize this trick and allow itself four times the amount of computation time to build sophisticated useful perceptions. The resultant latency of 120 msec would not be understood as an inevitable delay left over after trying to reduce it as much as possible. Instead, it would be better to say that there is selection pressure to maximize the delay for which the nervous system is able to compensate, thereby buying more time to prepare the perception. Counterintuitively, then, it may well be that the slower-to-react brains are the "smarter" ones.

## Visual prediction as indicated by perceptual adaptation to temporal delays and discrete stimulation

doi: 10.1017/S0140525X08003865

Douglas W. Cunningham
*Wilhelm-Schickard-Institut für Informatik, Graphisch-Interaktive Systeme, University of Tübingen, 72076 Tübingen, Germany.*
**douglas.cunningham@gris.uni-tuebingen.de**
**http://www.gris.uni-tuebingen.de/lehrstuhl/personen/mitarbeiter/36/cunningham.html**

**Abstract:** Analogous to prism adaptation, sensorimotor compensation for existing neural delays has been clearly demonstrated. This system can also adapt to new delays, both internal and external. This seems to occur at least partially in the sensor systems, and works for discrete, stationary events. This provides additional evidence for visual prediction, but not in a manner that is consistent with spatial extrapolation.

Nijhawan makes a compelling case that the compensation for neural delays must occur at some level, and possibly many levels, of the human central nervous system. In the several sections of this argument, he mentions research from prism adaptation. Prism adaptation is clear evidence that the human central nervous system compensates for its own internal spatial characteristics. If the sensorimotor system can compensate for internal spatial characteristics, might it also be able to compensate for internal temporal characteristics (i.e., neural delays)? As Nijhawan points out, it must, since failure to do so would yield an organism that cannot interact well with the world.

The field of prism adaptation starts with the observation that intersensory spatial arrangements are not constant over the course an organism's life, but change both slowly (e.g., during maturation) and rapidly (e.g., when reaching through water). Any organism that cannot adapt its compensatory mechanisms to deal with changes in intersensory relationships will have nearly as much difficulty interacting with the world as an organism that has no compensatory mechanisms. Prism adaptation shows that human sensorimotor systems can adjust to both rapid and slow changes in intersensory relationships (for reviews, see Bedford 1993; Welch 1978).

Neural delays also change both rapidly and slowly during each organism's lifetime (see, e.g., Ahissar & Ahissar 1994). Thus, if humans can compensate for the temporal characteristics of the sensory systems like they do for the spatial characteristics, might not this mechanism also be able flexible enough to respond to changes in neural delay? Since the 1950s, several researchers have shown that this does not seem to be the case (see, e.g., Sheridan & Ferrel 1963; Smith et al. 1962; 1963). It has, however, since been conclusively demonstrated that not only do humans compensate for external temporal shifts, but the mechanism for doing so is remarkably similar to that involved in prism adaptation (Cunningham et al. 2001a; 2001b). In general, the introduction of an external visual delay initially impairs performance, but a small amount of practice returns behavior to nearly normal levels, and the adapted state generalizes to similar situations. Subsequent removal of the delay produces a large renewed drop in performance as the newly adopted state of temporal compensation ceases to be appropriate (this negative aftereffect is the hallmark of the semipermanent nature of sensorimotor adaptation). Subsequent work has confirmed this effect (Fajen 2007; Fujisaki et al. 2004; Miall & Jackson 2006; Navarra et al. 2007; Stetson et al. 2006; Vatakis et al. 2004; 2007). It is critical here to note that, unlike the visual prediction effect speculated to exist for moving objects, the temporal adaptation effect cannot be completely explained by spatial extrapolation for many reasons, including the fact that it occurs for discrete, stationary stimuli.

Nijhawan focuses explicitly on compensating for continuous events, with special emphasis on moving objects. There must, however, be some form of temporal compensation that also



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

works for discrete events. The reason for this is simple: neural latencies differ drastically for different sensory modalities as well as within a modality for different stimuli (e.g., Bolz et al. 1982; Sestokas & Lehmkuhle 1986). The absence of a coordinating influence that can also compensate for brief, non-moving events would necessarily lead to a fragmentation of the phenomenally experienced world. For example, the brighter portions of a briefly seen object would seem to have occurred at different time than the darker portions. As multimodal, briefly presented, nonmoving objects are not perceptually fragmented, and because many of the demonstrations of temporal adaptation have successfully used stationary, briefly presented, discrete stimuli, there must be some form of temporal compensation that does not involve spatial prediction.

Finally, in their experiments, Cunningham et al. (2001a; 2001b) noted that many participants complained about apparent perceptual changes during temporal adaptation. More specifically, as behavior in the delayed conditions approached normal levels, the participants began to indicate that the visual and proprioceptive events, which were separated by up to 400 msec, seemed to occur simultaneously. A number of studies have since empirically confirmed this perceptual shift in the point of subjective simultaneity (Fujisaki et al. 2004; Navarra et al. 2007; Vatakis et al. 2007; Vroomen et al. 2004). Several of these studies used brief, discrete, stationary stimuli. In general, these results suggest that adaptation to temporal delays is at least partially perceptual in nature. Nijhawan argues that prism adaptation, to the degree that it occurs in the visual system, is support for the theory that the visual system is engaged in prediction. Thus, the evidence that humans can both behaviorally and perceptually compensate for changes in delayed information would seem to be support for the claim of visual prediction in general, if not for spatial extrapolation in particular.

In sum, the human sensorimotor systems can compensate not only for their own internal characteristics, but also for external changes in both the spatial and temporal relationships between the senses. Moreover, at least for temporal adaptation, this compensation occurs partially in the perceptual systems, implying that they are engaging in some form of prediction. As temporal adaptation also works for discrete as well as continuous events and is not necessarily motion based, it cannot be fully explained by spatial extrapolation. Nijhawan states that, "The goal of visual prediction is to use priors contained in the unfolding visual stimulus to create a perceived state of the world that matches, as far as possible, the actual state of the world" (sect. 4, para. 2). Given what is known about how spatial and temporal compensation mechanisms adapt, this statement might be amended to make the strong involvement of previous experience more explicit: The goal of visual prediction is to use priors acquired from both previous experience and the currently unfolding visual stimulus to create a perceived state of the world that matches, as far as possible, the actual state of the world.

## Asynchronous neural integration: Compensation or computational tolerance and skill acquisition?

doi: 10.1017/S0140525X08003877

James E. Cutting

*Department of Psychology, Cornell University, Ithaca, NY 14853-7601.*
**jec7@cornell.edu**
**http://people.psych.cornell.edu/~jec7/index.htm**

**Abstract:** Nijhawan argues that neural compensation is necessary to account for couplings of perception and action. Although perhaps true in some cases, computational tolerance for asynchronously arriving continuous information is of more importance. Moreover, some of the everyday venues Nijhawan uses to argue for the relevance of prediction and compensation can be better ascribed to skill.

> *It is the mark of an educated man to look for precision in each class of things just so far as the as the nature of the subject admits . . . and not to look for precision in all things alike.*
>
> — Aristotle, *Nicomachean Ethics*, I, 1094b24 & 1098b5

Nijhawan reviews the need for theoretical consideration of delays brought about by differential neural transmission time. Such a need is unequivocal in some situations, and the flash-lag effect, which Nijhawan has championed for some time, seems reasonably accounted for in such terms.

However, not all neural delays need be compensated for. Instead, there is often wide tolerance for integrating diverse continuous signals. As an example, consider first the asynchronous presentation of auditory speech and its visual signal. What are the tolerances in the perceiver for detecting that anything is amiss? As it turns out, the window is fairly large and asymmetrical. The auditory delays referenced to the visual signal can be between −45 msec (an auditory lead) and +200 msec (an auditory lag; Grant et al. 2004). Rewritten, this is $\sim 80 \pm 120$ msec. The 80 msec value can be taken as a difference in the delay between the auditory and visual systems. It might need to be "compensated" for, and it is consistent with Nijhawan's report. However, the tolerance window of $\pm 120$ msec from that mean value suggests that no compensation is necessary. Instead, information in both signals is temporally distributed and, arriving differentially from auditory and visual modalities, is seamlessly integrated within a fairly wide temporal window.

This first example, however, is not about perception and action, with which Nijhawan is most concerned. Therefore, consider a second case – the disfluency effects of delayed auditory feedback while speaking. Stuart et al. (2002) reported no performance differences among 0, 25, 50 msec delays, but there were many disfluencies at 200 msec delays. In other words, the coupling of auditory perception and vocal action is no more affected by a 50 msec delay than by listening to oneself normally. This too suggests tolerance, not compensation.

Of course, this second line of evidence concerns audition, which is not the modality of Nijhawan's focus. Therefore, let us consider, third, the results from the literature on delayed visual feedback. In a functional magnetic resonance imaging (fMRI) study, Farrer et al (2007) showed that subjects had no differential awareness of a 0 versus 100 msec visual delay in watching themselves pluck pegs from a board, and their performance did not suffer either. Moreover, the angular gyrus, which seems to register awareness of such asynchrony, was not substantially engaged until delays of 150 msec and longer. Again, this suggests tolerance for décalage in signals, this time visual and motor.

Nijhawan's primary interest concerns the prediction of the visual location of a moving object in space. The most critical everyday venues for such needs are in hunting and in ball games. Among the latter – whether baseball, cricket, tennis, or ping pong – one can rightly argue for a required accuracy of $\sim 5$ msec when hitting a ball with appropriate control (Bootsma & van Wieringen 1990; Land & McLeod 2000; Lobjois et al. 2006; Tresilian 1993; Watts & Bahill 1990). Nijhawan's point is that the visual system – retina to V1 and beyond – is slow by comparison.

Nonetheless, in all of these sports, the best participants are involved in highly stereotyped, highly practiced situations. Moreover, the relevant behavior of the opponent and the flight of the ball are distributed over as much as a second and often more. In baseball, those deemed excellent hitters – that is, they can accurately hit the ball 3 out of 10 times against the best pitchers – earn well more than US$1000 per swing of their bat.

Skill accrued over many years – and a background of appropriate genetics – is required to perform at this level. Moreover, there is ample evidence that hitters cannot follow the ball near



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

the critical time when they need to hit it (Land & McLeod 2000; Watts & Bahill 1990), and visual acuity in the periphery, where the ball arrives when it is hit, is extremely poor. Hitting skills are not due to higher cognition. Indeed, thinking often interferes with performance (e.g., Beilock et al. 2004). But such performance is also not due to the kind of neural compensation that Nijhawan addresses. Instead, it is the result of a skill-driven refinement of perception and action in the context of the information distributed in time.

A kind of compensation does occur in a number of situations. In archery, and in trap and skeet shooting, the sportsperson must aim to lead the target that is being shot at. Yet it seems unlikely that one should look for pure sensory or motor accounts in such feats. Instead, and again, it is the skill of the individual calibrated over repeated practice that allows targets to be hit.

None of this is to denigrate Nijhawan's attempt at laying out the needs for a coherent theory of dynamic neural organization. My purpose is simply to state that computational tolerance for the time-staggered arrival of integrable neural information, and skill acquired over many long bouts of practice, are likely to be more important than is compensation.

## Prediction and postdiction: Two frameworks with the goal of delay compensation

doi: 10.1017/S0140525X08003889

David M. Eagleman

*Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030.*
**eagleman@bcm.edu**
**http://neuro.bcm.edu/eagleman**

**Abstract:** Although prediction is one of the key tasks of intelligent brains, it often proves impossible in an unpredictably changing world. Hence, brains often decide what happened retrospectively. This framework of postdiction, the opposite of prediction, stands as an alternative or complimentary framework to prediction. I further show how motor-sensory recalibration demonstrates delay compensation at the perceptual level.

To construct an up-to-date, dynamic picture of the outside world, brains have to overcome a problem: different sensory signals are processed at different speeds in different parts of the brain. As Nijhawan emphasizes, all this processing takes time – time during which the outside world continues to change. One clever method by which nervous systems can reduce this problem is by predicting, whenever possible, the next state of the world (see Hawkins & Blakeslee 2004; Nijhawan 1994; Suddendorf & Corballis 2007; and the present target article). Nijhawan has long championed the idea that prediction happens at the level of the perceptual system (e.g., Nijhawan 1994), mitigating temporal delays at early, premotor stages. My comments here reinforce an emphasis on the brain's need to compensate for delays, but I offer some additional methods that the brain may use beyond prediction.

***Prediction and postdiction: Two frameworks with the goal of compensation.*** In conjunction with any prediction it attempts, the brain also appears to implement a *postdictive* strategy (Eagleman & Sejnowski 2000). This term simply means that the brain continues to collect information after an event (such as a flash) while settling on its best estimate about events and their timing in the outside world. Postdiction is easily illustrated by apparent motion: When the first stimulus disappears, one cannot perceive the motion until after the second stimulus has appeared (in an unpredicted location). Nonetheless, one has the illusion of seeing it move on the fly. Obviously, this can only happen when perception is generated retrospectively (Eagleman 2001).

In the context of the flash-lag effect (FLE), three empirical findings seem to support postdiction: the flash-initiated (Khurana & Nijhawan 1995), flash terminated (Nijhawan 1992), and motion-reversing (Whitney & Murakami 1998) conditions – all of which illustrate that the motion information after the flash, not before it, drives the effect (Eagleman & Sejnowski 2000). Nijhawan suggests that these challenges might be accounted for by a rapidly arbitrated competition between rivaling neural representations. Although he avoids the term postdiction, a biased competition model nonetheless requires it. The competition between a predictive model and new incoming information cannot be tipped until the new information is collected after the flash. No matter how rapidly the arbitration occurs, it still is postdictive.

There are competing explanations for the FLE that do not require the predictive component at all. For example, one alternative draws on the fact that motion signals shift localization judgments (for a review, see Whitney 2002). We have shown that the motion signals collected after an event are the ones that bias perceived position (Eagleman & Sejnowski 2007). Our evidence supports the framework that incoming motion information updates localizations so that the data do not become stale (for details, see Eagleman & Sejnowski 2007). This is a purely postdictive explanatory framework for the FLE.

To understand why postdiction is necessary, one only has to examine the tasks and resources of the visual system. As one of its tasks, the visual cortex tries to get the timing of outside events correct; for its resources, it has to deal with the foibles of the populations that feed it. The retina and visual thalamus have circuitries that cause incoming signals to temporally spread from the first stages of the visual system based, for example, on their luminance (Gawne et al. 1996; Lee at al. 2007; Maunsell et al. 1999). For the visual brain to correctly align the timing of events in the world, it may have to wait ~100 msec for the slowest information to arrive – thereby allowing the visual system to discount different delays imposed by the early stages.

The reason for the vigorous debate between prediction and postdiction is that both sides enjoy strong support: On the one hand, brains use every available opportunity to predict what is coming next, thus saving valuable processing time; on the other hand, experiments show that perceptions can be changed retrospectively, as may be necessitated by an unpredictably changing world.

The important thing to note is that both the predictive and postdictive frameworks have the same goal: to perceptually place the location of a moving object closer to its real-world position. Prediction does this by guessing ahead; postdiction does this by allowing incoming motion signals to advantageously adjust location judgments. Both techniques lead to the perception of a moving object closer to its real-world position. The advantage of postdiction is its ability to naturally account for sudden motion reversals such as bounces and ricochets; however, it is possible that both mechanisms are implemented in different contexts or in different areas of the nervous system. Future experiments are required to determine these details.

***Perceptual compensation for temporal delays.*** Finally, I want to expand on part of Nijhawan's argument. He correctly points out that compensation can take place at the level of the motor output or the perceptual input (or anywhere in between). He argues on the side of perceptual compensation, a position that has traditionally received little attention, in part because of the paucity of direct experimental support. However, in his target article he leaves out a critical example that supports perceptual compensation: the recalibration of motor sensory systems when exposed to delayed feedback.

The recalibration story begins with a mystery: Given that multisensory signals arrive in the brain at different times, how can the brain decide which events were supposed to be simultaneous in the outside world? We have proposed that the brain perceptually recalibrates its expectations about the arrival times of signals by employing a simple assumption: when the motor system executes an act (such as knocking on a door), all the resulting feedback should be assumed to be simultaneous, and any delays should



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

be perceptually adjusted until simultaneity is perceived (Stetson et al. 2006). Although this recalibration is normally adaptive, note a strange consequence: Imagine we repeatedly inject a tiny delay (e.g., 100 msec) between a subject's key-press and a subsequent flash for a dozen trials, and then we suddenly remove the delay. Because the perceptual systems have recalibrated to compensate for the delay, the subject now perceives that the flash occurred before the key-press: an illusory reversal of action and sensation (Cunningham et al. 2001a; Stetson et al. 2006). Note that this recalibration is no mere party trick – it is critical to solving the problem of causality, which requires, at bottom, temporal order judgments. The only way causality can be accurately determined in a multisensory brain is to keep the temporal delay of signals calibrated in the face of different sensory pathways operating at different speeds. This example serves to buttress Nijhawan's general argument that compensation for delays can take place at very early, perceptual levels.

## Transient signals per se do not disrupt the flash-lag effect

doi: 10.1017/S0140525X08003890

Piers D. Howe, Todd S. Horowitz, and Jeremy M. Wolfe
*Visual Attention Lab, Brigham and Women's Hospital, Cambridge, MA 02139.*
**howe@search.bwh.harvard.edu**
**toddh@search.bwh.harvard.edu**
**wolfe@search.bwh.harvard.edu**
**http://search.bwh.harvard.edu/**

**Abstract:** Nijhawan's theory rests on the assumption that transient signals compete with predictive signals to generate the visual percept. We describe experiments that show that this assumption is incorrect. Our results are consistent with an alternative theory that proposes that vision is instead *postdictive*, in that the perception of an event is influenced by occurrences after the event.

Nijhawan has presented a predictive theory of visual motion perception. He notes that because of the delays inherent in visual processing, the visual system cannot detect events at the instant at which they occur. He argues that the visual system compensates by using the information it has to predict the current state of the world. In other words, he suggests that we do not see reality, but rather, the visual system's best guess at reality is based on slightly out-of-date information. To support this theory, Nijhawan relies heavily on a particular visual illusion – the flash-lag effect. Here, we briefly describe experiments that Nijhawan's account cannot handle. Although Nijhawan's predictive theory can readily account for the standard flash-lag paradigm, the flash-terminated version poses a problem. In this version of the paradigm, the moving object disappears at the time of the flash. As Nijhawan notes in the target article, "At the outset, spatial extrapolation should continue beyond the time of unpredictable disappearance of the object. Yet these displays produce no flash-lag effect" (sect. 5.2.2, para. 1). To explain why the flash-lag effect does not occur under these circumstances, Nijhawan makes an additional assumption. He proposes that the transient signal produced by the disappearance of the moving object generates a neural representation (perhaps thalamic) that competes with and eventually overwhelms the predictive representation (perhaps cortical) that would normally cause the flash-lag effect. We will refer to this as the transient signal assumption.

The support for this assumption comes from Maus and Nijhawan (2006, reviewed in section 5.2.2 of the target article). That study did not employ a flash-lag design. Instead, Maus and Nijhawan demonstrated that when a moving object disappears gradually by moving behind a variable neutral density filter, it remains visible past its standard luminance threshold. This finding could be caused by hysteresis (Palmer 1999). In any case, it does not address the issue of whether the presence of a transient signal eliminates the flag-lag effect (i.e., the transient signal assumption).

To test this assumption, we used a reversed-contrast version of the flash-lag paradigm in which the background was gray and the moving object changed from black to white at the moment of the flash. The resultant transient signal must be at least as large as in the flash-terminated version (probably larger, since the neurons in the thalamus are sensitive to contrast polarity; Hubel & Wiesel 1961). If Nijhawan's transient signal assumption is correct, this should abolish the flash-lag effect. Nevertheless, the magnitude of the effect was just as strong with our stimulus as in the standard flash-lag paradigm (i.e., with no transient).

Perhaps Nijhawan's proposed transient signals do not originate in the thalamus and are generated by cells that are not sensitive to contrast polarity (Hubel & Wiesel 1962). Such cells respond as strongly to a white dot as they do to a black dot and so may not generate a transient signal in the reversed-contrast version of the flash-lag effect. If it is assumed that it is these cells that need to generate the relevant transient signals, then this might explain why the flash-lag effect still occurs in our reversed-contrast version of the paradigm. We tested this hypothesis by comparing two versions of the flash-lag paradigm. In one case, the Michelson contrast of the moving object was reduced from 0.30 to 0.04 at the moment of the flash. In the other case, it was reduced from 0.04 to 0; that is, it disappeared. The first condition should produce a larger transient than the second, yet we found the flash-lag effect to be of normal strength in the first case but entirely absent in the second. It appears that, contrary to the transient signal assumption, the flash-lag effect is not disrupted by transients but only by the actual disappearance of the moving object.

Why, then, does the flash-lag effect vanish in the flash-terminated condition? Nijhawan proposes his transient signal assumption as the reason, but our data fail to support that account. Our data are more compatible with the *postdictive* theory of Eagleman and Sejnowski (2007). In brief, this theory postulates that the apparent position of the moving object at the time of the flash is influenced by the motion of the object after the flash. In other words, consistent with our observations, the flash-lag effect should occur in any situation in which the moving object does not disappear, irrespective of any transient signals that might be generated at the time of the flash. This theory can readily account for all our data without invoking any additional mechanisms.

## Mental and sensorimotor extrapolation fare better than motion extrapolation in the offset condition

doi: 10.1017/S0140525X08003907

Dirk Kerzel[a] and Jochen Müsseler[b]
[a]*Faculté de Psychologie et des Sciences de l'Éducation, Université de Genève, 1205 Genève, Switzerland;* [b]*Psychology Department, RWTH Aachen University, 52056 Aachen, Germany.*
**dirk.kerzel@pse.unige.ch**
**http://www.unige.ch/fapse/PSY/persons/kerzel/**
**muesseler@psych.rwth-aachen.de**
**http://www.psych.rwth-aachen.de/ifp-zentral/front_content.php?idcat=222**

**Abstract:** Evidence for motion extrapolation at motion offset is scarce. In contrast, there is abundant evidence that subjects mentally extrapolate the future trajectory of weak motion signals at motion offset. Further, pointing movements overshoot at motion offset. We believe that mental and sensorimotor extrapolation is sufficient to solve the problem of perceptual latencies. Both present the advantage of being much more flexible than motion extrapolation.

Nijhawan claims that the offset of a smoothly moving object masks the extrapolated trajectory of the moving object. Therefore, the flash-lag effect is suppressed in the flash-terminated



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

condition. In an experiment using gradual variation of luminance across the trajectory, the final position was found to be misperceived in the direction of motion (Maus & Nijhawan 2006). The reason given in this study was that the gradual variation of luminance reduced the transient at stimulus offset. Contrary to Nijhawan's hypothesis, forward displacement of the final position of a moving target has been repeatedly observed with strong offsets. Jennifer Freyd (1987) was the first to observe displacement of the final position of an object undergoing implied motion (see overview in Hubbard 2005). She presented a stationary rectangle for 0.25 sec at three different orientations. The successive views implied the rotation of the object in a certain direction. Each view of the object was separated by a blank interval of 0.25 sec, and the remembered final position was probed after a retention interval of 0.25 sec. It is unlikely that low-level motion receptors were stimulated in this paradigm. The stimuli did not even reliably evoke an impression of apparent motion. Thus, the observed localization error has to be due to high-level, cognitive processes. In fact, when the motion type was systematically varied from implied to smooth motion by reducing the stimulus-onset-asynchrony (SOA) of consecutive presentations, the forward extrapolation was found to decrease: The worse the quality of the motion signals, the larger the error (Kerzel 2003). The interpretation was that observers who are confronted with intermittent motion signals, may engage in "mental extrapolation" where they try to predict the upcoming target positions. Mentally extrapolating the future trajectory induces errors in judging the last seen object position.

Nijhawan argues that a gradual reduction of luminance contrast reveals motion extrapolation because masking by the transient at target offset is reduced (Maus & Nijhawan 2006). However, it is just as likely that these conditions reveal mental extrapolation that is induced by the absence of a clearly visible offset. Under conditions of high uncertainty, observers are more likely to predict the stimulus position based on what they expect, rather on the basis of what they see. Thus, the gradual luminance variation may bring high-level and not low-level extrapolation mechanisms to the fore.

Further, one may wonder how much motion extrapolation contributes to accurate sensorimotor control. Let us consider the case of a subject placed in front of a touch-sensitive monitor. When prompted to rapidly hit a moving object with his index finger, the subject will take ~0.3 sec to initiate the movement, and another ~0.35 sec to transport the hand from a home-key 20 cm in front of the monitor to the screen surface. Thus, a total of ~0.65 msec elapses from the moment the imperative signal is presented to the moment the screen is touched. It is clear that subjects who plan to hit the currently visible position will miss the target considerably (at the most: 0.65 sec $\times$ $v$, where $v$ is target velocity). Hence, accurate interception of moving objects necessarily implies prediction of future positions. An intriguing hypothesis is that sensorimotor computations factor in visual delays. That is, interceptive actions are directed at the position the target will have reached in the time required to move the hand to that position plus visual delays (at the most: $[0.65 + 0.1$ sec$] \times v$, with visual delays equaling 0.1 sec). Pointing to the offset position of a moving object should isolate sensorimotor compensation for visual delays, as the system does not need to take into account the future trajectory. In fact, pointing movements overshoot the final position by a distance that roughly corresponds to the flash-lag effect (Kerzel & Gegenfurtner 2003). In contrast, perceptual judgments were centered on the true offset position. Thus, we believe that visual delays are not treated differently from other types of delays (time to decide, time to move, etc.). The sensorimotor system aims at positions further ahead that correspond to the respective delays.

A characteristic of sensorimotor compensation is that it is flexible and task-dependent. In a set of experiments on the Fröhlich illusion (Fröhlich 1923), we observed that localization judgments (mouse pointing) were affected by rapidly changing visuomotor strategies. In the classical Fröhlich illusion, the perceived onset position of a moving object is displaced in the direction of motion. Possibly, the first part of the trajectory is missed because of attentional latencies and metacontrast. When the onset position is predictable because the target always appears in two narrow regions of space, this error is reproduced with pointing movements. However, when the onset position is highly unpredictable because the target appears randomly in a large region of space, the error is eliminated or even slightly reversed (Müsseler & Kerzel 2004). Our interpretation was that with high uncertainty, subjects try to correct for having missed the initial part of the trajectory by pointing to positions opposite to the direction of motion. Analysis of the time course showed that the effect of uncertainty emerged after only about 10 trials. That is, subjects changed their response strategy as soon as they noticed that the target position appeared in random places. Further, perceptual judgments, which were unaffected by predictability when run in a separate block of trials, showed the same effect of predictability when randomly intermingled with pointing movements (Müsseler et al., in press). Thus, visuomotor strategies may affect the way retinal stimulation is evaluated.

Given the capacity to rapidly adapt visuomotor translation to changing circumstances, one may wonder whether motion extrapolation is needed to deal with visual latencies. Although it cannot be ruled out that such a mechanism exists, assuming its existence does not seem parsimonious. Other, more flexible mechanisms may do the job. Our proposal is that mental extrapolation of target motion and sensorimotor predictions solve the problem of visual latencies.

ACKNOWLEDGMENTS

Dirk Kerzel was supported by the Swiss National Foundation 10011-107768/1 and PDFM1-114417.

## What's in a name change? Visual prediction makes extrapolation real and functional

doi: 10.1017/S0140525X08003919

Beena Khurana

*Department of Psychology, University of Sussex, Falmer, East Sussex, BN1 9QH, United Kingdom.*

**beenak@sussex.ac.uk**

**http://www.sussex.ac.uk/psychology/profile116416.html**

**Abstract:** Nijhawan redraws our attention to the problem of accurately perceiving an ever-changing visual world via a sensory system that has finite and significant communication times. The quandary is compelling and stark, but the suggestion that the visual system can compensate for these transmission delays by *extrapolating the present* is not so unequivocal. However, in this current airing of contradictory issues, accounts, and findings, Nijhawan trades *spatial extrapolation* – a rather specific concept introduced earlier (in Nijhawan 1994) for *visual prediction* – a far more expansive notion that forces the issue of both the perceived reality and functional significance of compensation.

The discrepancy between visual information being delayed and behavior being accurate is unquestionable. It therefore might seem odd to take up the functionality of visual compensation. However, why should there be sensory compensation when later end effectors can compensate adequately? Of course, here one might ponder whether perceiving at all has functional significance, given that actions can be undertaken in the absence of accurate vision in terms of position information (Goodale et al. 1986). It is at this juncture in the present exposition that re-branding spatial extrapolation as visual prediction makes a tangible and productive difference.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

In its current conception visual prediction does not treat vision as separate and independent of action. In fact, Nijhawan now argues for extrapolation being a two-stage process in that compensation first occurs in the dorsal action pathway and then is communicated to the ventral perception pathway. In this way, any conflict between action and perception systems is obviated. While the workings of the action system are grounded in accurate behavior, how does one go about proving the reality of extrapolated percepts?

The tack that I favor is to look for perception that varies lawfully as a function of spatial offset. If certain spatial percepts without recourse to extrapolation result in already established patterns of performance, and these very percepts can be created via extrapolation to render equivalent patterns of performance, then extrapolated percepts are no different than non-extrapolated ones. Furthermore, if the perceptual reality of the non-extrapolated percepts is unquestionable, then logically the same must hold for extrapolated percepts.

A spatially modulated phenomenon in face processing provides the perfect candidate for assessing the perceived reality of extrapolated vision. When two face halves belonging to two different individuals are spatially aligned, the resulting "facial chimera" is difficult to read for the identity of either face half (Young et al. 1987). However, the introduction of a spatial offset or misalignment greatly aids identification of the constituent face halves. Observers viewed facial chimeras that consisted of one moving face half and one flashed face half (the flash-lag effect [FLE] facial chimera) in a flash-initiated display with instructions to report either the offset or identify of the flashed half. Observers both "perceived" the face halves to be spatially offset and also produced identification times that were faster, as is typical for constantly present spatially offset face halves. Now, if one does not question the account given for the chimeric face effect without extrapolated stimuli, then the same should apply to stimuli that are a result of extrapolation. The observed reaction time pattern with FLE facial chimeras is clearly indicative of extrapolation being part and parcel of the ventral system and perception (Khurana et al. 2006). The analogous logic applies to a situation in which the extrapolated percept is one of alignment and observers both perceive stimuli to be aligned and produce responses that are concordant with the percept.

Although these findings with FLE facial chimeras provide clear evidence for extrapolation in the ventral stream by measuring not only perceived offset but also the knock-on effects on identification latencies, other findings do not sit so comfortably with the notion that the flow of extrapolated information is from the dorsal pathway to the ventral pathway. It was also shown that when the two face halves are of the same familiar individual, the FLE is lawfully smaller. In the present context, these findings serve as counter examples. If $R_D$ is the dorsal representation of FLE facial chimeras and $R_V$ is the ventral representation in higher cortical face areas, then the smaller perceived spatial offset between face halves that belong to the same individual versus those that belong to different individuals can be accounted for as follows. The extrapolated moving face half makes contact with an underlying complete representation of the known face, thus priming the registration of the flashed face half. A reduction in the perceived spatial offset of same face halves is indicative of influences of ventral face processing on spatial dorsal computations. It might be that Nijhawan's proposed two-stage account is in actual fact two parallel processes – one dorsal and one ventral, each equally capable of influencing the other.

Until recently, the face data could be thought of as an existence proof of ventral influences on FLE; however, other well-learned representations can also reduce the magnitude of the FLE (Noguchi & Kakigi 2008). For Japanese observers, it was shown that perceived FLE for Kanji character segments was significantly reduced relative to non-meaningful geometrical segments. Importantly, this diminution was neither observed for pseudo-Kanji characters viewed by Japanese observers nor for Kanji characters viewed by non-Japanese English-speakers with no knowledge of Kanji. Thus, both faces and meaningful complex visual forms can affect spatial computations.

Although the flash-lag effect is often likened to other non-veridical percepts such as the Fröhlich effect and representation momentum, the functional significance of extrapolation lies in the visual system generating a percept that matches reality at the time of perceptual awareness. Both the examples above of cortical post-constancy late representations superseding retinal pre-constancy early representations in visual awareness are indicative of extrapolated percepts, having real consequences for behavior.

## Perception of direction is not compensated for neural latency

doi: 10.1017/S0140525X08003920

Bart Krekelberg

*Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102.*

**bart@rutgers.edu**

**http://vision.rutgers.edu/klab**

**Abstract:** Neural activity in the middle temporal area (MT) is strongly correlated with motion perception. I analyzed the temporal relationship between the representation of direction in MT and the actual direction of a stimulus that continuously changed direction. The representation in MT lagged the stimulus by 45 msec. Hence, as far as the perception of direction is concerned, the hypothesis of lag compensation can be rejected.

It is a truth universally acknowledged, that a brain aiming to survive in a fast-changing environment must compensate for time delays. The bone of contention is the processing stage at which this compensation takes place. Is what we perceive compensated for delays? Or do we perceive the world as it was some time ago and merely adjust our actions so that they are appropriate now?

Motion perception is a domain in which latency compensation could be useful. Imagine a situation where a monkey is chased by a tiger. The tiger circles the monkey and continuously changes its direction. It is advantageous to the monkey to extrapolate the motion changes, such that it can react in a timely manner. Doing this extrapolation in the motor system would provide all the benefits to survival. Nijhawan's hypothesis is that the monkey perceives the tiger to move in the direction it moves "now," rather than the direction in which it moved some time ago.

One major advantage of investigating delay compensation with motion perception is that – unlike the perception of position studied in the flash-lag effect – motion perception has been studied extensively at the neural level. There is convincing evidence to link the middle temporal area (MT) to the perception of the direction of motion (for review, see Born & Bradley 2005). This strong association between MT activity and perceived direction allows me to pose the question whether visual delay compensation exists, in a very specific manner as follows: When a stimulus continuously changes direction, does MT encode the current stimulus direction, or a direction of motion that happened earlier?

Figure 1A shows the responses of a single MT neuron to a random dot pattern that started out moving downward, but smoothly changed its direction of motion in a counterclockwise manner. During one second, all 360 degrees of motion direction were traced out (Krekelberg & Albright 2005; Schoppmann & Hoffmann 1976).

The perception of direction is related to MT activity through a labeled line model (Born & Bradley 2005). In this model, each



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

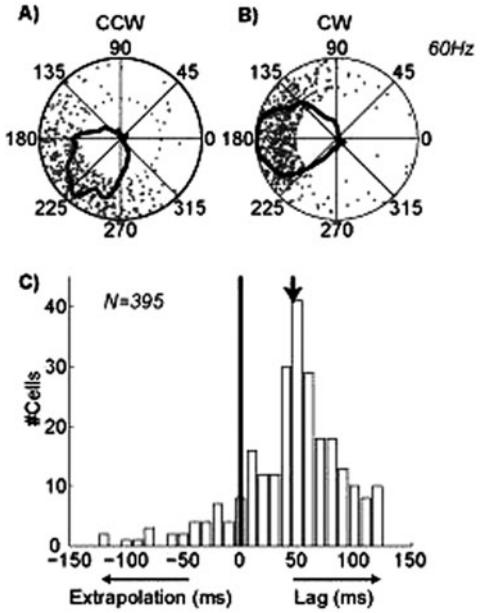



Figure 1 (Krekelberg). The lagging neural representation of stimulus direction in the middle temporal area. (**A**) Counterclockwise (CCW) changes in direction. The direction of stimulus motion is shown on the angular axis; dots represent spikes recorded while the stimulus moved in a given direction. Each concentric circle of dots represents a separate trial. The solid line shows the neural response averaged over trials; the radial distance represents the firing rate. (**B**) Responses to clockwise (CW) changes in direction. (**C**) Histogram of lags in the representation of the direction of motion by single middle temporal area neurons. On average (arrow), the population of middle temporal area cells lagged 45 msec behind the stimulus direction.

spike counts as a vote in favor of a perceived direction of motion that corresponds to the label of the cell. I will show how the assumption of latency compensation in this model leads to a contradiction.

I choose the label of the neuron in Figure 1A such that its spikes represent the direction of motion as it is "now," rather than how it was some time ago. It is easy to see that this lag-compensated label should be ~225 degrees – the direction of motion for which most spikes are recorded.

Now consider the response recorded when the same visual pattern initially moved upward but then gradually changed its direction of motion in a clockwise direction (panel B). For this trajectory, the largest number of spikes occurred when the stimulus moved in the 180 degrees direction. The label of the neuron does not change; hence, this peak activity signaled that a stimulus moved towards 225 degrees. Because stimulus direction changes at a rate of 360 degrees per second, the neuron is (180°–225°)/360°/s = 125 msec too late in signaling this direction of motion.

Clearly the interpretation that a neuron would compensate for its response latency for clockwise changes in direction, but then respond 125 msec too late for counterclockwise motion, is nonsensical. The correct interpretation must be that the neuron's true label is the direction that lies halfway between the maximum response direction for clockwise (CW) and counterclockwise (CCW) motion (~203°), and the response lags the stimulus direction equally (62.5 msec) for both clockwise and counterclockwise motion.

This analysis does not suffer from the problems associated with comparing flashed and continuous stimuli, because it involves no onset or offset transients but compares two equivalent patterns of motion. If a neuron perfectly compensated for lag, the two tuning curves for clockwise and counterclockwise circular pathways shown in panels A and B should be the same. To quantify the similarity between the curves, I calculated the difference between the stimulus directions at the peaks of the average response curves. Half this difference, divided by the speed of the change in direction of motion, equals the time by which the neural representation lags the stimulus. Figure 1C shows a population histogram of lags derived from 395 MT neurons recorded in three awake, fixating macaques. In this histogram, neurons that compensated for neural latencies are found at lag = 0. Clearly, there were few such neurons, and on average the MT population lagged the stimulus direction by 45 msec.

Neural activity in MT faithfully tracks the percept of motion, even when that motion is illusory (Krekelberg et al. 2003; 2006; Schlack & Albright 2007). As such, it is clearly a perceptual area. The analysis presented here, however, demonstrates that the representation of motion in MT lags 45 msec behind stimulus direction. If this lag for a predictable stimulus were less than the lag for an unpredictable stimulus such as a flash, one could still argue that MT does some (albeit incomplete) lag compensation. Data from many laboratories, however, show that abrupt visual onsets can reach MT cells in approximately 40 msec (Bair et al. 2002; Petersen et al. 1985; Price et al. 2005). This suggests that – as a population – MT does not compensate for lag at all.

These findings do not exclude the possibility that there is some compensation for time delays in other perceptual subsystems. However, motion perception is a subsystem in which compensation might be expected to benefit the organism most. If we cannot find evidence for lag compensation here, it seems prudent to conclude that while we may act in the present, we perceive the past.

ACKNOWLEDGMENTS

I thank Till Hartmann, Markus Lappe, and Tamara Watson for their helpful comments and discussion, and gratefully acknowledge the financial support of the Pew Foundations and the National Eye Institute (1R01EY017605).

## Unconscious inference and conscious representation: Why primary visual cortex (V1) is directly involved in visual awareness

doi: 10.1017/S0140525X08003932

Zhicheng Lin

*Department of Psychology, University of Minnesota, Minneapolis, MN 55455.*
**linxx443@umn.edu**
**http://zhichenglin.googlepages.com/**

**Abstract:** The extent to which visual processing can proceed in the visual hierarchy without awareness determines the magnitude of perceptual delay. Increasing data demonstrate that primary visual cortex (V1) is involved in consciousness, constraining the magnitude of visual delay. This makes it possible that visual delay is actually within the optimal lengths to allow sufficient computation; thus it might be unnecessary to compensate for visual delay.

The time delay problem – that perception lives slightly in the past as a result of neural conduction – has recently attracted a considerate amount of attention in the context of the flash-lag effect. The effect refers to a visual illusion wherein a brief flash of light and a continuously moving object that physically align in space and time are perceived to be displaced from one another – the flashed stimulus appears to lag behind the moving object (Krekelberg & Lappe 2001). In the target article, Nijhawan compellingly argues that delay compensation could be undertaken by a predictive process in the feedforward pathways in the vision system. Before jumping into the quest for the mechanism of delay compensation, however, I would like to



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

argue that the magnitude of delay has been overestimated, and that it might even be unnecessary to compensate for such a delay.

Although time delay poses a difficult issue for visual motion perception in particular, there is nevertheless the benefit to allow large amounts of computation to take place unconsciously. Presumably, evolution must have set optimal delay lengths that are short enough to avoid too much delay, but long enough to allow sufficient amounts of computation. If visual delay is within such optimal lengths, compensation might be unnecessary. The key issue, then, is to determine the magnitude of delay inherent in the visual system. As mentioned by Nijhawan, the perceptual lag of a moving object due to neural delay can be calculated as $v \times t_{max}$, where $v$ is object velocity and $t_{max}$ is the maximum cumulative delay in the vision pathway before conscious percepts emerge. The magnitude of delay is therefore determined by $t_{max}$. This boils down to this question: To what extent can the visual process proceed in the visual hierarchy without awareness? The farther the unconscious visual process proceeds, the larger is the perceptual lag that occurs.

Because humans are not aware of the delay compensation process, it is tempting to speculate that the brain areas subserving it might be unrelated to awareness. Indeed, neural activity at the level of photoreceptors is not sufficient to yield visual awareness, and some researchers further argue that later stages after the retina, such as primary visual cortex (V1), are not involved in consciousness. This logic is problematic, however, and can be illustrated more clearly when we take a step back to examine a more general problem in visual perception – what Hermann von Helmholtz called unconscious inference (i.e., before percepts emerge, huge amounts of computational processes take place without consciousness). A key notion here is that unconscious inference does not imply that the neural substrates are unrelated to consciousness. For example, in the binding problem, a distributed hierarchical network in the brain processes different features of an object. The neural substrates that correctly select and integrate separate neural representations of features to form coherent object representations are believed to comprise both parietal and frontal cortex (Treisman 1996). Can we thus infer that activity in parietal and frontal cortex is unrelated to consciousness? Apparently this is not the case (Lin & He, in press).

Yet, the notion that observers are not directly aware of neural activity in the optic nerve, the lateral geniculate nucleus (LGN), and even V1 has been popular among many researchers (e.g., sect. 3 of the target article), especially after the appearance of the influential paper by Crick and Koch in 1995. Although this idea (Crick & Koch 1995) was initially based on behavioral studies that show equal adaptation strengths for both visible and invisible conditions (e.g., He et al. 1996), recent evidence demonstrates, instead, that V1 is directly involved in visual awareness (for a review, see Lin & He, in press). Monkey single-unit recordings, human electroencephalogram (EEG) recordings, and brain imaging converge to reveal robust awareness-dependent modulations in V1; neural events in V1 are actually attenuated in response to suppressed (vs. non-suppressed) visual stimuli. Convergent psychophysical evidence also comes from *binocular rivalry*, the phenomenon that when each of the two eyes views dissimilar images, the two images rival for visibility, with one temporarily dominating perception for several seconds before being replaced in dominance by the other. With adaptation, a recent study shows that rivalry suppression weakens the strengths of aftereffect for low-contrast adaptors (Blake et al. 2006), which is believed to transpire in V1. It appears that although unconscious inference per se is beyond consciousness, neural activity in V1 is actually correlated with consciousness.

Where, then, does consciousness enter into the information processing sequence? How can one distinguish conscious processing from unconscious processing in the brain? It is important to know which neural substrates of unconscious inference are conscious and which are unconscious, as measured by correlations between neural activity and conscious experiences. This appears extremely difficult, given the unclear mechanisms of how the brain computes unconscious inference. In addition, it is almost impossible to isolate the inference processes (e.g., feature binding), since we only know that there are such processes after we are conscious of the objects. Moreover, although many studies focus on brain activation differences between conscious and unconscious experiences, it is an oversimplification to classify certain areas as conscious or unconscious based on such comparison. For example, a recent study shows that although V1 activity is closely correlated with conscious experience, even a considerable difference in cortical activation in many visual cortical areas does not necessarily lead to different conscious experiences (Jiang et al. 2007). It therefore appears that no single area, or even network, is sufficient for consciousness. A more dynamic approach to distinguish consciousness from unconsciousness is to go beyond isolated activity of certain brain areas, and, instead, focus on how different brain areas interact with each other. Lamme and Roelfsema (2000) suggest that when information first enters the early visual cortex in an feedforward sweep, it is unconscious; when such information is attentively grouped through recurrent processing (by means of horizontal connections and feedforward/feedback projections), it gives rise to consciousness. This might provide important clues as to what give rise to consciousness.

In sum, although the neural correlates of consciousness are still elusive, it is now well documented that V1 is directly related to awareness (Lin & He, in press). This has important implications for the magnitude of time delay by constraining $t_{max}$ (the maximum cumulative delay in the vision pathway before conscious percepts emerge) to less than 72 msec on average in V1 neurons (Lamme & Roelfsema 2000). Such a short delay might be within optimal delay lengths to allow sufficient computations to take place, and might even make it unnecessary to compensate for the delay.

ACKNOWLEDGMENTS

The preparation of this commentary was supported by a University of Minnesota Graduate School Block Grant to Zhicheng Lin, and by the James S. McDonnell foundation and the U.S. National Institutes of Health (EY-015261-01) grants to Dr. Shang He (Lin's current advisor).

## Neuronal adaptation: Delay compensation at the level of single neurons?

doi: 10.1017/S0140525X08003944

J. Patrick Mayo and Marc A. Sommer

*Department of Neuroscience, Center for the Neural Basis of Cognition, Center for Neuroscience, University of Pittsburgh, Pittsburgh, PA 15260.*

**jpm49@pitt.edu**

**masommer@pitt.edu**

**Abstract:** Saccades divide visual input into rapid, discontinuous periods of stimulation on the retina. The response of single neurons to such sequential stimuli is neuronal adaptation; a robust first response followed by an interval-dependent diminished second response. Adaptation is pervasive in both early and late stages of visual processing. Given its inherent coding of brief time intervals, neuronal adaptation may play a fundamental role in compensating for visual delays.

Nijhawan presents a compelling argument in support of the existence of delay compensation in visual processing. Specifically, he draws a distinction between "discrete" and "continuous" stimuli and focuses on the flash-lag effect as a useful illustration of differential processing between the two. Our laboratory investigates an important class of stimuli that extends the discrete-continuous continuum. We use sequential stimuli, consisting of consecutive stimulation separated by



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

brief blank periods. This kind of visual stimulation is frequently encountered by the retina.

Humans make rapid eye movements, or saccades, approximately three times per second. These movements shift the position of the world on the retina. Thus, the seemingly simple act of looking at an object, such as a face, involves the successive repositioning of its image features on disparate locations of the retina. Importantly, we do not perceive these shifts of the visual world or the period of lost time during a saccade, because of a general phenomenon known as saccadic suppression. Natural vision is therefore largely comprised of stable visual input that is frequently interrupted by brief periods of saccadic suppression (although continuous stimulation resulting from object tracking is possible in the form of "smooth pursuit" movements; for review, see Krauzlis 2004). Saccadic suppression adds appreciable "dead time" to the normal delays that Nijhawan reviews; yet we fail to notice all of the lost perceptual time just as we fail to be disturbed by the afferent delays. Instead, we perceive a world that is spatially and temporally stable.

How do we maintain this two-dimensional visual stability in the face of the brain's physical limitations? Nijhawan discusses feedforward signaling as one way of coping with visual delays, and our earlier work demonstrated the existence of such signaling in a pathway from the midbrain to prefrontal cortex (Sommer & Wurtz 2002; 2006). This pathway seems to help keep the brain and retinotopic visual world in spatial register before and after a saccade. Our recent work has focused on how the brain maintains temporal stability (Mayo & Sommer, submitted). We sequentially present two visual probes in the receptive field of neurons in awake, fixating monkeys. We systematically vary the amount of time between probe presentations (interstimulus interval, or ISI). The purpose is to mimic the fixation-saccade-fixation timing of natural vision. We have recorded neuronal activity in two brain areas: the retinally recipient superficial layers of the superior colliculus (supSC) and the frontal eye field (FEF), a "higher" visuomotor area in prefrontal cortex.

Sequential stimulation elicited sequential visual responses in neurons, with each response separated by a period of baseline activity equivalent to the ISI. Therefore, the timing of neuronal responses accurately encodes the timing of real-world visual stimulation. However, the relative magnitude of the neuronal responses also varied significantly with ISI (Fig. 1). Second visual responses were only a fraction of the first; this is a well-known phenomenon called neuronal adaptation.

We suggest that neuronal adaptation may be useful for timing, because it carries information about the ISI during sequential visual stimulation. Extremely short ISIs (<50 msec) lead to negligible second neuronal responses (Fig. 1, left), whereas longer ISIs lead to monotonically increasing responses (Fig. 1, right) until ISI ≈ 400 msec, where second responses attain normal magnitudes (Fig. 2). We found that adaptation followed the same time-course in FEF and supSC, but was stronger overall in FEF. As a point of reference with previous studies, our supSC data matched those reported by other laboratories (Robinson & Kertzman 1995; Wurtz et al. 1980).

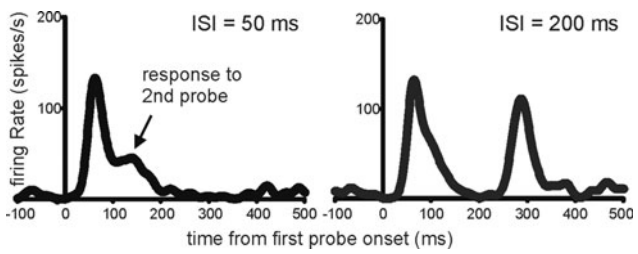



Figure 1 (Mayo & Sommers). Example visual responses from a small group of single neurons in superior colliculus. The results of using two probes presented at interstimulus intervals of 50 msec (left) and 200 msec (right) are shown.

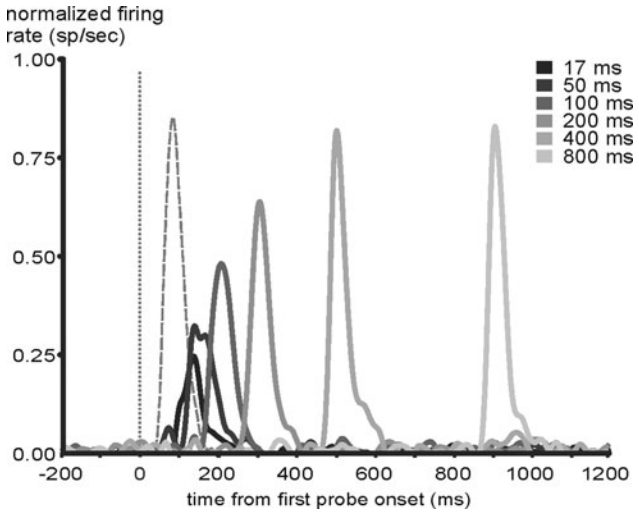



Figure 2 (Mayo & Sommers). Normalized visual responses in the superior colliculus population. Vertical dotted line indicates time of first probe onset. Grayscale lines depict mean responses to the second probe after sequential, paired stimulation, at various interstimulus intervals. Dashed line represents typical response to first probe.

In light of the apparent need for a compensatory physiological mechanism as described by Nijhawan, we posit that neuronal adaptation may act as an alternative source of latency encoding. As the target article explains, the exact timing of neuronal responses in identical conditions can vary by tens of milliseconds from trial to trial. Neurons downstream of these time-varying sensory responses therefore receive an imprecise measurement of stimulus onset. However, for sequential stimuli presented within 400 msec of each other, downstream neurons could corroborate the raw timing of the neuronal responses with relative magnitude comparisons of the incoming neuronal responses. Neuronal adaptation therefore may function as a complementary latency code. The presence of this code in ecologically relevant scenarios – in response to sequential stimuli, and at presentation times that mimic typical periods of fixation and saccades – lends further support to our hypothesis.

Using the relative magnitudes of sequential visual responses as a clue to the timing of visual stimuli requires an assumption that the sequential visual stimuli themselves are nearly identical. Accepting this assumption means that differences in visual responses that occur within 400 msec of each other must be due to timing. The assumption seems valid for many real-world situations, such as making saccades across continuously textured images or when an object under inspection is briefly occluded. In the latter case, neuronal adaptation would provide a robust encoding of the duration of occlusion.

Although it is still unclear how exactly neuronal adaptation could account for delay compensation, it has two important traits that are central to Nijhawan's proposal. First, neuronal adaptation is a sensory process that occurs only one synapse beyond the retina in cats and monkeys (Coenen & Eijkman 1972; Mayo & Sommer, submitted) and has been reported within the retina in lower species (Schwartz et al. 2007a). Thus, neuronal adaptation seems to be an immediate, automatic sensory mechanism. Second, neuronal adaptation is pervasive, seen at all stages of striate and extrastriate visual processing. This substantial concordance between adaptation and Nijhawan's proposal calls for additional research and computational modeling. An improved understanding of the neurophysiology of timing would provide an important complement to Nijhawan's thoughtful analysis.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

ACKNOWLEDGMENTS
This research was supported by the Alfred P. Sloan foundation and NIH R01-EY017592.

## Phase-alignment of delayed sensory signals by adaptive filters

doi: 10.1017/S0140525X08003956

Dennis J. McFarland
*Wadsworth Laboratories, New York State Department of Health, Empire State Plaza, Albany, NY 12201.*
**mcfarlan@wadsworth.org**

**Abstract:** Correction of sensory transmission delays is an intractable problem, since there is no absolute reference for calibration. Phase-alignment is a practical alternative solution and can be realized by adaptive filters that operate locally with simple error signals.

Nijhawan suggests that the visual system compensates for delays, with the perceived position of a moving object being based on a type of guess as to the object's actual position. This would be adaptive, as it is necessary that the output of the visual system be synchronized with the environment and behavior. However there is really no way that any part of the nervous system can be calibrated to correct for transmission delays, since there is no absolute reference available for calibration. As such, the calibration problem is intractable. This problem can be solved instead by aligning the relative phase of the signals in question. This is a common problem in engineering that has relatively simple and straightforward solutions.

Problems produced by transmission delays are frequently encountered in electrical engineering. Examples include cancellation of echoes produced in long-distance transmission lines and adaptive beam-forming from an array of sensors with delayed outputs (e.g., spatially separated sensors in a sonar system). These problems can be solved by means of adaptive filters with delayed inputs (Haykin 1996). The proper delay is selected by appropriate weighting on the delayed inputs. These weights are adjusted by some simple characteristic of the output, such as by minimizing the variance of the signal subject to the condition that the sum of the weights equals 1.

Adaptive filters capable of phase-aligning signals can be realized by simple operations that can be implemented either by algebraic equations or by a few elements of an artificial neural network. This means that the neural circuitry need not be complex. The cost function that tunes these filters can be very simple, so that it is not necessary to appeal to complex cognitive processes. Some characteristic of the fused signal, such as the variance, serves as an error signal to adjust the weights given to the delay elements. All that is required for phase-alignment is some sort of recursive feedback of the output of a perceptual subsystem that adjusts the connections of the delayed inputs. The simplest design is a tapped-delay line filter, although there is no unique solution to this problem.

The logical point for phase-alignment would be the step in processing that precedes the fusion of the to-be-aligned signals, as this would minimize the length of the path required for the recursive feedback and ensure the fidelity of the error signal. If filters are tuned by the output of areas far removed from the site of sensor fusion, the error signal will not closely reflect the characteristics of the fusion product. Hence, we would expect the site of phase-alignment to be the local network involving both the site of fusion and the immediate sources of its to-be-aligned inputs.

To the extent that the phase-alignment problem is solved with adaptive filters, we should expect that there would be no absolute delay correction employed by perceptual systems. Rather, the phase correction would depend on the problems encountered during the recent history of the various perceptual subsystems. An understanding of how this process works might best be obtained by observing the plasticity in the calibration of perceptual subsystems (i.e., adaptation to the statistics of the observer's environment). Fusion of audiovisual speech is an example.

Vatakis et al. (2007) measured judgments of the temporal order of auditory and visual speech tokens. Participants were concurrently exposed to audiovisual speech that was presented either in synchrony or with the auditory stream lagging by 300 msec. They found that exposure to asynchronous audiovisual speech shifted the point of subjective simultaneity in the temporal order task. These results demonstrate that even the brief exposure encountered in a typical experimental session is capable of recalibrating the relative phase of perceptual subsystems. Thus, phase-alignment in this case is an adaptive process.

Phase-alignment by adaptive filters can operate locally within a relatively simple network using simple error signals. As a result, the top-down input to perceptual systems does not need to involve complex cognitive processes from remote higher cortical regions. Furthermore, the input is in the form of adjustments to the strength of synaptic weights that tune perceptual filters. This adjustment occurs in a slow post-hoc manner so that changes resulting from any current mismatch do not affect on-line processing. This slow change would occur on a timescale much longer than that of individual perceptual events. This is in contrast to the view expressed by Nijhawan, where prediction is based on an interaction of visual systems with information from areas much farther downstream.

Prediction is a difficult problem that is not explained by simply appealing to top-down input to perceptual processing streams. Phase-alignment is a tractable problem with simple solutions. Prediction is based on a Cartesian view of the problem that considers the perception of motion to be the instantaneous value of a continuous variable that starts and stops a counter somewhere in the brain. Phase-alignment involves a filter that integrates input from sensors across time to produce an output, the nature of which is adjusted by prior sensory experience.

ACKNOWLEDGMENT
My thanks to Jonathan Carp for comments on an earlier version.

## Moving backward through perceptual compensation

doi: 10.1017/S0140525X08003968

Haluk Öğmen,[a] Saumil S. Patel,[b] Gopathy Purushothaman,[c] and Harold E. Bedell[d]
[a]*Department of Electrical and Computer Engineering, Center for Neuro-Engineering & Cognitive Science, University of Houston, Houston, TX 77204-4005;* [b]*Department of Neurobiology and Anatomy, University of Texas Medical School at Houston, Houston, TX 77030;* [c]*Department of Cell and Developmental Biology, Vanderbilt Vision Research Center, Vanderbilt University, T2304, Medical Center North, Nashville, TN 37232;* [d]*College of Optometry and Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-2020.*
**ogmen@uh.edu** **http://www.egr.uh.edu/ece/faculty/ogmen/**
**saumil@swbell.net**
**gopathy.purushothaman@vanderbilt.edu**
**Hbedell@optometry.uh.edu**

**Abstract:** In the target article Nijhawan speculates that visual perceptual mechanisms compensate for neural delays so that moving objects may be perceived closer to their physical locations. However, the vast majority of published psychophysical data are inconsistent with this speculation.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

The last decade has seen extensive debate on the interpretation of the flash-lag effect (FLE). Unfortunately, the target article fails to further scientific progress in this area. One problem is that many terms and concepts used in the article are not defined precisely. For example, the target article makes no clear distinction between "prediction" and "compensation" and claims that: "The notion that sensory processes per se may be predictive is not widely recognized. In fact, previous suggestions of sensory prediction (Nijhawan 1994) have led to controversy" (sect. 2, para. 4). That sensory processes may be predictive is neither new nor controversial. A multitude of mechanisms – from inhibition to light adaptation, occurring at neural loci from retina to cortex – have been discussed in the context of predictive processes and codes (e.g., Attneave 1954; Barlow 1961a; 1961b; 1961c; Hubel & Livingstone 1987; Grzywacz et al. 1995). What is controversial is the specific hypothesis proposed by Nijhawan, namely, compensation of neural delays at the perceptual level. A lack of conceptual clarity is exemplified also by the author's failure to distinguish between processes that contribute to compensation and the locus of compensation. The target article incorrectly attributes to us the suggestion that "visual processes do not contribute to compensation for neural delays" (sect. 6.1, para. 1). Our point was that the locus of compensation is primarily within the motor rather than the perceptual system (Purushothaman et al. 1998).

The target article is selective in addressing other published data and criticisms. For example, Murakami (2001) reported a FLE of approximately 60 msec for targets moving in random trajectories. Because the perceptual compensation model (PCM) proposed by Nijhawan relies on the predictability of motion trajectories, it cannot explain this outcome. Patel et al. (2000) and Öğmen et al. (2004) showed that the flash misalignment in a flash-lag paradigm depends on when the flash is presented relative to the onset of motion. If the flash was presented simultaneously with the onset of motion, a flash-lag was observed. If, using the same stimulus parameters, the flash was presented during continuous motion, a flash-lead was observed. Without ad hoc assumptions, this critical outcome cannot be explained by the PCM. Our review led to the conclusion that "given this large body of contradictory evidence, the lag-compensation through extrapolation model appears untenable" (Öğmen et al. 2004). We also showed that when the luminance of a flashed or a continuously moving target is changed, the magnitude of the FLE varies systematically in accordance with the predictions of the differential latency hypothesis (Kafalıgönül et al., in press; Ögmen et al. 2004; Patel et al. 2000; Purushothaman et al. 1998). This luminance-dependent variation in the flash-lag magnitude implies that the putative delay-compensation mechanism is unable to take into account stimulus-dependent variabilities in latencies.

Besides failing to address relevant published data and criticism, empirical data questioning the plausibility of the PCM based on substantial "compensation errors" are dismissed by stating that the model operates within an unspecified luminance range or that "over-extrapolation could be useful" (sect. 5.2.1.2, para. 3). The inability to specify after a decade of research even an approximate luminance range within which visual compensation works suggests that no such range exists. The Hess effect clearly demonstrates this point by showing that perceptual compensation for differential delays in targets of different luminance does not occur, within any range of target luminance (Williams & Lit 1983; Wilson & Anstis 1969).

Virtually all of the examples given to illustrate the need for a perceptual compensatory mechanism include motor action. The only unequivocal example offered for a perceptual spatial extrapolation of visual motion is the report by Maus and Nijhawan (2006). However, in order to demonstrate that perceptual spatial extrapolation occurs, one needs to show that the perceived trajectory of a moving object extends beyond its physical trajectory, that is, to locations where no retinal stimulation occurred. Maus and Nijhawan's findings can be explained by spatio-temporal integration along the motion path, which lowers the detection threshold for a target with a longer trajectory. Our analysis, shown in Figure 1, captures Maus and Nijhawan's findings quantitatively and illustrates that what they interpret as "extrapolation" is essentially a difference in sensitivity.

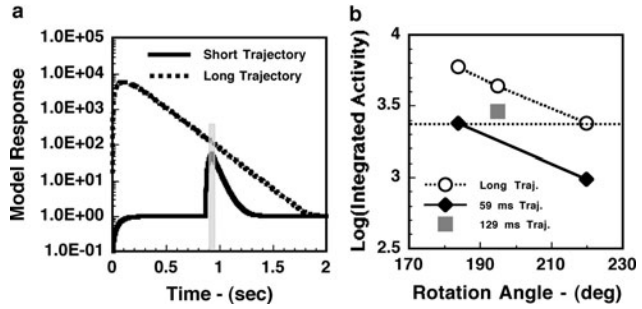



Figure 1 (Öğmen et al.). The results of simulating a unity gain, first-order low-pass temporal filter with stimuli used by Maus and Nijhawan (2006). (**a**) The dotted curve shows temporal filter (time constant = 63.5 msec) output for a stimulus that decreases smoothly in log luminance from a value of $10^4$ units during 360 degrees rotation. The solid curve represents the filter response to the same stimulus that moves only between angular positions 178 and 190 degrees. These simulate the *long trajectory* and *short trajectory conditions* in Maus and Nijhawan's Experiment 1. The gray vertical bar represents the temporal window within which the response is summed and compared for long and short trajectories, under the assumption that detectability increases with the log of the summed filter activity. (**b**) The logarithm of the summed output as a function of angular position during rotation, for the three conditions indicated in the inset. The activity produced at any angle of rotation is always less for short than long trajectory motion. Further, the integrated activity corresponding to short trajectories of motion at 184 and 195 degrees of rotation (59 and 129 msec motion durations) are very similar to that for the long trajectory of motion at 220 degrees rotation (horizontal dotted line). These three angular positions correspond to the positions reported by Maus and Nijhawan (2006) for the detection or disappearance of the moving spot, for the same durations of target motion.

Finally, the target article suggests that the proposed latency compensation mechanism can be viewed as producing a "lag constancy," analogous to a putative "sharpness constancy." It goes on to state that, in the absence of sharpness constancy, "the same moving object, traveling at different speeds, would appear more or less elongated (smeared)" (sect. 8, end of para.). Psychophysical data indicate that the length of perceived smear for moving targets varies systematically with speed (Bedell & Lott 1996; Chen et al. 1995). Because the target article provides no convincing evidence for visual compensation of latency, the appropriate analogy between these "constancies" is that neither has been shown to exist.

ACKNOWLEDGMENTS

We thank Srimant Tripathy and Ozgur Yilmaz for suggesting relevant literature.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

# Neurophysiology of compensation for time delays: Visual prediction is off track

doi: 10.1017/S0140525X0800397X

Gopathy Purushothaman,[a] Harold E. Bedell,[b] Haluk Öğmen,[c] and Saumil S. Patel[d]

[a]*Department of Cell and Developmental Biology, Vanderbilt Vision Research Center, Vanderbilt University, Nashville, TN 37232;* [b]*College of Optometry and Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-2020;* [c]*Department of Electrical and Computer Engineering, Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-4005;* [d]*Department of Neurobiology and Anatomy, University of Texas Medical School at Houston, Houston, TX-77030.*

**gopathy.purushothaman@vanderbilt.edu**
**Hbedell@optometry.uh.edu**
**ogmen@uh.edu**
**http://www.egr.uh.edu/ece/faculty/ogmen/saumil@swbell.net**

**Abstract:** Speculation by Nijhawan that visual perceptual mechanisms compensate for neural delays has no basis in the physiological properties of neurons known to be involved in motion perception and visuomotor control. Behavioral and physiological evidence is consistent with delay compensation mediated primarily by motor systems.

Prediction in skilled movements has been studied for more than 50 years (e.g., Gottsdanker 1952). Analysis of eye movements show that baseball and cricket batters and tennis players make an anticipatory saccade to the expected location of the ball and then pursue it foveally for a short period of time (Bahill & LaRitz 1984; Land & Furneaux 1997; Land & McLeod 2000). In conjunction with physiological evidence discussed below, these data support the view that compensation for neural latencies is based primarily on predictive motor plans (Kerzel & Gegenfurtner 2003; Öğmen et al. 2004; Patel et al. 2000; Purushothaman et al. 1998). Perceptual compensation in early visual areas, as suggested by Nijhawan in the target article, need not occur. Although Nijhawan links neurophysiological findings to "perceptual compensation," we show here that these links are misconstrued.

As Öğmen et al. point out in their commentary in this *BBS* issue, Nijhawan fails to distinguish between "prediction" and "compensation." As a consequence, he mistakes predictive phenomena, such as perisaccadic suppression of retinal motion and remapping of receptive fields (RFs), as evidence for "compensation of neural delays." For example, the author argues that the perisaccadic remapping of RFs in the lateral intraparietal (LIP) area (Duhamel et al. 1992) implies that "sensory processes participate in prediction" (target article, sect. 2, para. 4) and advances a model for perceptual compensation "based on the interaction of signals between the magnocellular and parvocellular systems" (sect. 5.2.1) within area V1 (see Fig. 6 of the target article). However, impending motor action is a prerequisite for the RF remapping reported by Duhamel et al (1992), which indicates that this phenomenon is linked to motor preparation and action and does not imply "compensation of neural delays" at the perceptual level. Indeed, RF remapping was shown not to occur in area V1 (Nakamura & Colby 2002). As discussed by Öğmen et al (this issue), the only example in the target article of purely perceptual extrapolation is from Maus and Nijhawan (2006), which is accounted for by differences in sensitivity.[1]

Nijhawan attempts to link the flash-lag effect (FLE) to the physiological findings of Berry et al. (1999), but the FLE increases linearly with the velocity of the moving object (Krekelberg & Lappe 1999; Nijhawan 1994); and Berry et al. showed that the neural response to a moving stimulus shifts increasingly in the opposite direction of motion as the velocity of the moving stimulus is increased. Whether compensation occurs at the perceptual or motor levels can be assessed by considering the dynamic response properties of neurons in areas associated with perception and motor control. Smooth pursuit eye movements exhibit approximately zero phase lag in response to predictable low-frequency target motion (see Carpenter 1988, pp. 58–63). However, in the absence of timing and direction cues, a pursuit latency of 100–150 msec occurs at motion onset, and this lag gradually approaches zero as the pursuit system evolves from its transient to steady-state mode of operation (Bahill & McDonald 1983; Barnes et al. 2000; Fukushima et al. 2002). In primates, the middle temporal (MT) area is closely associated with motion perception (see Britten 2004; Britten et al. 1996; Liu & Newsome 2005; Purushothaman & Bradley 2005), and the frontal eye field (FEF), a premotor area, is involved in generating eye movements (see Bruce et al. 2004; Schall 2004). Lesions to MT impair motion perception (Newsome & Paré 1988), whereas lesions to parts of FEF impair predictive smooth pursuit movements (Keating 1991; MacAvoy et al. 1991). Even with bilateral FEF lesions, monkeys can still detect target motion and attempt to track it, albeit with very low gain and high phase lag (Keating 1991). In addition, the stimulus-driven component of FEF neural responses is modulated significantly by the relevance of the stimulus for impending visuomotor behavior, and modulated little by the physical features of the stimulus (cf. Schall 2004). Finally, electrical stimulation of MT neurons biases perceived movement direction without causing eye movements (Nichols & Newsome 2002; Salzman et al. 1990; 1992), whereas stimulation of FEF neurons directly induces eye movements (Gottlieb et al. 1993; 1994). Thus, the activities of FEF neurons are closely associated with the generation of eye movements and are relatively unrelated to perceptual qualities of the tracking target.

MT neurons show response lags on the order of 100 msec for step-ramp smooth pursuit tasks (Komatsu & Wurtz 1988; Newsome et al. 1988). Despite complete compensation exhibited by smooth pursuit eye movements for predictable stimuli, these response lags do not diminish for similar stimuli (Lisberger & Movshon 1999), ruling out delay-compensation at or before this cortical site linked to motion perception. In the FEF, neurons active during smooth pursuit (pursuit neurons) veridically track predictable target motion (Gottlieb et al. 1993; 1994; MacAvoy et al. 1991; Tanaka & Fukushima 1998). However, no FLE occurs during accurate pursuit (Nijhawan 2001). Therefore, the linkage of the FLE to lag-compensation requires the analysis of conditions in which no smooth pursuit occurs. In the absence of active pursuit, pursuit neurons in FEF respond very little to target motion, predictable or otherwise (Gottlieb et al. 1993; 1994; MacAvoy et al. 1991; Tanaka & Fukushima 1998). Pursuit neurons in FEF that do respond to moving targets in the absence of active pursuit show "a clear phase lag due to visual latencies" (Fukushima et al. 2002), just like upstream MT and MST neurons. Unlike the motion-sensitive neurons in MT, this lag decreases after approximately one cycle of predictable motion (Fukushima et al. 2002).

In summary, the known physiology of motion perception and motor control is in close agreement with our proposal that the locus of compensation for neural time delays is primarily within the motor rather than the perceptual system (Öğmen et al. 2004; Patel et al. 2000; Purushothaman et al. 1998).

ACKNOWLEDGMENT

We thank Pierre Pouget for his suggestions and comments.

NOTE

**1.** The target article also attempts to link the FLE to other phenomena in which moving objects are perceived to shift in the direction of motion (e.g., Chung et al. 2007; DeValois & DeValois 1991; Ramachandran & Anstis 1990). However, these phenomena differ distinctly from the FLE in their dependence on velocity and other stimulus parameters, and are unlikely to share a common mechanism.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

## Single mechanism, divergent effects; multiple mechanisms, convergent effect

doi: 10.1017/S0140525X08003981

Bhavin R. Sheth[a,b] and Daw-An Wu[c]

[a]*Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204;* [b]*Center for NeuroEngineering and Cognitive Systems, University of Houston, Houston, TX 77204;* [c]*Department of Psychology, Harvard University, Cambridge, MA 02138.*

**brsheth@uh.edu** **daw-an@wjh.harvard.edu**

**Abstract:** It is commonplace for a single physiological mechanism to seed multiple phenomena, and for multiple mechanisms to contribute to a single phenomenon. We propose that the flash-lag effect should not be considered a phenomenon with a single cause. Instead, its various aspects arise from the convergence of a number of different mechanisms proposed in the literature. We further give an example of how a neuron's generic spatio-temporal response profile can form a physiological basis not only of "prediction," but also of many of the other proposed flash-lag mechanisms, thus recapitulating a spectrum of flash-lag phenomena. Finally, in agreeing that such basic predictive mechanisms are present throughout the brain, we argue that motor prediction contributes more to biological fitness than visual prediction.

It is likely that multiple mechanisms combine to create the flash-lag phenomenon: persistence, priming, backward masking, temporal dilation, and even attention have all been demonstrated in one study or another (Bachmann & Poder 2001; Baldo & Namba 2002; Kanai et al. 2004; Krekelberg & Lappe 2001; Namba & Baldo 2004; Sheth et al. 2000). It seems that cleverly designed experiments can prove the importance of one's favored model, but in vanishingly small parameter regimes. For example, experiments on the flash-terminated condition support extrapolation, but the results are limited to degraded, uncertain stimuli (Fu et al. 2004; Kanai et al. 2004). Other experiments support differential latency, but these use stimuli of much lower luminance (Patel et al. 2000; Purushothaman et al. 1998).

We have previously argued that a very basic consideration of neuronal response profiles can recapitulate a wide array of flash-lag related mechanisms and effects (Kanai et al. 2004). As a stimulus moves in physical space, it maps out a topographically corresponding path in cortical space. At a given time instant, there are the following components: (A) cells at the "current" location of the stimulus are the most active; (B) cells in the immediate past path of the motion contain residual activity; (C) cells in the distant past path contain below-baseline activity caused by adaptation and intracortical inhibition; and (D) cells in the family of future motion paths have above-baseline subthreshold activity through intracortical excitation. This pattern of activity arises from the basic temporal response profile of a single neuron to input, and from the fact that lateral connections between neighboring neurons tend to cause net excitation to weakly firing neurons and net inhibition to strongly firing neurons (Henry et al. 1978; Levitt & Lund 1997; Somers et al. 1998; Stemmler et al. 1995). These four components of the spatiotemporal response profile have strengths that depend not only on factors intrinsic to the neuronal network, but also on stimulus parameters such as luminance, speed, and so on.

These components can implement various mechanisms related to flash lag and motion processing. Component D could be descriptively labeled as priming, and if the activity in D is high enough to shift the centroid of the activity distribution forward, it could partially underlie a motion extrapolation mechanism. C could be a critical part of the neural basis for motion deblurring. When component B is prominent, differential latency for motion and flash arises: The spatiotemporal integral of the activity of A+B will reach perceptual threshold faster than a temporal integral of a stationary flash. Finally, stimulus conditions such as uncertainty will determine whether the activity in A alone suffices for awareness, or whether B needs to be added; this is a plausible neural basis for two different Bayesian estimators – conditional mean and maximum likelihood. Thus, the tuning of a simple neural mechanism can give rise to myriad psychophysical phenomena and high-level models.

When distilled down to the idea of lateral propagation of cortical activity, we agree that prediction is intuitive and should be neurally omnipresent. The above properties of neurons are generic and found in almost all networks – sensory and motor. One question that arises then is: What is the relative contribution of sensory and motor prediction to successful behavior?

We argue that prediction in the motor realm seems to be more effective and useful. First, visual prediction is applicable if a target moves with uniform velocity, but motion is hardly ever uniform in real life – physical (friction) and internal (attention, interest) factors often disrupt the smooth flow of motion. Second, motor prediction does not need to be as accurate as visual prediction. The agent can often over-compensate for the movements of the target, thus arriving at a common intersection point some time before the target. This allows the agent some slop, and with it, the flexibility to compensate for change in target speed, and for relatively small synaptic delays within its own nervous system. All delays – visual, synaptic, and of the muscle or tool-based effector – are available in a lump sum and are undifferentiated to the motor system as motor error. Motor systems routinely compensate for delays of the order of seconds, which arise from slow effectors. Such a system should be well-equipped to accommodate 100 msec of visual synaptic delay. Thus, the motor system seems to be the workhorse. Although this is but an isolated example, we note that prism adaptation begins in the motor system; one's motor system compensates for errors weeks before one begins to correctly perceive the world.

Visual prediction at the neural level is then just one of many important mechanisms in two senses: it is only one of the mechanisms which contribute to the flash-lag effect, and it is only one of the types of "neural prediction" which contribute to our biological fitness. In the case of flash-lag, variations in stimulus conditions can dictate the relative importance of visual prediction. In the case of biological fitness, it seems that visual prediction is just a small jumpstart – a small, subthreshold benefit to the organism in comparison to other predictive brain mechanisms.

## The mechanisms responsible for the flash-lag effect cannot provide the motor prediction that we need in daily life

doi: 10.1017/S0140525X08003993

Jeroen B. J. Smeets and Eli Brenner

*Research Institute MOVE, Faculty of Human Movement Sciences, VU University, NL-1081 BT Amsterdam, The Netherlands.*

**j.smeets@fbw.vu.nl** **e.brenner@fbw.vu.nl**

**http://www.fbw.vu.nl/~JSmeets/**

**Abstract:** The visual prediction that Nijhawan proposes cannot explain why the flash-lag effect depends on what happens after the flash. Moreover, using a visual prediction based on retinal image motion to compensate for neuronal time delays will seldom be of any use for motor control, because one normally pursues objects with which one intends to interact with ones eyes.

In his target article, Nijhawan proposes that early visual processing provides the prediction that is needed to deal with sensorymotor delays when we interact with moving objects, rather than such prediction arising from complex motor strategies as is generally assumed. He argues that the flash-lag effect and related phenomena illustrate the visual basis of such prediction. In his discussion of the extensive literature on this topic, he ignores several findings that show that the flash-lag effect cannot be



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

caused by a visual prediction based on the preceding object motion.

Several experiments have been performed in which a target moves both before and after the flash, but changes its speed or direction of motion at an unpredictable moment around the time of the flash. According to Nijhawan's account of visual prediction, the target's motion after the flash should be irrelevant for its perceived position at the time of the flash. However, the perceived position has been shown to depend on the target's motion up to 80 msec after the flash (Brenner & Smeets 2000; Eagleman & Sejnowski 2000; Whitney & Murakami 1998). This result is inconsistent with any kind of motion extrapolation. It is also unlikely that it is primarily caused by neuronal signals pertaining to the flashed target taking longer to reach the brain than ones pertaining to the moving target (Whitney & Murakami 1998), because the flash-lag effect can be demonstrated with a very bright flash and a dimly lit moving object.

The dependence of the flash-lag effect on what happens after the flash can readily be explained if one regards perception as an active process (O'Regan & Noe 2001). If so, the location of the moving object is not evaluated continuously. It is only evaluated when one needs to know it. The flash indicates that this is the case. As determining the position in response to the flash takes time, the result is a judged position that the object only reaches some time after the flash. The fact that the moving object is perceived ahead of its location at the time of the flash is therefore not due to extrapolation, but to sampling its position too late. This implies that the flash-lag effect should decrease if one can convince subjects to start evaluating the location of the moving object before the flash is registered. A way to achieve this earlier sampling is by making the moment of interest more predictable. Indeed, the flash-lag effect is reduced (and even absent in some subjects) under such conditions (Brenner & Smeets 2000).

Besides the doubts about the role of visual prediction in the flash-lag phenomenon, there is also a more fundamental problem with the main claim of the target article. Nijhawan's interesting claim is that visual prediction provides the prediction needed to compensate for neuronal delays when interacting with moving objects. However, when trying to intercept a moving target, subjects tend to pursue the target with their eyes. This is so not only in laboratory conditions (Mrotek & Soechting 2007) but also, for instance, during the final approach phase when hitting a ball in cricket (Land & McLeod 2000). Moreover, subjects are better at an interception task when they pursue the target with their eyes than when they fixate somewhere near the point of interception (Brenner & Smeets 2007). One reason for pursuing the target is that pursuit eliminates the blur caused by retinal motion, leading to more precise vision. However, the lack of retinal motion means that the predictive mechanism proposed in the target article will not be working. Therefore, in the situations in which prediction is needed most in daily life, the proposed mechanism cannot contribute to such prediction.

The way in which subjects pursue moving targets can give us insight into how prediction works. It is known that pseudo-random smooth target motion is pursued with delays of more than 200 msec (Collewijn & Tamminga 1984; Koken & Erkelens 1992). Targets moving at a constant – and therefore predictable – speed are pursued with a negligible delay (Barnes & Asselman 1991). If this reduction in visuomotor delay were caused by the kind of visual prediction proposed in the target article, it would only work when the target motion is constant. This is not the case: Negligible delays are also found when the target motion is predictable, but not on the basis of the directly preceding visual information (Thier & Ilg 2005). For instance, humans can pursue sinusoidal motion with minimal delays. It only takes about half a cycle of the sinusoidal target motion to achieve the minimal tracking delay. If the target disappears, or changes its motion, the sinusoidal eye movement continues for about half a cycle (van den Berg 1988). Additional evidence against the proposed visual prediction is that the prediction in pursuit is task-specific.

When following a target with their eyes, subjects make errors in the smooth pursuit that are corrected by catch-up saccades that are predictive: They compensate for the errors that develop during their programming and execution. These catch-up saccades could be based on a visual prediction, or on a motor prediction specific to the pursuit. In the former case, the errors in pursuing a smoothly moving target should also be compensated for when making a saccade in response to a sudden jump of the target. However, in such an experiment, the saccade amplitude is matched to the target jump (Smeets & Bekkering 2000). So the prediction that subjects make in order to be able to track the moving target is specific to pursuit.

Our conclusion is that if the low-level predictive mechanisms proposed by Nijhawan exist, they are responsible neither for the flash lag effect nor for the motion extrapolation in our interaction with moving objects.

## Anticipating synchronization as an alternative to the internal model

doi: 10.1017/S0140525X08004007

Nigel Stepp and Michael T. Turvey

*Center for the Ecological Study of Perception and Action, University of Connecticut, Storrs, CT 06269-1020; and Haskins Laboratories, New Haven, CT 06511.*

**nigel.stepp@uconn.edu** **michael.turvey@uconn.edu**
**http://ione.psy/uconn.edu/mturvey/**

**Abstract:** The fundamental assumption of compensation for visual delays states that, since delays are dealt with, there must be compensatory mechanisms. These mechanisms are taken to be internal models. Alternatives for delay compensation exist, suggesting that this assumption may not be fundamental, and nor should the existence of internal models be assumed. Delays may even be employed in their own compensation.

A case is made for the ubiquity of anticipatory behavior by the sensory-motor system, including visual perception. There is no question that anticipation is ubiquitous; however, that ubiquity tends to prompt assumptions about the world which may not be warranted. For instance, "*the fundamental assumption of compensation for visual delays*, [. . .] which states: 'In the absence of mechanisms compensating for visual delays, many behaviors in otherwise healthy animals would be disrupted.'"(target article, sect. 6.1, para. 2; italics in original).

The nature of these compensating mechanisms is of particular importance. If a tacit assumption which goes along with "the fundamental assumption" is that those compensation mechanisms are forward models, then the assumption may not be fundamental. The existence of alternative explanations for anticipatory behavior strongly suggests that the assumption should not be regarded as axiomatic. Moreover, the term *compensation* connotes both the presence of error and active correction. The term *anticipation* will be used to avoid these implications.

Here, we present an alternative paradigm for anticipatory systems, which may be useful when considering the anticipation for visual delays. This alternative comes in the form of anticipating synchronization (Ciszak et al. 2004; Voss 2000). In one of its simplest manifestations, anticipating synchronization is a coupling of two dynamical systems such that the slave system is able to synchronize with the future of the master system. The general form of this coupling is shown in the following system of equations:

$$\dot{x} = f(x(t))$$
$$\dot{y} = f(y(t)) + kx(t) - y(t - \tau))$$



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

When two dynamical systems are coupled in this way, values of $y$ synchronize (after a transient) with the future of $x$. Put another way, $y$ becomes very close to a phase-shifted copy of $x$. This method of anticipating synchronization is an instance of the theoretical notion of strong anticipation (Dubois 2003; Stepp & Turvey 2007). As opposed to weak anticipation, which involves the use of models for explicit prediction, strong anticipation utilizes law-based facts together with ordinary intrinsic dynamics. In other words, the system above is anticipatory because it has no opportunity not to be.

The delay term $y(t - \tau)$, representing delayed feedback in the slave, should attract special attention. It is notable that without this delayed feedback, the system above would not exhibit anticipatory behavior. Typically, delays are taken to be a liability, for which there must be explicit compensation. The alternative perspective suggests that delays are an asset, which can be the source of anticipatory behavior. In the succeeding paragraphs, one may see how delay-coupling addresses the features of behavior generally taken to imply a forward model.

Returning to the "fundamental assumption" itself, consider the fact that increased delay induces a breakdown of anticipatory ability, for example, demyelination in multiple sclerosis patients. This fact is cited as the starting point for the assumption of compensating mechanisms. Increasing the $\tau$ parameter of a delay-coupled system is known to show this same feature. Figure 1 shows the $(k,\tau)$ parameter space associated with a delay-coupled system such, where the function $f$ is the Rössler system. There is a sharp drop (shift from white to black) in anticipatory ability on increasing$\tau$. Thus, anticipatory ability is lost after $\tau$ is increased past a critical value. The effect does not require an internal model.

Nijhawan makes the important point that mysterious self-anticipation applies to the entire sensory-motor system. For a delay-coupled system showing anticipating synchronization, this behavior is expected by definition. Considering again the system of equations above, the delay is inherent in the slave system, $y$ – the same system which is anticipating its master, $x$. Therefore, $y$ is indeed anticipating its own delay. Furthermore, this anticipation is made possible not through an internal model, but falls out of the normal functioning of the correctly organized system.

Knowing something about delay-coupled systems and anticipating synchronization, we may turn to the flash-lag effect. The pertinent aspect of this effect is that continuously moving stimuli are led by the perceiver, whereas flashes are not. As noted by Nijhawan, this effect shows itself in animal behavior as a virtual lack of delay while tracking a moving object. Viewing through the lens of anticipating synchronization, this also comes as no surprise. In fact, one might be surprised if this effect were not the case. Recall, anticipating synchronization, like all dynamical synchronizations, goes through an initial transient. A flash, because of its short duration, does not allow this transient to take place. Alternately, in the case of continuous stimuli, we expect that the transient has taken place, allowing for anticipation.

To reiterate, the evidence presented for the existence of anticipation of visual delays is sound. However, the existence of anticipating synchronization in particular, and strong anticipation in general, restricts assumptions that can be made about the nature of that anticipation. What is seen as "compensation" by an internal forward model may simply fall out of the ordinary dynamics of the components in question because of the way in which they are organized with respect to their stimulus. The above may be taken both as a counterexample to the necessity of compensatory forward models and as an alternative framework for investigating anticipatory behavior.

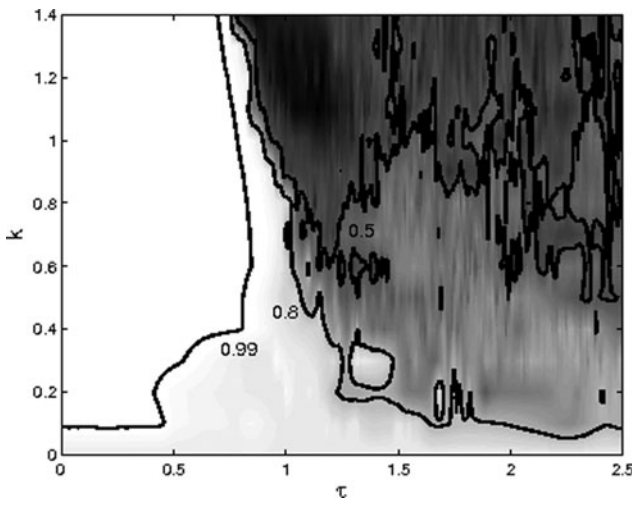



Figure 1 (Stepp & Turvey). For many $(k,\tau)$ combinations, the slave system $(y)$ was incrementally shifted relative to the master system $(x)$, and maximum correlations recorded. White corresponds to a maximum correlation of 1.0; black corresponds to 0.0.

## Empirically testable models are needed for understanding visual prediction

doi: 10.1017/S0140525X08004019

Giuseppe Trautteur,[a] Edoardo Datteri,[b] and Matteo Santoro[c]

[a]*Dipartimento di Scienze Fisiche, Università di Napoli Federico II, Complesso Universitario Monte Sant'Angelo, NA 80126 Napoli, Italy;* [b]*Dipartimento di Scienze Umane per la Formazione "Riccardo Massa", Università degli Studi di Milano-Bicocca, MI 20126 Milano, Italy;* [c]*DISI, Universita' degli Studi di Genova, GE 16146 Genova, Italy.*

**trau@na.infn.it**
**edoardo.datteri@unimib.it**
**http://ethicbots.na.infn.it/datteri/**
**matteo.santoro@disi.unige.it**

**Abstract:** Nijhawan argues convincingly that predictive mechanisms are pervasive in the central nervous system (CNS). However, scientific understanding of visual prediction requires one to formulate empirically testable neurophysiological models. The author's suggestions in this direction are to be evaluated on the basis of more realistic experimental methodologies and more plausible assumptions on the hierarchical character of the human visual cortex.

The target article succeeds in convincing the reader that prediction mechanisms are pervasive in the central nervous system (CNS), and that prediction is involved even in the early stages of sensory processing. Indeed, these results are not surprising for informed readers, as discussions on this topic have been prevalent for decades (Butz et al. 2007). Nonetheless, after the neat survey by Nijhawan, it is likely – and indeed, highly desirable – that most of the controversies about the predictive character of visual processing will cease.

At the same time, we feel that the wealth of psychophysical results discussed by the author – which may constitute a strong basis for predicting behavioral performances in this kind of tasks – fails to further significantly our understanding of the phenomenon of "perception" in the brain: Without a suitable proposal about how anticipation is implemented in the CNS, we can hardly assume we have won one more point in this direction. One must acknowledge that the target article argues forcefully for the existence of predictive capabilities in visual processing. However, what we need now in order to explain the behaviors observed in psychophysics experiments is to shift from the level of psychophysical analysis to the level of neuroscientific modeling, and analyze the biological mechanisms



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

responsible for these behaviors. "Understanding visual prediction" (target article, Abstract) requires identifying parts of the target nervous system, studying their behavior, their interconnections, and their functional role with respect to the visual processing capacity under investigation (Craver and Darden 2001; Cummins 1983; Glennan 2005). In the lack of sufficiently precise theoretical models, we can hardly evaluate the biological plausibility of some of the author's claims, especially the highly speculative claim, discussed in the final sections of the article, that "the goal of any type of compensation is to produce coordination between the time-varying neural representations that exist in parts of the CNS separated by significant delays" (sect. 9.1, para. 1). Similarly, in the lack of theoretical models, we can hardly appreciate the relationship between the hypotheses proposed by the author and other models available in the literature: Notably, what is the distinctive character of the "internal model for visual prediction (IMVP)" hypothesis discussed in the target article (see sect. 7.2) with respect to the internal-model-based hypothesis discussed by Wolpert et al. (1995; quoted in the article) which involves sensory anticipation as well?

This is not to deny that conjectures on the involvement of some neural structures in anticipation-based capacities may play important roles in the discovery of anticipation mechanisms. As has been pointed out (Machamer et al. 2000), the formulation of mechanistic models of biological behaviors often involves the formulation of mechanism sketches, which provide partial or underspecified accounts of the target mechanism. Often one starts with the isolation of some parts of the target system which are supposed to play key functional roles with respect to the capacity under investigation, and proceeds by identifying, in a stepwise fashion, other functional components (and their material instantiation) which are needed to obtain a "working" mechanism. One must acknowledge that Nijhawan, albeit being chiefly concerned with attesting the "existence" of visual prediction, does occasionally formulate sketches of the neural mechanisms that could produce the behaviors under investigation.

However, we surmise that some of the author's sketches are based on too simplistic representations of the architecture of the visual cortex, which undermines the possibility of extending the proposed hypothesis to the study of other predictive phenomena in the CNS. In fact, the article focuses on psychophysical experiments in which only simple stimuli are involved. If we consider more complex classes of stimuli, then the strictly retinotopic organization of the visual cortex – which plays a key role in the model – is no longer sustainable. Instead, according to a wealth of physiological data reported in Serre et al. (2005, and references therein), the main feature of the cortical architecture is a hierarchical organization, which aims (in a series of stages) to increase the invariance to object transformations and to tune to more specific features. It is acknowledged that one of the main functions of the ventral stream pathway is to achieve an exquisite trade-off between selectivity and invariance, which results in an improved capability to accomplish readily many recognition tasks. As a possible side effect of such a trade-off, it is very likely that strict information about the position of the objects becomes less significant (in the higher levels of the cortical hierarchy). If we assume that predictive phenomena also play a role in this context – and we do – then it is not clear how extrapolation can be of use. In our opinion, different phenomena should be considered, such as, for example, the effect of the temporal and spatial integration of the neural signal from the lower layers of the hierarchy to the higher ones. That could have induced a kind of "characteristic time" for each layer insofar as different classes of objects would be affected by different amounts of delays, and consequently, distinct anticipation mechanisms.

In conclusion, we feel the author argues convincingly that predictive mechanisms must be implemented in the CNS at various levels. This claim is consistent with a wealth of literature on sensory processing and sensory-motor coordination. However, as far as scientific understanding of these phenomena is concerned, empirically testable theoretical models of visual prediction, at a neurophysiological level, are needed. In our opinion, the author's suggestions in this direction are to be evaluated on the basis of more realistic experimental methodologies and by taking into account the hierarchical character of the human visual cortex.

## Motion as a reference for positions

doi: 10.1017/S0140525X08004020

Wim van de Grind

*Department of Functional Neurobiology, Utrecht University, 3584 CH Utrecht, The Netherlands.*

**w.a.vandegrind@uu.nl**

**Abstract:** Is the position of a moving target "predicted"? I argue that we should regard moving targets as the natural (veridical) position references. Motion is probably perceptually absolute, whereas position and time are relative quantities, as in physics. According to this view, processing delays are incorporated in the abstract local signs of motion signals. The flash-lag effect is one case in point.

The usual way of looking at the flash-lag effect (FLE) is to (implicitly) assume that the brain contains fixed position-maps (e.g., retina-centric, head-centric, etc.) and that incoming retinal signals need to be assigned a place on these maps. Moving objects are in the habit of changing their position while your brain is still trying to pinpoint them on a map. Therefore, it would be a smart move if the brain were to place them on the map at the position they will have reached at pinpointing time. Nijhawan calls this *visual prediction*. It is relatively easy to envision such a visual prediction for the whereabouts of moving targets. Motion sensors report velocity vectors, and the corresponding position shifts in the mapping process can be (re-)calibrated during the visuomotor experience. It would be smart to do this in the front end of the visual system, so that both the action system and the perception system can profit from the shifted mapping. Lankheet and I have developed a simple implementation of this idea (van de Grind 2006, Fig. 1; cf. Lankheet & van de Grind, in press) in which motion stop signals can curtail the mapping shift. Motion prediction as well as start and stop signals are already available at the retinal level (Berry et al. 1999; Schwartz et al. 2007b). Further refinement is possible in motion sensors in V1 and middle temporal area (MT) and in the subsequent mapping on a head-centric position map. Cantor and Schor (2007) presented arguments and evidence that further support such low-level explanations. Echoes of these ideas can be found in the target article, but in the end Nijhawan prefers an exclusively high-level model, including feedback from the visuomotor action-system to the perceptual system. Here I want to sketch a rather more grandiose low- to middle-level alternative for the prediction hypothesis.

A neurophysiologist can see how a position in the environment or on the retina is mapped onto a position in the brain of a studied subject, but such a comparison is not possible for the subject. Without a homunculus, we cannot read physical map positions in our own brain, so external positions have to be more abstractly coded in terms of processable neural activity. It has been shown by Koenderink (1984a; 1984b) that it is possible to neurally construct such an abstract "local sign" system from scratch, by analyzing incoming information with simple operators. Receptive fields and brain maps set up correlation constraints because of overlap and inclusion relations, and local signs are embedded in the correlation-structure of activity patterns. This requires interaction with the environment to (re-)calibrate local sign information in the brain. Similar



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

remarks hold for the time sense (van de Grind 2002; 2006). Static coordinates are not suitable as references for the construction and maintenance of such a local sign system. In natural vision, everything always flows (or jumps) on the retina. Therefore, motion information is the best anchor for action control, setting the momentary coordinates. Let us hypothesize that motion is used to build up and continuously recalibrate the local sign system. This makes motion an "absolute" quantity, and position and time relative quantities, as in physics. For the FLE that would mean that the position map is always tuned to motion, so that moving objects are normally seen veridically at their instantaneous physical position, regardless of their processing time. Moving objects thus form the position reference for all non-moving objects. The position of the flash in the FLE is therefore correctly seen to lag the position of the moving reference – a shift that for logical reasons appears to equal the speed times latency difference. The flash localization in the FLE is correct, not an illusion, because at the time that we can report the flash position, that position indeed lags the motion reference position, in terms of momentary physical coordinates.

Interestingly, this was the implicit assumption in some of the earliest papers on the flash-lag effect, by Hazelhoff and Wiersma (1924; 1925), based on Hazelhoff's Ph.D. thesis of 1923. Hazelhoff wanted to measure the perception time of a flash; that is, how long after a brief flash do we consciously perceive it? He invented a new method of measuring this latency. His subjects had to pursue a black vertical bar moving along a horizontal scale with a speed of 23 degrees per second. When the bar passed point 0, a white line was flashed (10–20 msec), centered on the slightly wider black bar, and the subject had to report the bar's scale-position at the moment the flash was seen. Hazelhoff calculated the presumed perception times for a range of flash luminances and for three subjects. The correspondence to modern findings is striking. Hazelhoff reasoned that the pursued object is always seen at the fovea and is "on" all the time. The flash-signal, on the other hand, reaches some position comparator with a head-centric position label that was valid one perception-time back for the moving eye. Metzger (1932) showed that pursuit of the moving target is not essential in the Hazelhoff effect. If you fixate and the scale moves continuously, a flash will be shifted the same amount. Metzger criticized Hazelhoff's interpretation by pointing out that motion processing also has a latency, and came up with the latency-difference explanation that is still popular today. My aforementioned hypothesis, that motion provides the reference signal for position and time, means that I think that Metzger was wrong. Of course motion has a processing time, but that can simply be incorporated in the corresponding local sign. Now that study of the FLE has grown into a field with many stake-holders, subtle experimental paradigms, grossly incompatible explanations, and dozens of papers, it may be time to turn to the underlying question: What is the reference for positions in visual processing? My guess is: motion.

## Perception-action as reciprocal, continuous, and prospective

doi: 10.1017/S0140525X08004032

Jeffrey B. Wagman

*Department of Psychology, Illinois State University, Normal, IL 61790.*
**jbwagma@ilstu.edu**
**http://www.psychology.ilstu.edu/jbwagma/**

**Abstract:** From the perspective of ecological psychology, perception and action are not separate, linear, and mechanistic processes that refer to the immediate present. Rather, they are reciprocal and continuous and refer to the impending future. Therefore, from the perspective of ecological psychology, delays in perception and action are impossible, and delay compensation mechanisms are unnecessary.

Nijhawan argues that delays inherent in the transmission of information introduce error into perceptual and motor processes. Such delays are primarily a consequence of the time required to transmit information from the world to the central nervous system (perception) and from the central nervous system to the world (action). These delays create error between the actual state of the world and the state of the world perceived (and acted on) by the nervous system. As a result, delay compensation mechanisms are required to correct this error, and such mechanisms are a pervasive feature of the nervous system (particularly the visual and motor planning systems). I propose that delay compensation mechanisms are necessary only if perception and action are (mis)characterized (a) as fundamentally separate linear and mechanistic processes, and (b) as processes that (primarily) occur with reference to the immediate present. From the perspective of ecological psychology (J. Gibson 1979; Reed 1996), both (a) and (b) characterizations are rejected. Thus, from this perspective, the proposed delays do not exist, and delay compensation mechanisms are unnecessary.

***Perception-action as reciprocal and continuous.*** The delays described by Nijhawan reflect those that are inherent in the receiving and sending of information. As a result, such delays (and the need for delay compensation mechanisms) are dependent upon characterizations of perception and action as processes in which information is received and sent, respectively. That is, they are dependent upon a characterization of perception and action as information processing. From such a perspective, perception and action are inherently separate processes in which information is transmitted in a sequence of linear and mechanistic steps. Therefore, perception and action are subject to the time delays inherent to such processes.

From the perspective of ecological psychology, far from being separate, linear, and mechanistic, perception and action are reciprocal and continuous. Perception and action are circularly causal in that perceiving refers to acting, and acting refers to perceiving (Shaw & Wagman 2001). It is inappropriate to label perception as "input" and action as "output," because perception and action are multidirectional and mutually constraining. Perception is action, in that exploratory behaviors (e.g., head movements or manipulation of a handheld object) reveal properties of the world that would otherwise remain hidden. To paraphrase James Gibson (1979, p. 223), "As much as we must perceive in order to move, we must also move in order to perceive.

To ecological psychologists, perception-action is a continuous event with no clear beginning or end. The relationship between perception and action is more like the continuum that exists between the sides of a Möbius band than like the relationships among the parts of a clock (Turvey 2004). For this reason, ecological psychologists refer to the relationship between perceiving and acting as "perception-action" much as physicists refer to the relationship between space and time as "space-time." If perception-action is characterized in this way, then delays between perception and action are impossible because perception-action is a continuous event (Turvey 2004). No transmission of information is necessary (or even possible), and delays between perception and action do not (and cannot) occur. Therefore, delay compensation mechanisms are unnecessary.

***Perception-action as a prospective act.*** According to Nijhawan, the delays created by the transmission of information in the nervous system are problematic because they produce an error between the "instantaneous registered state" of the world and the actual state of the world. As a result, perception necessarily "lags behind" the world and must be "brought up to speed" by the pervasive delay compensation mechanisms in the nervous system. Such delays seem particularly problematic in situations in which a behavior requires prediction of a future event (such as a batsman attempting to hit a ball thrown by a bowler). However,



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

in addition to being dependent on a characterization of perception as information processing, such delays (and the need for delay compensation mechanisms) are also dependent upon a characterization of perception and action as processes that (primarily) occur with reference to the immediate (i.e., instantaneous) present.

From the perspective of ecological psychology, perception-action primarily occurs with reference to the (impending) future (Wagman & Malek, in press). In order to successfully achieve a behavioral goal (e.g., reaching for a cup of coffee or hitting a thrown ball), perceiver-actors must be able to perceive whether that (future) behavior is possible, and (if so) they must be able to perceive how to control their (future) movements such that this possibility is realized (Shaw & Turvey 1999). Thus, perception-action is inherently a prospective act (Turvey 1992). If perception-action is inherently prospective, there is no need for the nervous system to bring the perceiver-actor "up to speed" because perception-action places awareness "ahead of the world."

The prospectivity of perception-action is considered by some to be one of the fundamental hallmarks of a psychological being (E. Gibson 1994). From the perspective of ecological psychology, the stimulation variables that support such prospectivity are not the static and isolated variables of standard physics (so-called "lower-order" stimulation variables) but, rather, are the dynamic and relational variables of an ecological physics (so-called "higher order" stimulation variables) (Turvey & Shaw 1999). For example, a handheld object's resistance to rotational acceleration in different directions not only informs a perceiver about whether that object can be used to achieve a particular goal (e.g., striking another object) but also about how that object should be used to do so (Wagman & Carello 2001; 2003). If perception-action is characterized as a prospective act, then there is no need for delay compensation mechanisms in perception-action because higher-order relational variables are sufficient to specify impending states of affairs without the need for mediating processes.

Delays that are inherent in the sending and receiving of information create an explanatory gap in a scientific understanding of perception and action. However, rather than fill that gap with specialized delay compensation mechanisms, I propose that perception and action be (re)characterized in a way in which such delays are an impossibility and the explanatory gap dissolves. The ecological approach to perception-action provides such a (re)characterization.

## Visuomotor extrapolation

doi: 10.1017/S0140525X08004044

David Whitney

*Department of Psychology, and The Center for Mind and Brain, University of California, Davis, CA 95618.*

**dwhitney@ucdavis.edu**

**http://mindbrain.ucdavis.edu/content/Labs/Whitney/**

**Abstract:** Accurate perception of moving objects would be useful; accurate visually guided action is crucial. Visual motion across the scene influences perceived object location and the trajectory of reaching movements to objects. In this commentary, I propose that the visual system assigns the position of any object based on the predominant motion present in the scene, and that this is used to guide reaching movements to compensate for delays in visuomotor processing.

Nijhawan's article provides evidence for compensation mechanisms in visual perception and visually guided action. Most of this evidence is drawn from the flash-lag effect, where a single object moves across the retina. There are several other illusions, some of which are briefly mentioned in the target article, which might also support Nijhawan's position (De Valois & De Valois 1991; Hess 1904; Matin et al. 1976; Nishida & Johnston 1999; Ramachandran & Anstis 1990; Regan & Beverley 1984; Snowden 1998; Whitaker et al. 1999; Whitney & Cavanagh 2000). For a review of these illusions, see Whitney (2002). The strongest support for compensation in the perceptual system (e.g., extrapolation) comes from the displacement of stationary edges by motion. For example, visual motion viewed through a static aperture causes the aperture to appear shifted in the direction of the motion (De Valois & De Valois 1991; Ramachandran & Anstis 1990; Regan & Beverley 1984); the motion aftereffect is accompanied by a concurrent shift in the apparent position of a static test stimulus (Nishida & Johnston 1999; Snowden 1998; Whitaker et al. 1999); and, static flashed objects appear shifted in the direction of nearby motion (Whitney & Cavanagh 2000). Whereas the flash-lag effect may be due to differential latencies for moving and flashed objects (Ogmen et al. 2004; Purushothaman et al. 1998; Whitney & Murakami 1998), these other mislocalizations of static edges by visual motion cannot be caused by temporal mechanisms such as differential latencies (Whitney 2002).

Although these mislocalizations of static edges by visual motion provide the strongest support for perceptual extrapolation (i.e., compensation for neural delays in the perceptual system), some of these illusions greatly complicate things: Several papers have shown that flashed objects appear shifted forward, in a direction consistent with any nearby visual motion, even when that motion is several degrees away from the flash. This has been called the "flash-drag" effect or the "flash-shift" effect (Durant & Johnston 2004; Eagleman & Sejnowski 2007; Watanabe et al. 2002; 2003; Whitney 2006; Whitney & Cavanagh 2000; 2003). Because the flash is not moving, and it is distantly separated from the moving object, it does not immediately make sense why the flash should appear shifted (or extrapolated) in the direction of nearby motion. This result is somewhat difficult to reconcile with the notion of compensation for moving object positions, but it is not entirely incompatible. In fact, this flash-drag effect suggests that the sort of compensation that Nijhawan describes for a single moving object extends to all objects, and may be a far more pervasive and important mechanism than simply allowing us to perceptually extrapolate a baseball or other moving object's position.

In Nijhawan's article, the primary case that is considered is one in which a single moving object needs to be perceived or grasped. This is a relatively rare situation compared to what normally happens: usually, there is image motion across the entire retina, not just a single moving object. Normally the world is physically stationary, and it is we (our eyes, heads, or bodies) that move around; and it is our movement which generates retinal image motion. For example, when we reach to any object, we usually make an eye or head movement during or just before the reach. In this case, there is retinal motion of the scene and the target object. On account of delays in visual processing, delays in coordinate transformations, and other factors such as imperfect efference copy signals (Bridgeman 1995) – along with the fact that targets of reaching movements are coded in eye-centered coordinates (Buneo et al. 2002; Crawford et al. 2004; Henriques et al. 1998) – our visuomotor system faces a somewhat similar challenge to the one outlined by Nijhawan, but on a much grander scale. Because of these visuomotor delays, we should miss-direct nearly every reaching movement we make to virtually any object. Every time we reach toward our coffee cup, we should either hit the cup, knocking it over, or fall short of the cup – all because of sluggish visual and motor processing.

How does the visuomotor system avoid these errors? In a recent series of studies, we found that the visuomotor system samples motion across the visual field and then shifts the trajectory of the hand in the direction of that motion when reaching to any object in the scene (Whitney & Goodale 2005; Whitney et al.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

2003; 2007). This effect was recently called *the manual following response* (Gomi et al. 2006; Saijo et al. 2005) and reveals an adaptive mechanism: The visuomotor system uses retinal motion to gauge movements of the eye and body (probably because it is as fast or faster than using vestibular or proprioceptive cues), and then adjusts the trajectory of the reach based on this information to improve the accuracy of goal-directed action. In support of this, when subjects were passively rotated, the presence of background retinal image motion improved the accuracy of reaching movements compared to cases in which only static information, or nothing, was visible (Whitney et al. 2003). The manual following response is conceptually similar to the "flash-drag effect" described above, and it suggests that the visual and visuomotor systems use retinal image motion (the kind generated every time we move our eyes) to update/extrapolate/shift the representations of object position (causing objects to appear shifted in position) – and this allows us to guide our hand more accurately than would otherwise be possible.

This visuomotor extrapolation model has the advantage that it accounts for several psychophysical findings that are discrepant with the perceptual extrapolation model; and it also has the advantage that it explains accurate visuomotor behavior under the most common circumstances – where the world is stationary and we are moving.

## Compensation for time delays is better achieved in time than in space

doi: 10.1017/S0140525X08004056

Myrka Zago[a] and Francesco Lacquaniti[b]

[a]*Scientific Institute Santa Lucia, 00179 Rome, Italy;* [b]*University of Rome "Tor Vergata" and Scientific Institute Santa Lucia, 00179 Rome, Italy.*

**m.zago@hsantalucia.it**

**lacquaniti@caspur.it**

**Abstract:** Mechanisms of visual prediction based on spatial extrapolation work only for targets moving at constant speed, but do not easily accommodate accelerating or decelerating motion. We argue that mechanisms based on temporal extrapolation deal with both uniform and non-uniform motion. We provide behavioural examples from interception of falling objects and suggest possible neurophysiological substrates of time extrapolation.

Nijhawan makes a clear case for the need to compensate for delays arising from processing and transmission times. The evidence for compensation in perceptual decision and visual awareness appears somewhat controversial (Eagleman & Sejnowski 2000; Krekelberg & Lappe 2001), but the evidence for compensation for motor reactions to a rapidly changing sensory stimulus is uncontroversial. Typical visuomotor delays in ballistic interception of fast targets (such as in catching or hitting in ball games) are about 200 msec – at least an order of magnitude longer than the temporal accuracy required for interception (about $\pm$10 msec). Unless the nervous system has built-in mechanisms to compensate for such delays, the interception program would be based on obsolete visual information about target motion, and, as a consequence, the target would be badly missed.

Nijhawan proposes a mechanism for neural compensation of delays that is based on a spatial extrapolation linearly related to the time delay. According to his hypothesis, visual prediction would be concerned primarily with horizontal processes, which transmit neural information between two neighbouring retinotopic sites. The speed of neural transmission and the distance between neighbouring neurons along the horizontal direction would jointly determine the amount of spatial and temporal extrapolation. Another mechanism could consist in a shift of the receptive field in response to moving stimuli. Sundberg et al. (2006) found that neurons in monkey area V4 exhibit such a shift in response to a particular type of moving stimuli. The direction of the receptive field shift was opposite to the direction of target motion, as if the cell had been recruited by a wave of activity preceding the target. Ferrera and Barborica (2006) argued that a moving target would leave a trail of refractory neurons in its wake so that spiking activity would be shifted toward the leading edge.

Interestingly, mechanisms of visual prediction based on spatial extrapolation, such as those mentioned above, work only for targets moving at constant speed (uniform motion), because the spatial shifts co-vary with the time samples in a fixed manner. Most targets, however, accelerate or decelerate to a variable extent. Let us consider a very common situation – that of motion affected by Earth's gravity, such as free-fall, ballistic, pendulum, or wave motion. Although all objects are accelerated downward by gravity at the same rate, the corresponding acceleration of the retinal image is not at all constant, being inversely related to the apparent viewing distance of the object. The question then is how the central nervous system (CNS) compensates for delays in the case of accelerating or decelerating motion. Here we show that temporal extrapolation rather than spatial extrapolation can more easily do the job.

Figure 1A depicts space–time plots similar to those of Figures 3 and 4 of Nijhawan, but for an object moving at constant acceleration (when the spatial variable decreases from right to left) or deceleration (when the spatial variable increases from left to right). The dashed curve depicts the physical trajectory, and the dotted curve depicts the corresponding trajectory "seen" by a neuron with a fixed visual delay. Clearly, the spatial shifts required to compensate for the visual delay (solid line segments connecting the two curves) are not constant anymore, as they were in the spatial extrapolation scheme proposed by Nijhawan.

In theory, a first-order model might be used to approximate a second-order motion. One such model is provided by the tau function, tau $= x(t)/v(t)$, where $x(t)$ is the spatial position of the target and $v(t)$ is the corresponding velocity (Lee 1976). However, it can be shown that, in case of free-fall motion from relatively short drop heights, such an approximation would imply significant temporal errors in interception (>50 msec), corresponding to the difference between the time-to-contact predicted by tau and the actual time-to-contact of the ball accelerated by gravity (Zago & Lacquaniti 2005). In fact, we know that unless taken by surprise, people can easily intercept targets descending along the vertical accelerated by gravity (Lacquaniti & Maioli 1989; Zago et al. 2004); they generally intercept

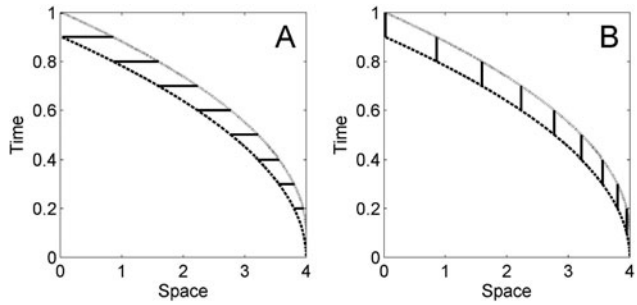



Figure 1 (Zago & Lacquaniti). Figure depicts space-time plots similar to those of Figures 3 and 4 in the target article, but for an object moving at constant acceleration (when the spatial variable decreases from right to left) or deceleration (when the spatial variable increases from left to right). The dashed curve depicts the physical trajectory and the dotted curve depicts the corresponding trajectory "seen" by a neuron with a fixed visual delay.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

such targets much better than similar targets descending at constant speed (McIntyre et al. 2001; Zago et al. 2004). The inescapable conclusion, therefore, is that neural compensation uses second-order information. However, direction, velocity, and acceleration are probably encoded together at the level of single neurons both in retinal ganglion cells (Thiel et al. 2007) and in visual motion area MT (Lisberger & Movshon 1999; Price et al. 2005; Schlack et al. 2007). It may not be an easy task for neural populations to disentangle acceleration estimates independently of velocity and to extrapolate target trajectory in a time-varying fashion, as required by the schematic graph of Figure 1A.

An alternative solution, however, is to extrapolate trajectories in time, rather than in space. This is shown in Figure 1B, where the same spatio-temporal trajectories of an accelerating or decelerating object as those of Figure 1A are plotted. Here, however, the compensation for the visual delay is implemented by means of a temporal shift (black segments connecting the two curves). The resulting correction at each sampled position of the target is constant throughout the trajectory. Needless to say, time extrapolation works equally well for uniform motion.

Behaviourally, it has been shown that when subjects must intercept novel types of vertical target motion, they learn to improve their performance by shortening or lengthening central processing time, with the direction of change depending on the task (Senot et al. 2005; Zago & Lacquaniti 2005; Zago et al. 2004; 2005). This is equivalent to modulating the time shifts in Figure 1B. We do not know where in the brain a time extrapolation mechanism might be implemented, though it could occur at the interface between sensory and motor systems. There is some evidence that neurons in posterior parietal cortex (area 7a and LIP) may encode elapsed time by means of ramping activity (Leon & Shadlen 2003; Maimon & Assad 2006; Merchant et al. 2004).

ACKNOWLEDGMENTS

This work was supported by a DCMC grant from the Italian Space Agency (ASI) and by a grant from the Italian Ministry of Health (RF ISPESL).

# Author's Response

## Predictive perceptions, predictive actions, and beyond

doi: 10.1017/S0140525X08004068

Romi Nijhawan

*Department of Psychology, University of Sussex, Falmer, East Sussex BN1 9QH, United Kingdom.*

**romin@sussex.ac.uk**

**http://www.sussex.ac.uk/psychology/profile116415.html**

**Abstract:** Challenges to visual prediction as an organizing concept come from three main sources: (1) from observations arising from the results of experiments employing unpredictable motion, (2) from the assertions that motor processes compensate for all neural delays, and (3) from multiple interpretations specific to the flash-lag effect. One clarification that has emerged is that visual prediction is a process that either complements or reflects non-visual (e.g., motor) prediction.

The hypothesis that visual prediction (VP) is a central component of visual perception attracts its fair share of nay-sayers. It has been over 15 years since I first set up a mechanical display in a basement at the University of California, Berkeley, using a 7-up soda can, and observed the phenomenon. Although I was intrigued at the time, I never anticipated that the next decade and a half would be spent sorting it out. The attraction for me, as I suspect it is for many of the commentators, lies in the way this simple phenomenon forces consideration of the relationship between time and space, the physical and the perceived, perception and action, and brain and mind. A more implicit theme emerging is consistent with Mach's notion of the "elements" (cf. Ratliff 1965) and "neutral monism" that reality is ultimately one, and it is neutral between the physical and the perceived.

My response is divided into six sections. The first two sections, R1 and R2, are on general philosophical, psychophysical, and neurophysiological issues concerning phenomenal consciousness (sect. R1) and the role of motion in defining perceptual space (sect. R2). Several commentators have suggested that motor prediction, which could take into account visual delays in addition to non-visual delays, renders visual prediction unnecessary. In the third section (R3), I address this alternative point of view, consider alternative accounts of flash-lag, clarify misinterpretations, and consider not just predictions but also the other side of the coin, violations of prediction. The next section (R4) addresses the points raised by commentators, ranging from suggestions that delays do not exist, to the proposal that delays exist but frequently do not require compensation, to the point that delays are optimal and may, in fact, be beneficial. The next section (R5) answers to several commentators who have questioned the applicability of the VP framework beyond single objects moving at constant velocity. The final section (R6) discusses commentaries that suggest mechanisms (e.g., adaptation) responsible for VP, or, instead, for alternative favored mechanisms for compensation of neural delays. It focuses on the role of feedback during interactions with objects in specifying retinal "local sign," and on temporal (as opposed to spatial) extrapolation.

### R1. Visual prediction (VP) is unaffected by the outcome of the functionalism versus epiphenomenalism debate

As **Khurana** points out, renaming "extrapolation" as *visual prediction* draws attention to its function. What is the function of phenomenal consciousness? Some scientists and philosophers ascribe functionality to consciousness, and some believe consciousness to be an epiphenomenon – that all mental states are causally inert and cannot impact physical (brain) states. The flash-lag effect is part of phenomenal experience and the visual prediction (VP) hypothesis arises from a claim about the phenomenal experience of moving objects. So, does the truth of VP depend on a functional account of phenomenal consciousness? The answer to this is, no. As Khurana suggests, in the newer rendition, VP is unaffected by the outcome of the debate. If one believes that mental things have a function, then a veridical perceived position of moving objects (with lag removed) would be useful for whatever phenomenal consciousness is useful for. If mental things do not have



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

a function, then a veridical perceived position of moving objects could simply be a byproduct of predictive visual neurons that supply other brain structures (e.g., motor planning centers) with compensated visual input.

## R2. Dynamic calibration of visual space: The fundamental role of motion in perception of space

**McFarland** points out a potential difficulty for compensation for neural delays in the processing of physical events, as the absolute reference for calibration is not known. **Van de Grind** offers a possible solution, arguing that moving objects do not require a reference, as moving objects *are* the reference. He proposes that position and time are relative quantities and motion is the absolute quantity. Van de Grind states: "For the FLE [flash-lag effect] that would mean that the position map is always tuned to motion, so that moving objects are normally seen veridically at their instantaneous physical position, regardless of their processing time."

This is a fascinating proposal. Arguments from developmental neurobiology and evolution bear on this. Movement has been identified as a fundamental sensory dimension (Nakayama 1985) because of its association with survival of the individual and the species. Many primitive visual systems do not respond to stimulation unless the stimulus is moving, and the ability of eyes to signal to the brain in the absence of movement, it seems, is present only in higher animals (Gregory 1979). Indeed, the more primitive extreme periphery of the human retina is only capable of signaling movement. Thus, there are visual systems that do not respond to stationary stimuli, but none that do not respond to movement.

Movement is also necessary for visual perception. Stationary stimuli disappear rapidly when their images are stabilized on the retina (Coppola & Purves, 1996). Although micro eye-movements allow for sustained vision in adults, these movements are not sufficient for strongly activating the cortical cells of a developing visual system. Many cortical cells in newborn mammals do not give a sustained response even to high-contrast patterned stimuli, unless the stimuli move (Wiesel & Hubel 1963). Consistent with this is the finding that newborn humans respond much more strongly to moving than to stationary stimuli (Wickelgren 1969). These findings suggest that visual motion must be the primary source of stimulation for neonatal visual systems, particularly the visual cortex (Hubel & Wiesel 1959). Because neural activity is essential in the normal development of the visual system (Blakemore & Cooper 1970), and cortical activity is essential for perception, this suggests that motion plays a primary role in first establishing the normal mapping between points on the retina and the corresponding perceptual positions in space (Fawcett & O'Leary 1985).

Movement plays a role in early development even before the onset of visual experience. Spontaneous ganglion-cell action potentials, seen as patterns of activity sweeping across the retina at different speeds, are essential to the normal development of the visual system (Butts et al. 1999). These sweeping patterns are similar to waves of spikes produced by moving optical images in ganglion cells of the mature retina (Berry et al. 1999; Meister, Wong, Baylor, & Shatz 1991). Thus, a significant proportion of pre- and postnatal neural activity of the immature visual system results from movement.

**Van de Grind** states that "the usual way of looking at the flash-lag effect (FLE) is to (implicitly) assume that the brain contains fixed position-maps (e.g. retinacentric, head-centric, etc.) and that incoming retinal signals need to be assigned a place on these maps," and he asks whether this is a correct description. Indeed, it may be incorrect to view such maps as fixed position-maps if motion-dominated early experience contributes in an essential manner to the mapping between retinal points and perceived positions. We have suggested elsewhere (Nijhawan & Khurana 2002) that in a developing visual system lacking neural specificity (Barlow & Pettigrew 1971), when a retinal point $(m, n)$ is stimulated by motion, the corresponding perceived object is localized at point $(m', n')$ in visual space. Once motion-based topographic maps are set up, these maps are further refined and are eventually used by the mature visual system to represent stationary stimuli. Thus, in the mature system, when retinal point $(m, n)$ is stimulated by a stationary object, it is also localized at position $(m', n')$. This proposal (Nijhawan & Khurana 2002), which we call the "dynamic calibration of visual space," reinforces van de Grind's suggestion that motion is primary, and that the position of stationary stimuli is determined by the principles established by moving objects – not the other way round, as is more commonly assumed.

**Whitney** also ascribes a fundamental role to motion, as motion in the scene can influence the position of not just moving objects, but also that of other static objects (Whitney 2002). What determines the direction of influence? Because visual delays should cause a backward shift in the perceived position of moving objects, a forward shift in the presence of motion immediately suggests processes geared to undoing the backward shift. The observation that moving objects cause nearby flashes to shift in the direction of motion is not well understood, but it seems likely that whatever causes moving objects to appear shifted forward also causes the flashes to appear shifted. Contrary to frequent assumptions that a flash is a static object (e.g., Whitney & Cavanagh 2000), a brief flash neither moves nor remains stationary; neither does it move relative to the retina when the eye is stationary, nor when the eye is moving. Thus, in the realm of moving/nonmoving stimuli, a brief flash may be characterized as an ambiguous object. Similar to all ambiguous stimuli, then, flashes should be susceptible to contextual effects. To the extent that moving objects are forward-shifted, and flashes are influenced by motion signals, flashes will also appear forward-shifted (but see Shi & Nijhawan, under review). Truly static objects, such as motion-defined stationary Gabor patches (De Valois & De Valois 1991) and longer duration flashes, are seen as shifted forward because these objects do not contain strong cues as to their position. I would argue, however, that flashes and motion-defined patches are typically not shifted forward to the same extent as moving objects because stationary reference objects (e.g., the computer monitor) reduce the forward shift.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

## R3. Predictive perception, alternative interpretations, misinterpretations, and the flash-lag effect

### *R3.1. Motor compensation and accounts of flash-lag*

There are many types of prediction – motor, visual, and cognitive – and **Kerzel & Müsseler** underscore this generality. However, phenomena of such generality would require mechanisms of comparable diversity. Consequently, predictive mechanisms may be widespread within the nervous system. This also stands to reason, as the brain is a massively interacting computational device. Any significant neural function in one part of the brain (e.g., motor) can influence another part of the brain, such that both parts will show that function. But Kerzel & Müsseler suggest that visual mechanisms do not contribute to prediction. This is unparsimonious, as more that 50% of the cortex serves a visual function.

**Sheth & Wu** suggest that, "The agent can often overcompensate for the movements of the target, thus arriving at a common intersection point some time before the target," in claiming that motor prediction is more flexible in compensating for delays in some situations, as it can tolerate slop. A clear example where compensation for delays may not be necessary is where a stationary animal hunting for fish is positioned in a narrow, fast-flowing stream. However, some tasks, such as a frog striking a flying insect with its tongue, would be impossible without extrapolation in visual neurons (possibly in the tectum). Furthermore, in the given fishing example, and myriad other similar examples, even skill (which **Cutting** favors) and motor compensation (which Sheth & Wu favor) are not required. Just as this example does not diminish the importance of skill and motor compensation generally, examples in which visual compensation may seem unnecessary do not diminish the importance of visual compensation generally. The animal needs to localize not only objects, but also its end-effectors. Prediction in the motor domain but not in the visual domain would result in a spatial mismatch between visual objects and end-effectors, which would in turn interfere with the formation of a unified spatial map. A mechanism for "motor-sensory unification" is considered in section R6.5.

**Purushothaman, Bedell, Öğmen, & Patel** [**Purushothaman et al.**] argue the point that compensation for delays is carried out by motor processes. It must be noted, however, that there is debate neither in the literature nor in the target article challenging the extremely important role of motor processes. The examples these commentators give are useful; however, these do not rule out the possibility of predictive visual processes. Although Purushothaman et al. are correct in saying that most of the evidence supporting prediction involves behavior (e.g., saccades), they fail to note that "purely" visual processes are not as divorced from motor processes as was once believed (Wilson & Knoblich 2005).

**Smeets & Brenner** misattribute the following claim to the target article: "In his target article, Nijhawan proposes that early visual processing provides the prediction that is needed to deal with sensory-motor delays when we interact with moving objects, rather than such prediction arising from complex motor strategies as is generally assumed." Visual prediction does not replace motor strategies, but rather, suggests that prediction in vision either complements or reflects prediction in other (nonvisual) processes. In addition, VP could provide crucial data not just for interactions but also for decision processes such as when to act, and whether to act or withhold action.

**Sheth & Wu** suggest that VP makes only a minor contribution to the flash-lag effect, invoking multiple other factors. In some cases, large magnitude effects, for example, some geometrical illusions, have been difficult to account for with a single mechanism, so multiple mechanisms have been invoked (Coren & Girgus 1973). However, there are still unresolved issues of basic causes versus modulatory influences for the flash-lag effect.

**Baldo & Klein** have championed an attentional framework for understanding the flash-lag effect. There is no doubt that attention has a role to play in the flash-lag effect, however, what this role is remains unclear, as there are discordant findings (Baldo & Klein 1995; Khurana et al. 2000; Khurana & Nijhawan 1995). Another difficulty for the attentional framework is that there are many forward-shift effects that do not involve flashes (e.g., see Maus & Nijhawan 2006), so movement of attention is not present, while forward-shift effects are.

### *R3.2. Additional delays in the localization of moving objects would be dangerous*

**Smeets & Brenner** claim that the visual system determines the precise location of a moving object only when there is a need to know the object's location, not throughout the object's motion trajectory. They suggest that the flash-lag effect occurs because the observer needs to determine the position of the object after the flash is presented.

I claim that the visual system acquires and has continuous position information of a smoothly moving object. This is because precise position of a moving object could be required to guide action at an unpredictable instant. Localization is a fundamental task, and the mechanisms serving localization are frequently tested (and selected) in trying situations, such as predator–prey interactions. On account of survival pressure, animal nervous systems cannot tolerate any added delays in localization of moving objects. Continuous location information concerning a moving object has a survival advantage, as the relevance of a moving object can change drastically from one moment to the next.

Imagine a monkey being hunted by a tiger, which is slowly approaching from the right. The monkey is not unduly concerned, as there is a sure getaway – there is a large tree near the monkey to the left that it can quickly climb. Now imagine that there is second tiger hiding behind the tree, whose presence is given away by the sound of a twig breaking under the tiger's paw. Under this scenario, and innumerable similar scenarios in nature, the animal that already has location information regarding an important moving object (in this example, the approaching tiger from the right) will be able to plan its next escape strategy more quickly, and thus have a survival advantage over an animal that must first establish the location of the moving object following an unpredictable event (in this example, the sound of the twig).



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

### *R3.3. The flash$_\phi$/flash$_\psi$ muddle*

The non-controversial aspect of the various accounts of the flash-lag effect is the assumption of a significant delay between the presentation of a flash and its perception. However, keeping the two events, flash presentation and perception, separate is as difficult as it is important to be explicit as to which of the two meanings one intends when using the term *flash*. In order to be more precise, I will denote the event physical-flash by flash$_\phi$ and the event psychophysical-flash by flash$_\psi$. Note that the event flash$_\psi$ is short for the perception of the flash.

Now let us turn to the moving object. Visual input delay ($\Delta t_i$) should cause an observer to perceive a continuously moving object in a position that it occupied $\Delta t_i$ seconds ago. The initial impetus for the suggestion that this statement is incorrect came from the flash-lag effect (Nijhawan 1994), and motion extrapolation was suggested as a process that undoes this lag. Happily, **Eagleman** suggests that *postdiction*, which until now has been considered to be the antithesis of prediction, can also achieve this at least partially.

The initial basis for the postdiction account was the observation that the flash-terminated condition does not lead to the flash-lag effect (Eagleman & Sejnowski 2000). Postdiction suggested that the visual system integrates information after an event, and this integrated information can impact perceptual decisions about an event in the immediate past. A semantic analogy would be a word appearing later in a sentence giving added precision to (or disambiguating) the meaning of a word appearing earlier in the sentence. **Eagleman** suggests that continuous input from future events could update past delayed information to align the perceived position of a moving object with its actual position. But what happens when a flash is presented? On the current rendition of postdiction, when the moving object is physically aligned with flash$_\phi$, it is also perceptually aligned with it (as the delay between the perceived and physical position of the moving object has been removed by the postdictive process). Thus, when event flash$_\psi$ occurs, the perceived position of the moving object is ahead, which explains the flash-lag effect. However, on this view, a flash-lag effect should also occur in the flash-terminated condition. The lack of effect in the flash-terminated condition now poses a problem for postdiction, even though it was the initial impetus for it.

Now let us consider the flash-initiated condition. Although it is correct that motion information after event flash$_\phi$ (which is what is commonly meant by the term "flash") drives the flash-lag effect, it is equally correct to say that motion input before event flash$_\psi$ drives the flash-lag effect. The suggestion by **Smeets & Brenner** that the dependence of the flash-lag effect on what happens after the "flash" (more precisely, flash$_\phi$) opposes VP, is incorrect. By definition, in the flash-lag effect, the observers report that flash$_\psi$ lags behind the moving object, so flash$_\psi$ is the all-important event. Therefore, this event should play the central role in all analyses. According to the analysis presented in section 5.2.1 of the target article (also see Figure 7 there), more than 98% of the time period immediately before flash$_\psi$, during which the moving object is in view, is available for motion extrapolation. So, it is the past, not the future, motion that contributes to the flash-lag effect.

**Eagleman** believes apparent motion to be a quintessential postdictive phenomenon. In apparent motion, two dots, dot 1 and dot 2, are flashed in two positions in quick succession. The perception is that dot 1 appears to move to dot 2. Movement of the dot could have started only after dot 2 appears, and yet the movement of the dot implies that it occupies positions between dot 1 and dot 2, which are past positions. This phenomenon has puzzled many researchers, and Eagleman has identified a possible postdictive account of this. I believe this phenomenon seems puzzling due to acquiescence. Influential scientists have continually reminded us that we have no direct immediate access to the world (Neisser 1976) and that visual perception is a generative process that "creates" the visual image (Kandel & Wurtz 2000). In the absence of such reminders, we acquiesce into believing that our visual percepts directly reflect the state of the world.

There is a temporal version of this acquiescence, which results in the flash$_\phi$/flash$_\psi$ muddle and engenders a postdictive account of apparent motion. It is well known that motion can be perceived even when there is no object appearing to move. Sitting in a quiet room, we can sometimes have the experience that something moved, without perceiving what moved. In the far periphery of our retina, a stimulus is perceived only when it is moving; as soon as movement stops, the object vanishes. This is what happens in apparent motion: say event flash1$_\phi$ occurs at time 1 and flash2$_\phi$ quickly follows at time 2. This leads to neural signaling of motion. These motion signals occur before event flash2$_\psi$ (because flash2$_\psi$ is subject to conduction delays). Thus, at the time of flash2$_\psi$, the motion computation has already occurred (on the basis of flash2$_\phi$), so the motion percept is interjected before the percept of the second flash. The seemingly postdictive nature of apparent motion comes about because we forget that flash2$_\phi$ and flash2$_\psi$ are distinct events.

In conclusion, postdiction does not cause the flash-lag effect, but it does contribute to its elimination it in the flash-terminated condition. Backward masking in which a later stimulus renders an earlier stimulus invisible is another excellent example of a postdictive phenomenon. By letting later information have an impact, postdiction has the advantage of obtaining more reliable data about the world, and disambiguating information when the data are noisy.

### *R3.4. Going nowhere by misinterpreting and mislabeling visual prediction*

**Öğmen, Patel, Purushothaman, & Bedell** [**Öğmen et al.**] and **Purushothaman et al.** appear to have misinterpreted the basics of VP. They even mislabel it as "perceptual compensation model" (PCM), and then proceed with criticisms that are directed at PCM and not VP. At first, it might seem that the authors may have simply switched the word "perceptual" for "visual," as one may interchangeably use the terms in describing the perceptual system or the visual system. This would not be a serious problem. However, there is a deeper issue, as revealed in several statements by Öğmen et al. and Purushothaman et al. Consider the following statement by Purushothaman et al.: "Perceptual compensation in early visual areas, as suggested by the target article, need not occur."



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

In the target article I suggested that perception is the outcome of visually sensitive neurons that are capable of compensating for visual delays, or that compensatory motor processes can influence perception (e.g., via feedback from motor areas to visual areas). The percepts reveal this compensation (e.g., in the flash-lag effect) because of compensation mechanisms in the feedforward vision pathway and/or motor pathways. Thus, perception is *not* a process that produces compensation, as the authors seem to suggest. Rather perception is the outcome of compensation mechanisms.

**Öğmen et al.**'s misinterpretation of VP is underscored when they draw a contrast between perceptual compensation (their term) and sensory compensation. They say: "That sensory processes may be predictive is neither new nor controversial. . . . What is controversial is the specific hypothesis proposed by Nijhawan, namely, compensation of neural delays at the perceptual level." It is, however, highly unlikely that sensory processes can be predictive while the perceptions that these sensory processes contribute to are not.

Contrary to what **Purushothaman et al.** report, not only does pursuit eye-movement yield a flash-lag, but the magnitude of the effect with pursuit is about the same as in the standard object motion case (Nijhawan 2001; also see **van de Grind**'s commentary). Purushothaman et al.'s suggestion that other motion-bias phenomena (De Valois & De Valois 1991; Ramachandran & Anstis 1990) do not support the notion of compensation for delays, is not widely shared (see, e.g., commentary by **Whitney**). Finally, **Öğmen et al.**'s favored explanation of our fading motion data in terms of lowered detection threshold resulting from integration along the motion path, is incorrect; we explicitly tested this hypothesis by doubling the trajectory length of the motion stimulus and found the detection threshold to not change significantly (Maus & Nijhawan 2006, p. 4379).

**Öğmen et al.** claim that the target article is selective, as it does not address the luminance-dependent modulation of the flash-lag effect. This is surprising, as section 6.1. of the target article is devoted to this very issue. I would like to add that many everyday objects consist of parts of different luminance (e.g., a fair-skinned person with black hair), and movement of such objects does not lead to spatial distortions. This supports the view that the visual system takes differential latency (due to variations in luminance) into account. This brings me to a related point concerning the "flash-lead" effect (Purushothaman et al. 1998), which the commentators mention. The authors have claimed that a flash-lead effect occurs when the display combines a moving object of low luminance contrast with a bright flash. Let me say that in all my experience of working with the flash-lag effect since 1990, I have never observed a flash-lead effect based on luminance differences. This is also true of displays that combine a moving object that is close to absolute detection threshold and a stroboscopically generated flash that is a million times brighter! This has also been noted by others (see, e.g., Brenner & Smeets 2000, and **Smeets & Brenner**'s commentary).

### *R3.5. Neural response to the instantaneous value of continuous input*

In the interaction between a monkey and a tiger that **Krekelberg** considers, the smooth changes in direction that the tiger produces are more to test the monkey's endurance than to challenge its visual system. What is demanding for the visual system in predator–prey interactions is the sudden changes in movement direction that the animals often produce. Such energetically costly movements are selected precisely because smooth changes in movement are predictable.

There have been several studies on motion per se, showing predictive responses of neurons (e.g., Fukushima et al. 2002). **Krekelberg**'s interpretation that a change in the response maximum of a middle temporal area (MT) neuron to different motions shows that this neuron is unable to compensate for the delay, is incorrect. The possibility of relating the neural response of a cell to the "wrong" instantaneous value of a continuously changing input was discussed in the target article for orientation tuning in primary visual cortex (V1) (see sect. 1.3). There is a relatively straightforward interpretation of the interesting results shown by Krekelberg. The so-called directional label of a MT neuron depends not just on the intrinsic properties of the neuron, but also on the context of the stimulus. In the present case, the directional tuning curve of an MT neuron depends on the history of the stimulus. In general, if the movement direction of the stimulus changes continuously, it will produce a different tuning curve than if the stimulus moves in a fixed direction within the MT cell's receptive field (in the vast majority of studies, fixed direction stimulus is used). Furthermore, the tuning for clockwise motion versus counterclockwise motion would be different, as the history of the stimuli is different for the two motions. Continuous input to a cell leads to adaptation in the cell's response, and clockwise (CW) and counterclockwise (CCW) motions will cause differential adaptations. In addition, a single neuron does not receive input only from the retina, but also from other neighboring cells, which are differently tuned. These factors, and not differential delay in the cell's peak response to CW and CCW motions, are the likely cause of the observed difference in the peak response of the cell described in the commentary.

### *R3.6. Prediction and violation of prediction in the flash-lag phenomenon*

**Baldo & Klein** suggest that the original motion extrapolation account (Nijhawan 1994) nicely explains the flash-lag effect in its purest form. They state, however, that:

> in order to accommodate that explanation with discordant empirical findings, Nijhawan, [. . .] now resorts to at least three different putative mechanisms: (1) the role of lateral communications between neighboring retinotopic locations, (2) the impact of visual transients generated by offsets, and (3) a dynamical interplay between motor and visual processes.

Baldo & Klein are correct in saying that the target article focuses more on new physiological mechanisms, but neither have these mechanisms been mentioned here for the first time, nor are they introduced to explain discordant findings.

Take the role of transients in the localization of moving objects, which was first introduced by Nijhawan (2002). In nature, predictable events and events that violate predictions go hand in hand, so one goal of the target article was to look at both sides of the coin. Prey animals often



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

produce abrupt trajectory changes in an attempt to throw off a following predator. There is recent neurophysiological evidence that the retina not only predicts moving objects but also strongly signals violations of prediction (Holy 2007; Schwartz et al. 2007b) resulting from sudden changes in the trajectory of moving objects. It is natural to think of an abruptly disappearing moving object in the flash-terminated condition, or sudden change in object trajectory in the motion-reversal condition, as violating prediction. I suggest that the strong transient signals generated by violations of prediction interact postdictively (Eagleman & Sejnowski 2007) with predictive signals and overwrite them. We tested this claim by making the stimulus disappear without this violation and found the predictive-overshoot (Maus & Nijhawan 2006) that is typically missing in the flash-terminated condition. Furthermore, when the strength of the transients is manipulated, transients of intermediate levels result in intermediate forward displacements (Maus & Nijhawan, in press).

**Howe, Horowitz, & Wolfe** [**Howe et al.**] tested our account by doing the opposite experiment: they attempted to eliminate the flash-lag effect by introducing a transient. In their experiment, as an object moved on a gray background, it abruptly reversed contrast at the time of the flash from black to white. The observers reported a flash-lag effect. The commentators claim that this contradicts the proposal by Nijhawan (2002) and Maus and Nijhawan (2006) that the transient in the flash-terminated condition eliminates the predictive-overshoot. Before analyzing Howe et al.'s interesting experiment, it is important to note that there is a transient at the beginning of motion in the flash-initiated condition, which fails to eliminate the flash-lag effect (Khurana & Nijhawan 1995). So, clearly, the timing of the transient is crucial.

One can think of the contrast reversal experiment as consisting of two phases. The first phase, with a black moving object, comprises the flash-terminated condition, and the second phase, with a white moving object, comprises the flash-initiated condition. We know that the latter condition does and the former condition does not produce a flash-lag. For a black-to-white transition, the important question to ask, however, is: Did the flash appear to lag the black or the white moving object? (It is not sufficient simply to report a flash-lag effect here). If the white object appeared ahead of the flash (which would be my guess), then this is not surprising, as this is just the flash-initiated condition producing the effect. In the biased competition model presented in the target article (see sect. 5.2.2.) the visual system gives higher weighting to newer retinal input. In the case of the flash-initiated condition, the initial transient resulting from motion onset does not fix the position of the moving object because there is newer (motion) input after the transient.

It would be of great interest if the black object appeared ahead of the flash. This would be the flash-terminated condition producing a predictive-overshoot (which is normally not seen). Such a finding would mean that the motion input resulting from the white object following contrast reversal weakens the black-to-white transient. A weakened transient, in turn, is less able to suppress the prediction-based representation of the black object. In conclusion, a transient will fix the location of a moving object only if (1) it comes at the end of an event (as a final event), so that it has the opportunity to compete strongly with, and mask, the previously setup prediction-based representation, and (2) no new input is there to suppress the transient.

### *R3.7. TTC models and VP*

**Benguigui, Baurès, & Le Runigo** [**Benguigui et al.**] suggest that the visual system is not an accurate predictor of moving objects if the last part of their trajectory is occluded for more than 200 msec, contrary to the claim made in the target article, which suggests that the time period is much longer. However, the 200 msec limit does not significantly change my conclusion. Time-to-collision (TTC) theorists suggest that the nervous system uses information contained in the changing retinal image to compute TTC. In the target article, I claim (and make a case) that a corollary of the TTC approach is that visual mechanisms compensate for the delays that would be incurred in the transmission of the changing retinal image to the central nervous system (CNS). This claim will only be affected if accuracy in prediction is compromised for an occlusion duration that is close to the duration of visual processing delay. Since, according to Benguigui et al. prediction ability is compromised for occlusion durations of 200 msec and beyond, which is longer than visual processing latency, my claim is unaffected. So the restated claim would read: Because TTC can be predicted accurately 200 msec into the future on the basis of retinal information, the position of the moving object can certainly be predicted ~80 msec into the future, as required by the motion extrapolation account of the flash-lag effect.

## R4. Are neural delays optimal, and therefore compensation unnecessary?

### *R4.1. Are tolerance for delays and skill enough to account for human performance?*

**McFarland** points out that the absolute reference location of a moving object is unavailable. Adding to the suggestions of **van de Grind**, I argue that the need for a reference for locating moving objects may not arise, as motion is absolute (see sect. R2). Alternatively, I suggest that a position assigned to moving objects by the visual system is determined by the success or failure of the animal that uses that information during action; if the behavior produced by the animal on the basis of position information is successful, contributing to the animal's survival, then the position is correct (and it is the same as that of the end-effector engaged in the production of action; see sect. R6.5). Otherwise, it is incorrect. McFarland claims that compensation may be unnecessary, and phase-alignment between different signals may be sufficient. Phase-alignment may solve one problem, but it raises another. Even if phase-alignment causes fusion of the to-be-aligned signals (say, vision and touch) the question still remains: How is the animal able to successfully interact with a visuo-tactile object, for example, an insect moving rapidly on the skin, given that both vision and touch give delayed position information?



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Both **McFarland** and **Cutting** describe an experiment on audiovisual speech where the auditory and visual (lip movements) signals are asynchronous. The findings are that observers fail to detect moderately large asynchronies. Cutting suggests that the examples of large tolerances show that compensation is not necessary. I have an alternative interpretation of these interesting findings: The nervous system reduces discrepancies between two (or more) neural representations that belong to the same world event. This function of VP is outlined in the target article (see section 9.1. "Coordination of neural representations"). A previous study has shown that the brain can use visual distance to calibrate the simultaneity of auditory and visual signals of audio-visual objects (Sugita & Suzuki 2003). Thus, the asynchrony between auditory speech and its visual signal is undetected not because there is a wide range of tolerance, but because the brain actively coordinates the auditory and visual signals that likely belong to the same physical event.

There is a more fundamental aspect of this coordinating influence. There are many neurons in the CNS that receive convergent input from more than one modality, for example, audition and vision. Consider the consequences of unchecked lags between auditory and visual channels from a neurophysiological perspective. As vision is slower than audition, differential processing delays would mean that the input to bimodal neurons will be asynchronous, and for moving audiovisual objects, the two signals will be spatially offset in location maps. But, most importantly, an asynchrony would mean that the convergent inputs from moving objects would repeatedly fail to activate such bimodal neurons. This has implications not just for perception but also for the survival of such neurons. One possible solution could be a delay introduced in the faster modality; however, this would mean that early warning from an audiovisual object would cause the animal to respond with the latency of the slower modality. A much more adaptive solution would be for the nervous system to extrapolate the slower modality.

**Cutting** argues that compensation for delays is not necessary for fine-timed interceptions and claims that behaviors in sports are better understood in terms of skill acquisition over many years (combined with a genetic predisposition). There is no doubt that skill and genetics are important in everyday interactions and crucial in modern-day sports. However, the suggestion that skill can render compensation for sensory delays unnecessary is probably wrong. Highly skilled actions have features in common with ballistic actions (such as saccades); in both cases, the role of sensory data is reduced/altered. Hence, one may argue that compensation for sensory input delays during such open-loop actions becomes less relevant. However, in extreme cases of skilled performance, one may conjecture that no sensory input, let alone compensated sensory input, is required. Consider the performance of a knife-throwing artist. With enough skill, the artist can throw knives blindfolded. However, it is known from research on deafferented patients, for example, that there are limitations to open-loop actions; while some aspects of open-loop actions seem unaffected others aspects are compromised (Sarlegna et al. 2006). So sensory data are required even in highly skilled acts, and compensation for sensory delays is important.

**Wagman**, however, denies the very existence of neural delays. He challenges the view that perception and action are information-processing tasks, and so transmission of information between perception-to-action is not even possible. I am happy to treat perception-action as a hyphenated quantity, and I agree that the wide conceptual gap between perception and action that is sometimes assumed to exist is unjustified. Furthermore, I agree that the relational variables (e.g., in the expanding retinal image) are the most relevant in guiding actions. However, I do not see how these positions lead to the conclusion that delays in perception and action do not exist. In any event, the issue of delays discussed at length in the target article is not one of delays between perception and action but, rather, between a physical event and its perception, and a physical event and an action directed at it. So it is not clear whether commentator Wagman is addressing the main issues raised there.

### *R4.2. The faster the better, but not too fast please!*

**Lin** suggests that the target article overestimates the visual processing delays. He further claims that the need for compensation depends on the assumption that perception occurs beyond V1. I disagree. The requirement for compensation is virtually independent of which segment of the vision pathway is considered necessary and sufficient for visual perception. This is so because it is well established that visual delays are nontrivial, even at the first processing level; photoreceptor hyperpolarization peak to a bright flash occurs with a latency of about 50 msec. Many actions in high-speed ballgames require a temporal window that is narrower than photoreceptor delays (Tresilian 1993). Delays quickly become large as they accrue along the vision pathway. Hence, there is no question of compensation being unnecessary for delays before V1.

**Lin**'s suggestion that neural delays are optimal – they are long enough to allow sufficient computation time, but not so large so as to make the responses too slow – is interesting. Similarly, **Changizi** points out that contrary to the assumption that faster neural processes would always confer an advantage to the animal, there may be a trade-off between computation speed and complexity. Changizi makes the point with an insightful analogy between nervous systems and computers. Computation speeds have increased by several orders of magnitude in the past two decades, and yet we still have to wait for software to load. Clearly, modern programs are doing a lot more. This behavior (rather than instantaneous loading of programs that do a lot less) is selected for in the computer world. In the real world, some behaviors can be produced with little computation – and relatively quickly – whereas others require elaborate computations. The knee-jerk reflex, for example, requires a single synapse and can be produced within 40 msec. However, such responses are stereotyped and have a limited (though important) function. The plasticity, flexibility, and complexity characteristic of many primate behaviors requires significant processing and elaborate networks. At the level of perception, the issue of compensation for delays may arise precisely because complex and time-consuming computations (e.g., integration of retinal



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

image size and retinal disparity carried out by cortical cells in the computation of size constancy) are required for accurate localization of moving objects in three-dimensional space.

However, there are two things to note here. First, there is no doubt that increased speed of transmission of unmodified information over long distances, for example by peripheral sensory-motor nerves, can only be beneficial. Second, there are several other types of trade-off that are relevant. One is speed and size of the nervous system. Although certain invertebrate nervous systems are capable of producing fast reactions (e.g., the escape reaction in the squid), these reactions are possible only by the virtue of large-diameter nerve fibers. Myelination of neurons in vertebrates results in signal transmission that is more than an order of magnitude faster without the need to support much larger fibers (and brains). Therefore, although all this speed may not translate into faster behavior of vertebrates, generally speaking there is certainly some speeding up. Thus, it may be correct that software-loading times have not gone down linearly with the increase in computation speed but there is some general speeding up of computers.

Perhaps the most significant type of trade-off is revealed by comparison of processing speeds for predictable versus unpredictable events. In animal behavior, unpredictable events are of great importance. For example, in predator–prey interactions, each animal attempts to make gains over the opponent by producing sudden changes in trajectory. Such behaviors involving high accelerations are energetically expensive, so the fact that they are common in nature underscores their importance; these behaviors exist because the opponent's nervous system takes time to process them. It is likely that the upper limit of how much delay the animal's visual system will tolerate depends on how quickly the animal can respond to unpredictable visual events. A brain that is slower because of its fancier computations would take a long time in communicating an important, unpredictable, event (such as the predator adopting an attack posture). Thus, slowness due to fancier computations that may allow for extrapolation farther into the future, as suggested by **Changizi**, may indeed have no cost for highly predictable events, but the resulting slowness in the registration of unpredictable events (that are presumably processed by the same brain structures) will confer a huge disadvantage.

## R5. How does the VP framework fare with multiple objects, observer motion, and accelerating motion?

Movement velocity is hardly ever constant in nature, and there are always multiple objects present. Does VP come into play in such real-world scenarios?

**Smeets & Brenner** claim that in typical interactions with moving objects, the subject attempts to smoothly pursue the object. This eliminates retinal image motion, and obviates the need for VP in typical interaction scenarios. Animals, however, rarely have the luxury of tracking a single moving object with their eyes. There are usually many relevant objects in the scene, both stationary and those moving at various velocities. Smooth pursuit of one object will cause other stationary objects to move on the retina. So, invariably, there will be retinal image motion, and the issue of neural delays and compensation become relevant. Imagine a monkey being hunted by two tigers, one near by to the left and stationary, and the other farther away to the right, advancing toward the monkey. The monkey needs to maintain fixation on the nearby tiger to monitor its fine movements. The movement of the advancing tiger's retinal image, and the transmission delay of this image to the cortex, if uncompensated, would lead to an error in the visual position of the tiger. VP would provide compensated data relevant for the monkey in planning its escape manoeuvre.

**Sheth & Wu** and **Zago & Lacquaniti** suggest that the fact that common objects do not travel at constant speed is problematic for visual (spatial) prediction. However, when considering movement at the time scale of neural delays (~100 msec), velocity is approximately constant in most situations. Furthermore, for objects that produce a large change in velocity in 100 msec (e.g., when an animal jumps to escape a predator), even the predictive motor mechanisms, which Sheth & Wu favor, will fail. Failure of predictive motor mechanisms in such situations does not reduce the importance of motor mechanisms in other situations, in which acceleration is smaller or zero. Likewise, scenarios in which VP does not apply do not rule out the role of VP generally.

**Whitney** notes that single moving objects in the environment are rare, and furthermore, typically it is the observer that is moving. Therefore, VP may not apply in most real-world scenarios. In the target article, I briefly discuss the situations in which the observer's eyes, limbs, or the whole body are moving (see the beginning of sect. 7). The issues that arise for observer movements are very similar to those arising for object motion. When an animal moves its arm, or its whole body forward, what is its instantaneous position as represented by the continuously changing neural representation of the moving item? Cells in the hippocampus known as place cells, and entorhinal cortex, are responsible for coding the animal's location. Delays in neural transmission will cause the registered location to be delayed relative to the animal's physical location. A similar localization issue arises when the animal moves its sensors or end-effectors. For example, cells coding for head-direction in the horizontal plane (Taube et al. 1990) should respond with a lag in relation to the physical direction of the head. However, it is known that some head-direction neurons lead physical head direction of the animal (Blair and Sharp 1995). The target article focuses more on moving objects than on moving end-effectors because localizing moving limbs and eyes is considered more complex.

However, I had an even more important reason for addressing the simpler object motion case first. Many articles (e.g., see Krekelberg & Lappe 2001, p 335) begin by posing a puzzle: "Do we perceive the position of a moving object with a delay because of neuronal latencies?" The goal of the target article is not just to address accurate perception of baseballs and the like, but also to address the puzzle of visual awareness and its neural underpinnings. Yet another goal is to draw focused attention to the issues of perception and reality in the spirit of Ernst Mach, whose lifetime research goal was "to distinguish in any phenomenon, the properties of the observer from the properties of the thing observed" (Ratliff 1965, p 1).



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

## R6. Mechanisms for compensation

### *R6.1. Implementation of Visual Prediction*

I agree with **Trautteur, Datteri, & Santoro** [**Trautteur et al.**] that the first goal of the target article is to make a convincing case that prediction is a widespread function of the nervous system. This goal would be impossible to achieve without inclusion of visual processes – hence my focus on VP. Indeed, my hope is to reduce the opposition to VP that has existed in raging controversies over the past decade or so. How VP is implemented is the next exciting question, and I am happy that the authors think that there are some useful rough sketches in the target article. Toward this goal, I would like to ask: Might there be as much variety in mechanisms as there are time scales for performing different actions, as suggested by **Balkenius & Gärdenfors**? **Stepp & Turvey** outline a novel approach in their commentary that may show well-known phenomenon of neural synchronization (Sillito et al. 1994; Neuenschwander et al. 2002) in the new light of visual predictive phenomena. I think there is likely to be a great variety in mechanisms for prediction, only a few of which we have any details on.

### *R6.2. Early versus late compensation: Are internal models indispensable?*

**Stepp & Turvey** help to not only broaden the basis of anticipation (to use their term) but also bring clarity to the issues. First, I agree with these commentators (and with **Benguigui et al.**) that it is wrong to assume that internal models are the basis of all anticipation; it was not my intention to confer this exclusivity to anticipatory mechanisms. Indeed, there are anticipatory neural mechanisms that do not fit the description of internal models (Barlow 1953), and Stepp & Turvey mention an interesting example from physics, termed *anticipating synchronization*. Indeed, anticipatory phenomena may lie in realms beyond animal behavior, cellular networks, or single cells, and may be present in physical systems (e.g., in semiconductor lasers with optical feedback). This suggests an approach that may be termed *anticipatory behavior without internal models*, which is reminiscent of a previous approach to *intelligent behavior without internal representations* (Brooks 1991).

**Stepp & Turvey** further point out that anticipating synchronization shares features (e.g., delayed feedback) with the behavior of animals and nervous systems. In anticipatory synchronization there is a driver and a slave. The observation is that the slave can synchronize with the future state of the driver. Is such a system neurophysiologically plausible? The answer is clearly, yes. There are many examples of synchronization of neural responses in the retina, lateral geniculate nucleus (LGN) and cortex (Neuenschwander et al. 2002). Consider the LGN-V1 driver–slave pair, with LGN as the driver and V1 as the slave. In this system, V1 sends delayed feedback to LGN. It is not difficult to imagine V1 activity synchronizing with the future state of LGN activity (Sillito et al. 1994). It is noteworthy, however, that the basis of prediction may be triplets of neural locations (instead of pairs) that are capable of producing zero time-lag synchronization (Chawla et al. 2001).

**Stepp & Turvey** (and **van de Grind**) have suggested that I have portrayed VP as an exclusively high-level model. I disagree. In the target article, I argue that prediction is ubiquitous in the CNS, with both early and late mechanisms devoted to it; findings such as the flash-lag effect in the flash-initiated condition (Khurana & Nijhawan 1995) and the color decomposition effect (Nijhawan 1997) would be difficult to explain in terms of high-level mechanisms. Furthermore, there are neurophysiological considerations that support early mechanisms for prediction (see sect. R4.1). Nonetheless, anticipation can manifest itself at "high" levels where one cannot escape internal models (Erlhagen 2003; Snyder 1999), and synchronization and phase alignment become inappropriate metaphors. I would like to reiterate the proposal in the target article: For a function as fundamental as prediction, the nervous systems uses as many tricks as possible, and the phenomena revealing this function may show up in many guises.

### *R6.3. Time from adaptation*

Every time we make a saccadic eye-movement there is a discrete, but predictable, change in visual input. **Mayo & Sommer** describe a process in which adaptation from one input to the next could be used for the timing of the sequence of discrete stimuli. A modified version of such a timing mechanism could play a role in compensation for neural delays. The mechanism is revealed when two sequential stimuli are presented to neurons in the superior colliculus and the frontal eye field. For inter-stimulus intervals of less than 400 msec, the neural response magnitude to the second stimulus varies precisely with the delay between the first and the second stimulus.

**Mayo & Sommer** have wondered how such a mechanism could be adapted to compensate for delays. I speculate on one possibility for estimating delays involving three visually responsive neural sites (A, B, and C). Sites A and B are hierarchically organized, with site B farther away from the retina than site A. Site C receives input from both A and B. If a single stimulus is presented, neurons in site A will respond before those in site B. Thus, a single world event will lead to two successive neural events that are temporally separated by the delay between site A and site B. The delay between the responses of site A and site B can be estimated from the adaptation level (reduced response) of neurons in site C to the input from site B (second input). This would be one way in which the visual system could monitor its own delays, and not just delays between two external events.

**Sheth & Wu** describe a neural activity profile in a topographical (cortical) map generated by the movement of an object. This profile includes the neural activity correlated with the object's distant past, immediate past, current, and immediate future positions; the locus of maximum cortical activity in the map represents the "current" location of the object. These commentators point out that the neural mechanisms leading to this profile are basic to all neurons. However, there is another basic neural mechanism that the commentators have missed, and that is spatiotemporal filtering, which has been invoked as one of the mechanisms leading to spatial extrapolation. One mechanism that Sheth & Wu do mention is neural adaptation.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

**Mayo & Sommer** make a clever proposal that adaptation could be used by neural networks in the timing of sequential stimuli, and perhaps in compensation of neural delays.

Consider again **Sheth & Wu**'s statement that the maximum of neural activity in a topographic map represents the object's current location. There are multiple topographic maps in the brain, so the described profile of neural responses will be present in all of the maps stimulated by a moving object. However, as these maps are located at various distances from the receptor sheet (e.g., in the superior colliculus, cortex, etc.), the neural activity maxima representing the object's current position will correspond not to one but to many positions in space. Some of these maps (e.g., superior colliculus) serve mainly a motor function (e.g., eye-movements), but others might mainly serve perception. Input from two or more such maps to a target site that produces an average would be required to eliminate the multiplicity in perceived positions. With little modification, an extrapolation process could achieve the same goal.

### *R6.4. Modeling forward-shift effects*

**Berzhanskaya** presents a detailed computational model that might produce not only the flash-lag effect, but also several other well-known forward-shift effects. One may, however, ask why these computational models behave the way they do. Is there any functional significance to this behavior of the models, assuming that natural visual systems behave like this? Towards this, then, one may wonder whether prediction versus emergent property of directional selective mechanisms, as the commentary title suggests, is the correct contrast to draw. Consider an analogy: human photoreceptors are most sensitive to light around the yellow-green part of the spectrum, despite the fact that blue photons are more energetic. One might explain this in terms of the photochemical interactions and the neural networks tuned to produce a maximum response to yellow-green light. This answer is correct and suffices for certain purposes. However, the fact becomes more meaningful when one looks deeper into the cause. The facts that maximum energy in the solar spectrum is in the yellow-green region, and that the absorption coefficient of water, which the human eye contains, is smallest precisely over the range for which human vision is most sensitive, give a better understanding of the peak sensitivity in the yellow-green part of the spectrum (Nunes 2003). Likewise, the psychophysical phenomena or networks that produce forward shifts would be understood better in the context of neural delays and prediction.

**Berzhanskaya** has focused on direction selectivity. It is worth mentioning that the flash-lag effect has been observed in retinal ganglion cells of the rabbit and salamander that are directionally insensitive. Another point to note is related to the commentators's explanation of the flash-lag effect in terms of shift in position of the flash due to trailing inhibition left behind by the moving object. This predicts that the flash should be mislocalized relative to stationary objects in the direction opposite to object motion. We have found that flash mislocalization in the direction opposite to motion may contribute, however, it is not sufficient to explain the full flash-lag effect (Shi & Nijhawan, under review).

### *R6.5. Modifying perceptions*

**Balkenius & Gärdenfors** note that the task for predictive sensory-motor processes is more complex than is portrayed by the target article, and, in particular, that the target article underestimates the contribution of feedback from past interactions in creating a compensated percept of the moving object. I agree with the commentators. This is indeed an omission, which I will attempt to rectify here.

**Balkenius & Gärdenfors** propose that predictions concerning a moving object's visual position result in an anticipated now, but it is not known until later whether the prediction was correct or incorrect. In the target article, I say that: "during a catch, the visual system modifies the perceived position of the ball so that it matches the position of the moving hand." I now elaborate on this with the example of interceptive action directed at a moving object. A successful catch implies that the animal's CNS was able to issue motor commands to send the hand to the correct location. Now the question is: To which physical position in space did the visual neural representation of the moving object correspond at the time of interaction? Feedback (vision, touch, and sound) from the results of the action helps in answering this question for the CNS. The CNS uses the motor-sensory representation (motor command and proprioception) of the hand as a reference to which the visual representation is compared. Successful interactions lead to integration of hand and visual positions, and the moving object is assigned the same visual position as the hand. Helmholtz first suggested that the visual position code (local sign) for an object could be trained by the motor position code when an end-effector, for example, the subject's moving hand, stimulates the subject's own retina. There is evidence that cells in the posterior parietal cortex respond to visual stimulation resulting from the movement of limbs during action (Mountcastle et al. 1984). Because the feedback interactions occur in a plastic neural environment, the error between spatial position of the visual object and the hand is removed over multiple successful interactions.

### *R6.6. Extrapolation in time or space?*

**Zago & Lacquaniti** propose that extrapolation in the domain of time rather than space would be more useful for the visual system. They argue that time-extrapolation would be better adapted to handling moving objects in nature, as these objects rarely travel at constant velocities. However, a question immediately arises: Is a time-extrapolated input the same as a reduced latency one? (This of course raises the question of how this is different from the differential latency accounts by Whitney & Murakami [1998] and Purushothaman et al. [1998]). The main problem with time-extrapolation, as Zago & Lacquaniti recognize, is one of finding a mechanism. Again, if the suggestion is analogous to latency reduction, then this can be achieved by neural facilitation along the motion path (Grzywacz & Amthor 1993). However, this can only be achieved on the basis of lateral interaction, which is again a spatial process. **Mayo & Sommer** discuss a purely temporal mechanism based on neural adaptation, which is reliable and unaffected by trial-to-trial variations in response delays. It remains to be seen whether such



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

a mechanism would have the necessary speed to produce the interception of accelerating objects, as has been demonstrated by Lacquaniti and colleagues. A time-extrapolation could supplement the task of prediction, and work cooperatively with space-extrapolation to yield a more robust system than each type of extrapolation alone would.

**Cunningham** suggests that in addition to spatial compensation for delays, there should also be temporal compensation; and in addition to compensation for delays for continuous events, there should also be compensation for discrete events. There are perhaps two ways of describing temporal compensation. One is similar to whatz **Zago & Lacquaniti** might call time-extrapolation. This may be called *absolute temporal compensation*, as it affects delays between a physical event and its registration. The other is *relative temporal compensation*, where the nervous system actively coordinates sensory signals to compensate for temporal asynchronies between modalities (see commentary by **Cutting**).

Relative temporal compensation is what **Cunningham** and colleagues have shown in their interesting experiments. Let us consider absolute temporal compensation further. First, there is a form of absolute temporal compensation, as demonstrated by experiments on finger tapping to auditory tones. Within limits, human subjects can predict repetitive discrete tones and put their finger taps just in advance of the tones (Mates et al. 1994). This is a good example of where the sensorimotor system must have information about the actual time of an external event. But can absolute temporal compensation occur for perception (as opposed to behavior such as finger taps)? Can the visual system generate the percepts of repetitive flashes, for example, simultaneously with the actual flashes? The answer would appear to be no, and this is where spatial and temporal compensations differ. The claim of spatial compensation is that it can put the percept of the moving object closer to (or even ahead of) the actual position of the moving object. In contrast, temporal compensation cannot, it seems, put the perceptual event close to the time of the actual event.

I fully endorse the amendment suggested by **Cunningham** that: "The goal of visual prediction is to use priors acquired from both previous experience and the currently unfolding visual stimulus to create a perceived state of the world that matches, as far as possible, the actual state of the world."

### *R6.7. Mental extrapolation and (not "or") visual extrapolation*

**Kerzel & Müsseler** suggest that sensorimotor prediction and mental extrapolation, as opposed to visual extrapolation, can overcome perceptual latencies. There is no doubt that sensorimotor prediction is an important, and highly flexible, contributor to successful behavioral acts. The role of mental extrapolation in the context of flash-lag effect is, however, not as clear as these commentators propose. Kerzel & Müsseler claim that the missing predictive-overshoot in the flash-terminated condition opposes the visual extrapolation account, and they invoke mental extrapolation in explaining the forward-shift of fading moving objects (Maus & Nijhawan 2006). However, it is not clear how mental extrapolation escapes this very criticism. Why does mental extrapolation not lead to an overshoot in the flash-terminated condition?

Mental extrapolation falls in the general category of phenomena such as mental imagery and mental rotation. Researchers have investigated the neural basis of mental imagery. One of the key findings is that mental imagery tasks engage the primary visual cortex (Kosslyn & Sussman 1994). In addition, Kosslyn and colleagues have found a number of similarities between mental imagery and visual perception, such as the topographic nature of both representations (Kosslyn et al. 1995). Thus, the existence of mental extrapolation would predict the existence of visual extrapolation. **Kerzel & Müsseler**'s proposal that mental extrapolation exists but visual extrapolation does not is unparsimonious.

I claim that the task of mental extrapolation is not to solve the problem of neural conduction delays, but rather, it is to determine when a moving object, occluded by another object, will reappear (Wexler & Klam 2001). In the case of continuous sensory input from a moving object, the task of mental extrapolation is to determine the object's future position. The task of visual extrapolation is to use sensory input to determine the object's current position (after compensating for visual conduction delays). In past studies, Kerzel and colleagues have used either probe stimuli presented after a retention interval, or pointing movements, and so in effect asked for the remembered final position of the moving target. In flash-lag experiments, or in the task used by Maus and Nijhawan (2006), participants make an online perceptual judgment comparing the position of the moving target to a flash or to a static probe. Although obviously the observer's response is given after the visual offset, the judgment is based on simultaneously visible stimuli. It is likely that the two experimental methods differentially engage mental and perceptual extrapolation. In this context, it is interesting to note that the forward-shift effect of the fading moving object observed by Maus and Nijhawan (2006) is 175 msec, which is a much larger shift than the typical flash-lag effect of 80 msec. It is possible that this is a cumulative effect of both visual and mental extrapolation.

## Reference

**Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.**


Ahissar, E. & Ahissar, M. (1994) Plasticity in auditory cortical circuitry. *Current Opinion in Neurobiology* 4:580–87. [DWC]

Aho, A. C., Donner, K., Helenius, S., Larsen, L. O. & Reuter, T. (1993) Visual performance of the toad (*Bufo bufo*) at low light levels: Retinal ganglion cell responses and prey-catching accuracy. *Journal of Comparative Physiology A* 172(6):671–82. [aRN]

Alais, D. & Burr, D. (2003) The "flash-lag" effect occurs in audition and cross-modally. *Current Biology* 13(1):59–63. [MVCB, JB, aRN]

Anderson, B. L. & Barth, H. C. (1999) Motion-based mechanisms of illusory contour synthesis. *Neuron* 24:433–41. [JB]

Andersen, R. A. & Buneo, C. A. (2002) Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience* 25:189–220. [aRN]

Andersen, R. A., Snyder, L. H., Li, C. S. & Stricanne, B. (1993) Coordinate transformations in the representation of spatial information. *Current Opinion in Neurobiology* 3(2):171–76. [aRN]




https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Anderson, C. H. & Van Essen, D. C. (1987) Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Science USA* 84(17):6297–301. [aRN]

Anderson, C. H., Van Essen, D. C. & Gallant, J. L. (1990) Blur into focus. *Nature* 343(6257):419–20. [aRN]

Anstis, S. (2007) The flash-lag effect during illusory chopstick rotation. *Perception* 36:1043–48. [JB]

Anstis, S. M., Smith, D. R. & Mather, G. (2000) Luminance processing in apparent motion, Vernier offset and stereoscopic depth. *Vision Research* 40(6):657–75. [aRN]

Arbib, M. A. (1972) *The metaphorical brain: An introduction to cybernetics as artificial intelligence and brain theory*. Wiley Interscience. [aRN]

Attneave, F. (1954) Informational aspects of visual processing. *Psychological Review* 61:183–93. [HO]

Bachmann, T. & Poder, E. (2001) Change in feature space is not necessary for the flash-lag effect. *Vision Research* 41(9):1103–106. [BRS]

Bahill A. T. & LaRitz T. (1984) Why can't batters keep their eyes on the ball? *American Scientist* 72:249–53. [GP]

Bahill, A. T. & McDonald, J. D. (1983) Smooth pursuit eye movements in response to predictable target motions. *Vision Research* 23:1573–83. [GP]

Bair, W., Cavanaugh, J., Smith, M. & Movshon, J. (2002) The timing of response onset and offset in macaque visual neurons. *Journal of Neuroscience* 22(8):3189–205. [BKr]

Baldo, M. V. & Caticha, N. (2005) Computational neurobiology of the flash-lag effect. *Vision Research* 45(20):2620–30. [aRN, MVCB]

Baldo, M. V. C., Kihara, A. H., Namba, J. & Klein, S. A. (2002) Evidence for an attentional component of perceptual misalignment between moving and flashing stimuli *Perception* 31:17–30. [MVCB]

Baldo, M. V. C. & Klein, S. A. (1995) Extrapolation or attention shift? *Nature* 378(6557):565–66. [MVCB, rRN]

(in press) Paying attention to the flash-lag effect. In: *Space and time in perception and action*, ed. R. Nijhawan & B. Khurana. Cambridge University Press. [MVCB]

Baldo, M. V. C. & Namba, J. (2002) The attentional modulation of the flash-lag effect. *Brazilian Journal of Medical and Biological Research* 35:969–72. [MVCB, BRS]

Balkenius, C. & Johansson, B. (2007) Anticipatory models in gaze control: A developmental model. *Cognitive Processing* 8:167–74. [CB]

Barlow, H. B. (1953) Summation and inhibition in the frog's retina. *Journal of Physiology* 119:69–88. [arRN]

(1961a) Possible principles underlying the transformations of sensory messages. In: *Sensory communication*, ed. W. A. Rosenblith. Wiley. [aRN, HO]

(1961b) Three points about lateral inhibition. In: *Sensory communication*, ed. W. A. Rosenblith, pp. 782–86. MIT Press. [HO]

(1961c) The coding of sensory messages. In: *Current problems in animal behaviour*, ed. W. H. Thorpe & O. L. Zangwill, pp. 331–60. Cambridge University Press. [HO]

(1979) Reconstructing the visual image in space and time. *Nature* 279(5710):189–90. [aRN]

(1981) The Ferrier Lecture, 1980. Critical limiting factors in the design of the eye and visual cortex. *Proceedings of the Royal Society of London B: Biological Sciences* 212(1186):1–34. [aRN]

Barlow, H. B. & Pettigrew, J. D. (1971) Lack of specificity of neurones in the visual cortex of young kittens. *Journal of Physiology* 218(1):98P–100P. [rRN]

Barnes, G. R. & Asselman, P. T. (1991) The mechanism of prediction in human smooth pursuit eye-movements. *Journal of Physiology* 439:439–61. [JBJS]

Barnes, G. R., Barnes, D. M. & Chakraborti, S. R. (2000) Ocular pursuit responses to repeated, single-cycle sinusoids reveal behavior compatible with predictive pursuit. *Journal of Neurophysiology* 84:2340–55. [GP]

Batista, A. P., Buneo, C. A., Snyder, L. H. & Andersen, R. A. (1999) Reach plans in eye-centered coordinates. *Science* 285(5425):257–60. [aRN]

Bedell, H. E. & Lott, L. A. (1996) Suppression of motion-produced smear during smooth pursuit eye movements. *Current Biology* 6:1032–34. [HO]

Bedford, F. L. (1993) Perceptual learning. *Psychology of Learning and Motivation* 30:1–60. [DWC]

Beilock, S. L., Bertenthal, B. I., McCoy, A. M. & Carr, T. H. (2004) Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin and Review* 11:373–79. [JEC]

Benguigui, N., Broderick, M. P., Baurès, R. & Amorim, M. A. (in press) Motion prediction and the velocity effect in children. *British Journal of Developmental Psychology*. [NB]

Benguigui, N., Ripoll, H. & Broderick, M. P. (2003) Time-to-contact estimation of accelerated stimuli is based on first-order information. *Journal of Experimental Psychology: Human Perception and Performance* 29(6):1083–101. [NB]

Bennett, K. M. B. & Castiello, U. (1995) Reorganization of prehension components following perturbation of object size. *Psychology and Aging* 10:204–14. [NB]

Berry, M. J., Brivanlou, I. H., Jordan, T. A. & Meister, M. (1999) Anticipation of moving stimuli by the retina. *Nature* 398(6725):334–38. [MVCB, arRN, GP, WvdG]

Berzhanskaya, J., Grossberg, S. & Mingolla, E. (2004) Motion-to-form cortical projections and distortion of the position maps. Paper presented at the Vision Sciences Society Conference, Sarasota, FL, May 1–5, 2004, p. 149. [JB]

(2007) Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision* 20(4):337–95. [JB]

Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R. & Warland, D. (1991) Reading a neural code. *Science* 252(5014):1854–57. [aRN]

Blair, H. T. & Sharp, P. E. (1995) Anticipatory head direction signals in anterior thalamus: evidence for a thalamocortical circuit that integrates angular head motion to compute head direction. *Journal of Neuroscience* 15(9):6260–70. [rRN]

Blake, R., Tadin, D., Sobel, K. V., Raissian, T. A. & Chong, S. C. (2006) Strength of early visual adaptation depends on visual awareness. *Proceedings of the National Academy of Sciences USA* 103(12):4783–88. [ZL]

Blakemore, C. & Cooper, G. F. (1970) Development of the brain depends on the visual environment. *Nature* 228(5270):477–78. [rRN]

Blakemore, S. J., Wolpert, D. M. & Frith, C. D. (2002) Abnormalities in the awareness of action. *Trends in Cognitive Sciences* 6(6):237–42. [aRN]

Bolz, J., Rosner, G. & Wässle, H. (1982) Response latency of brisk-sustained x and brisk-sustained y cells in the cat retina. *Journal of Physiology (London)* 328:171–90. [DWC]

Bootsma, R. J. & Oudejans, R. R. (1993) Visual information about time-to-collision between two objects. *Journal of Experimental Psychology: Human Perception and Performance* 19(5):1041–52. [aRN]

Bootsma, R. J. & van Wieringen, P. C. W. (1990) Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance* 16:21–29. [NB, JEC]

Born, R. T. & Bradley, D. C. (2005) Structure and function of visual area MT. *Annual Review of Neuroscience* 28:157–89. [BKr]

Brenner, E. & Smeets, J. B. (2000) Motion extrapolation is not responsible for the flash-lag effect. *Vision Research* 40(13):1645–48. [arRN, JBJS]

(2007) Eye movements in a spatially and temporally demanding interception task. *Journal of Vision* 7(9):565. [JBJS]

Brenner, E., Smeets, J. B. J. & van den Berg, A. V. (2001) Smooth eye movements and spatial localization. *Vision Research* 41:2253–59. [aRN]

Bridgeman, B. (1995) A review of the role of efference copy in sensory and oculomotor control systems. *Annals of Biomedical Engineering* 23(4):409–22. [DW]

Bringuier, V., Chavane, F., Glaeser, L. & Fregnac, Y. (1999) Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science* 283(5402):695–99. [aRN]

Britten, K. H. (2004) The middle temporal area: Motion processing and the link to perception. In: *Visual neurosciences*, ed. L.M. Chalupa & J. S. Werner. MIT Press. [GP]

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience* 13:87–100. [GP]

Brooks, R. (1991) Intelligence without representation. *Artificial Intelligence* 47:139–59. [rRN]

Bruce, C. J., Friedman, H. R., Kraus, M. S. & Stanton, G. B. (2004) The primate frontal eye field. In: *Visual neurosciences*, ed. L. M. Chalupa & J. S. Werner. MIT Press. [GP]

Bullock, D. (2003) Motoneuron recruitment. In: *The handbook of brain theory and neural networks*, ed. M. A. Arbib. MIT Press. [aRN]

Buneo, C. A., Jarvis, M. R., Batista, A. P. & Andersen, R. A. (2002) Direct visuomotor transformations for reaching. *Nature* 416(6881):632–36. [DW]

Burr, D. C. (1980) Motion smear. *Nature* 284(5752):164–65. [aRN]

Burr, D. C. & Morgan, M. J. (1997) Motion deblurring in human vision. *Proceedings of the Royal Society of London B: Biological Sciences* 264(1380):431–36. [aRN]

Burr, D. C. & Ross, J. (1979) How does binocular delay give information about depth? *Vision Research* 19(5):523–32. [aRN]

Butts, D. A., Feller, M. B., Shatz, C. J. & Rokhsar, D. S. (1999) Retinal waves are governed by collective network properties. *Journal of Neuroscience* 19(9):3580–93. [rRN]

Butz, M. V., Sigaud, O., Pezzulo, G. & Baldassarre, G., eds. (2007) *Anticipatory behavior in adaptive learning systems: Advances in anticipatory processing*. Springer. [GT]

Cai, R. & Schlag, J. (2001) A new form of illusory conjunction between color and shape. *Journal of Vision* 1(3):127, 127a. [JB]

Cantor, C. R. L. & Schor, C. M. (2007) Stimulus dependence of the flash-lag effect. *Vision Research* 47:2841–54. [WvdG]

Carpenter, R. H. S. (1988) *Movements of the eyes*, 2nd edition. Pion Press. [GP]

Cavanagh, P. (1997) Predicting the present. *Nature* 386(6620):19, 21. [aRN]




https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Changizi, M. A. (2001) "Perceiving the present" as a framework for ecological explanations of the misperception of projected angle and angular size. *Perception* 30:195–208. [MAC]
(2003) *The brain from 25,000 feet: High level explorations of brain complexity, perception, induction and vagueness*. Kluwer Academic. [MAC]
Changizi, M. A., Hsieh, A., Nijhawan, R., Kanai, R. & Shimojo, S. (in press) Perceiving-the-present and a systematization of illusions. *Cognitive Science*. [MAC, aRN]
Changizi, M. A. & Widders, D. (2002) Latency correction explains the classical geometrical illusions. *Perception* 31:1241–62. [MAC]
Chappell, M., Hine, T. J., Acworth, C. & Hardwick, D. R. (2006) Attention "capture" by the flash-lag flash. *Vision Research* 46:3205–13. [MVCB]
Chawla, D., Friston, K. J. & Lumer, E. D. (2001) Zero-lag synchronous dynamics in triplets of interconnected cortical areas. *Neural Networks* 14(6–7):727–35. [rRN]
Chen, S., Bedell, H. E. & Öğmen, H. (1995) A target in real motion appears blurred in the absence of other proximal moving targets. *Vision Research* 35:2315–28. [HO]
Chung, S. T. L., Patel, S. S., Bedell, H. E. & Yilmaz, O. (2007) Spatial and temporal properties of the illusory motion-induced position shift for drifting stimuli. *Vision Research* 47:231–43. [GP]
Ciszak, M., Marino, F., Toral, R. & Balle, S. (2004) Dynamical mechanism of anticipating synchronization in excitable systems. *Physical Review Letters* 93:114102. [NS]
Coenen, A. M. & Eijkman, E. G. (1972) Cat optic tract and geniculate unit responses corresponding to human visual masking effects. *Experimental Brain Research* 15(5):441–51. [JPM]
Collewijn, H. & Tamminga, E. P. (1984) Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds. *Journal of Physiology* 351:217–50. [JBJS]
Coppola, D. & Purves, D. (1996) The extraordinarily rapid disappearance of entopic images. *Proceedings of the National Academy of Sciences USA* 93(15):8001–8004. [rRN]
Coren, S. & Girgus, J. S. (1973) Visual spatial illusions: Many explanations. *Science* 179(72):503–504. [rRN]
Craver, C. F. & Darden, L. (2001) Discovering mechanisms in neurobiology: The Case of spatial memory. In: *Theory and method in neuroscience*, ed. P.K. Machamer, R. Grush & P. McLaughlin, pp. 112–37. University of Pittsburgh Press. [GT]
Crawford, J. D., Medendorp, W. P. & Marotta, J. J. (2004) Spatial transformations for eye-hand coordination. *Journal of Neurophysiology* 92(1):10–19. [DW]
Crick, F. & Koch, C. (1995) Are we aware of neural activity in primary visual cortex? *Nature* 375(6527):121–23. [ZL, aRN]
Cudeiro, J. & Sillito, A. M. (2006) Looking back: Corticothalamic feedback and early visual processing. *Trends in Neuroscience* 29(6):298–306. [aRN]
Cummins, R. (1983) *The nature of psychological explanation*. MIT Press/Bradford Books. [GT]
Cunningham, D. W., Billock, V. A. & Tsou, B. H. (2001a) Sensorimotor adaptation to violations of temporal contiguity. *Psychological Science* 12:532–35. [DWC, DME]
Cunningham, D. W., Chatziasros, A., von der Heyde, M. & Bülthof, H.H. (2001b) Driving in the future: Temporal visuomotor adaptation and generalization. *Journal of Vision* 1(2):88–98. [DWC]
Cynader, M. & Berman, N. (1972) Receptive-field organization of monkey superior colliculus. *Journal of Neurophysiology* 35(2):187–201. [aRN]
Davidson, D. (1970) Mental events. In: *The nature of mind*, ed. D. Rosenthal. Oxford University Press. [aRN]
De Valois, R. L. & Cottaris, N. P. (1998) Inputs to directionally selective simple cells in macaque striate cortex. *Proceedings of the National Academy of Sciences USA* 95(24):14488–93. [aRN]
De Valois, R. L. & De Valois, K. K. (1991) Vernier acuity with stationary moving Gabors. *Vision Research* 31(9):1619–26. [JB, arRN, GP, DW]
Dean, P., Redgrave, P. & Westby, G. W. (1989) Event or emergency? Two response systems in the mammalian superior colliculus. *Trends in Neuroscience* 12(4):137–47. [aRN]
Dennett, D. C. & Kinsbourne, M. (1992) Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences* 15:183–247. [aRN]
Desimone, R. (1998) Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 353(1373):1245–55. [aRN]
Desimone, R. & Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18:193–222. [aRN]
DeYoe, E. A. & Van Essen, D. C. (1988) Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience* 11(5):219–26. [aRN]
DiCarlo, J. J. & Maunsell, J. H. (2005) Using neuronal latency to determine sensory-motor processing pathways in reaction time tasks. *Journal of Neurophysiology* 93(5):2974–86. [aRN]
Diedrichsen, J., Verstynen, T., Hon, A., Lehman, S. L. & Ivry, R. B. (2003) Anticipatory adjustments in the unloading task: Is an efference copy necessary for learning? *Experimental Brain Research* 148(2):272–76. [aRN]
Dowling, J. E. (1979) Information processing by local circuits: the vertebrate retina as a model system. In: *The neurosciences: Fourth study program*, ed. F. O. Schmitt & F. G. Worden. MIT Press. [aRN]
Dreher, B., Fukada, Y. & Rodieck, R. W. (1976) Identification, classification and anatomical segregation of cells with X-like and Y-like properties in the lateral geniculate nucleus of old-world primates. *Journal of Physiology* 258(2):433–52. [aRN]
Dubois, D. M. (2003) Mathematical foundations of discrete and functional systems with strong and weak anticipations. In: *Lecture notes in computer science*, *2684*, pp. 110–32. Springer. [NS]
DuBois, R. M. & Cohen, M. S. (2000) Spatiotopic organization in human superior colliculus observed with fMRI. *Neuroimage* 12(1):63–70. [aRN]
Duhamel, J.-R., Colby, C. L. & Goldberg, M. E. (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255:90–92. [aRN, GP]
Durant, S. & Johnston, A. (2004) Temporal dependence of local motion induced shifts in perceived position. *Vision Research* 44(4):357–66. [DW]
Eagleman, D. M. (2001) Visual illusions and neurobiology. *Nature Reviews Neuroscience* 2(12):920–26. [DME]
Eagleman, D. M. & Sejnowski, T. J. (2000) Motion integration and postdiction in visual awareness. *Science* 287(5460):2036–38. [DME, arRN, JBJS, MZ]
(2007) Motion signals bias localization judgments: a unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of Vision* 7(4):3, 1–12. [DME, PDH, rRN, DW]
Engel, A. K., Fries, P. & Singer, W. (2001) Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience* 2(10):704–16. [aRN]
Enns, J. T. & Oriet, C. (2004) Perceptual asynchrony: Modularity of consciousness or object updating? [Abstract]. *Journal of Vision* 4(8):27, 27a. Available at: http://journalofvision.org/4/8/27/. [MVCB]
Erlhagen, W. (2003) Internal models for visual perception. *Biological Cybernetics* 88(5):409–17. [MVCB, arRN]
Eskandar, E. N. & Assad, J. A. (1999) Dissociation of visual, motor and predictive signals in parietal cortex during visual guidance. *Nature Neuroscience* 2(1):88–93.
Fajen, B. R. (2007) Rapid recalibration based on optic flow in visually guided action. *Experimental Brain Research* 183:61–74. [DWC]
Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S. & Grafton, S. T. (2008) The angular gyrus computes action awareness representations. *Cerebral Cortex* 18(2):254–61. [JEC]
Fawcett, J. W. & O'Leary, D. M. (1985) The role of electrical activity in the formation of topographical maps in the nervous system. *Trends in Neuroscience* 8:201–206. [rRN]
Fein, A. & Szuts, E. Z. (1982) *Photoreceptors: Their role in vision*. Cambridge University Press. [aRN]
Felleman, D. J. & Van Essen, D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1(1):1–47. [aRN]
Ferrera, V. P. & Barborica, A. (2006) A flashing line can warp your mind. *Neuron* 49(3):327–29. [MZ]
Franz, V. H., Gegenfurtner, K. R., Bulthoff, H. H. & Fahle, M. (2000) Grasping visual illusions: no evidence for a dissociation between perception and action. *Psychological Science* 11(1):20–25. [aRN]
Freyd, J. J. (1987) Dynamic mental representations. *Psychological Review* 94(4):427–38. [DK, aRN]
Fröhlich, F. W. (1923) Über die Messung der Empfindungszeit. [On the measurement of sensation time]. *Zeitschrift für Sinnesphysiologie* 54:58–78. [DK]
Fu, Y. X., Shen, Y. & Dan, Y. (2001) Motion-induced perceptual extrapolation of blurred visual targets. *Journal of Neuroscience* 21(20):RC172. [aRN]
Fu, Y. X., Shen, Y., Gao, H. & Dan, Y. (2004) Asymmetry in visual cortical circuits underlying motion-induced perceptual mislocalization, *Journal of Neuroscience* 24:2165–71. [JB, BRS]
Fujisaki, W., Shimojo, S., Kashino, M. & Nishida, S. (2004) Recalibration of audiovisual simultaneity. *Nature Neuroscience* 7:773–78. [DWC]
Fukushima, K., Yamanobe, T., Shinmei, Y. & Fukushima, J. (2002) Predictive responses of periarcuate pursuit neurons to visual target motion. *Experimental Brain Research* 145(1):104–20. [rRN, GP]
Gawne, T. J., Kjaer, T. W. & Richmond, B. J. (1996) Latency: Another potential code for feature binding in striate cortex. *Journal of Neurophysiology* 76(2):1356–60. [DME]
Gegenfurtner, K. (1999) Neurobiology. The eyes have it! *Nature* 398(6725):291–92. [aRN]
Georgopoulos, A. P. (1986) On reaching. *Annual Review of Neuroscience* 9:147–70. [aRN]
Ghez, C. & Krakauer, J. (2000) The organization of movement. In: *Principles of neural science*, ed. E. R. Kandel, J. H. Schwartz & T. M. Jessell. McGraw Hill.



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Gibson, E. J. (1994) Has psychology a future? *Psychological Science* 5:69–76. [JBW]
Gibson, J. J. (1961) Ecological optics. *Vision Research* 1:253–62. [aRN]
(1979) *The ecological approach to visual perception.* Houghton Mifflin. [JBW]
Glennan, S. (2005) Modeling mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36:443–64. [GT]
Goel, A., Jiang, B., Xu, L. W., Song, L., Kirkwood, A. & Lee, H. K. (2006) Cross-modal regulation of synaptic AMPA receptors in primary sensory cortices by visual experience. *Nature Neuroscience* 9(8):1001–1003. [aRN]
Gomi, H., Abekawa, N. & Nishida, S. (2006) Spatiotemporal tuning of rapid interactions between visual-motion analysis and reaching movement. *Journal of Neuroscience* 26(20):5301–308. [DW]
Goodale, M. A. & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neurosciences* 15(1):20–25. [aRN]
Goodale, M. A., Pelisson, D. & Prablanc, C. (1986) Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature* 320(6064):748–50. [BKh, aRN]
Gottlieb, J. P., Bruce, C. J. & MacAvoy, M. G. (1993) Smooth eye movements elicited by microstimulation in the primate frontal eye field. *Journal of Neurophysiology* 69:786–99. [GP]
Gottlieb, J. P., MacAvoy, M. G. & Bruce, C. J. (1994) Neural responses related to smooth pursuit eye movements and their correspondence with electrically elicited slow eye movements in the primate frontal eye field. *Journal of Neurophysiology* 72:1634–53. [GP]
Gottsdanker, R. M. (1952) The accuracy of prediction motion. *Journal of Experimental Psychology* 43:26–36. [GP]
Grant, K. W., van Wassenhove, V. & Poeppel, D. (2004) Detection of auditory (cross-spectral) and auditory-visual (cross-modality) asynchrony. *Speech Communication* 44:43–53. [JEC]
Gregory, R. (1979) *Eye and brain*, Weidenfeld and Nicholson. [rRN]
Grossberg, S. & Mingolla, E. (1985) Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review* 92(2):173–211. [JB]
Grush, R. (2004) The emulator theory of representation: Motor control, imagery and perception. *Behavioral and Brain Sciences* 27:377–442. [CB]
Grzywacz, N. M. & Amthor, F. R. (1993) Facilitation in ON-OFF directionally selective ganglion cells of the rabbit retina. *Journal of Neurophysiology* 69(6):2188–99. [arRN]
Grzywacz, N. M., Watamaniuk, S. N. J. & McKee, S. P. (1995) Temporal coherence theory for the detection and measurement of visual motion. *Vision Research* 35:3183–203. [HO]
Harris, C. S. (1963) Adaptation to displaced vision: Visual, motor, or proprioceptive change? *Science* 140:812–13. [aRN]
(1980) Insight or out of sight? Two examples of perceptual plasticity in the human adult. In: *Visual coding and adaptability*, ed. C. S. Harris. Erlbaum. [aRN]
Hawkins, J. & Blakeslee, S. (2004) *On intelligence*. Owl Press. [DME]
Haykin, S. (1996) *Adaptive filter theory*. Prentice Hall. [DJM]
Hazelhoff, F. & Wiersma, H. (1924) Die Wahrnehmungszeit I. *Zeitschrift für Psychologie* 96:171–88. [WvdG, aRN]
(1925) Die Wahrnehmungszeit II. *Zeitschrift für Psychologie* 97:174–90. [WvdG]
He, S., Cavanagh, P. & Intriligator, J. (1996) Attentional resolution and the locus of visual awareness. *Nature* 383(6598):334–37. [ZL, aRN]
Held, R. & Freedman, S. J. (1963) Plasticity in human sensorimotor control. *Science* 142:455–62. [aRN]
Henriques, D. Y., Klier, E. M., Smith, M. A., Lowy, D. & Crawford, J. D. (1998) Gaze-centered remapping of remembered visual space in an open-loop pointing task. *Journal of Neuroscience* 18(4):1583–94. [DW]
Henry, G. H., Goodwin, A. W. & Bishop, P. O. (1978) Spatial summation of responses in receptive fields of single cells in cat striate cortex. *Experimental Brain Research* 32(2):245–66. [BRS]
Hess, C. V. (1904) Untersuchungen über den Erregungsvorgang in Sehorgan bei Kurz- und bei länger dauernder Reizung. *Pflügers Archiv für die gesammte Physiologie des Menschen und Thiere* 101:226–62. [DW]
Hess, E. H. (1956) Space perception in the chick. *Scientific American* 195:71–80. [aRN]
Holy, T. E. (2007) A public confession: The retina trumpets its failed predictions. *Neuron* 55(6):831–32. [rRN]
Houtkamp, R., Spekreijse, H. & Roelfsema, P. R. (2003) A gradual spread of attention during mental curve tracing. *Perception and Psychophysics* 65:1136–44. [MVCB]
Hubbard, T. L. (2005) Representational momentum and related displacement in spatial memory: A review of the findings. *Psychonomic Bulletin & Review* 12(5):822–51. [DK]
Hubel, D. H. & Livingstone, M. S. (1987) Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience* 7:3378–415. [HO]
Hubel, D. H. & Wiesel, T. N. (1959) Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology* 148:574–91. [rRN]
(1961) Integrative action in the cat's lateral geniculate body. *Journal of Physiology* 155:385–98. [PDH]
(1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* 160:106–54. [PDH, aRN]
Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. & Bullier, J. (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394(6695):784–87. [aRN]
Jacobs, D. M. & Michaels, C. F. (2006) Lateral interception I: Operative optical variables, attunement, and calibration. *Journal of Experimental Psychology: Human Perception and Performance* 32(2):443–58. [NB]
Jagadeesh, B., Wheat, H. S., Kontsevich, L. L., Tyler, C. W. & Ferster, D. (1997) Direction selectivity of synaptic potentials in simple cells of the cat visual cortex. *Journal of Neurophysiology* 78(5):2772–89. [JB]
Jancke, D., Erlhagen, W., Dinse, H. R., Akhavan, A. C., Giese, M., Steinhage, A. & Schoner, G. (1999) Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *Journal of Neuroscience* 19(20):9016–28. [aRN]
Jeannerod, M., Kennedy, H. & Magnin, M. (1979) Corollary discharge: Its possible implications in visual and oculomotor interactions. *Neuropsychologia* 17(2):241–58. [aRN]
Jiang, Y., Zhou, K. & He, S. (2007) Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience* 10(5):657–62. [ZL]
Johansson, R. S. & Westling, G. (1988) Programmed and triggered actions to rapid load changes during precision grip. *Experimental Brain Research* 71(1):72–86. [aRN]
Kafalıgönül, H., Patel, S. S., Öğmen, H., Bedell, H. E. & Purushothaman, G. (in press) Perceptual asynchronies and the dual-channel differential latency hypothesis. In: *Space and time in perception and action*, ed. R. Nijhawan & B. Khurana. Cambridge University Press. [HO]
Kanai, R., Sheth, B. R. & Shimojo, S. (2004) Stopping the motion and sleuthing the flash-lag effect: spatial uncertainty is the key to perceptual mislocalization. *Vision Research* 44(22):2605–19. [MVCB, aRN, BRS]
Kandel, E. R. & Wurtz, R. H. (2000) Constructing the visual image. In: *Principles of neural science*, ed. E. R. Kandel, J. H. Schwartz & T. M. Jessell. McGraw Hill. [arRN]
Kaplan, E. & Shapley, R. M. (1982) X and Y cells in the lateral geniculate nucleus of macaque monkeys. *Journal of Physiology* 330:125–43. [aRN]
Kawato, M. (1999) Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology* 9(6):718–27. [CB, aRN]
Kawato, M., Furukawa, K. & Suzuki, R. (1987) A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics* 57(3):169–85. [aRN]
Keating, E. G. (1991) Frontal eye field lesions impair predictive and visually guided pursuit eye movements. *Experimental Brain Research* 86:311–23. [GP]
Kerzel, D. (2003) Mental extrapolation of target position is strongest with weak motion signals and motor responses. *Vision Research* 43(25):2623–35. [DK]
Kerzel, D. & Gegenfurtner, K. R. (2003) Neuronal processing delays are compensated in the sensorimotor branch of the visual system. *Current Biology* 13(22):1975–78. [DK, GP]
Keysers, C. & Perrett, D. I. (2002) Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences* 6(3):120–25. [aRN]
Khurana, B., Carter, R. M., Watanabe, K. & Nijhawan, R. (2006) Flash-lag chimeras: The role of perceived alignment in the composite face effect. *Vision Research* 46(17):2757–72. [BKh, aRN]
Khurana, B. & Nijhawan, R. (1995) Extrapolation or attention shift? Reply to Baldo and Klein. *Nature* 378:565–66. [MVCB, DME, arRN]
Khurana, B., Watanabe, K. & Nijhawan, R. (2000) The role of attention in motion extrapolation: Are moving objects "corrected" or flashed objects attentionally delayed? *Perception* 29(6):675–92. [arRN]
Kirschfeld, K. (1983) Are photoreceptors optimal? *Trends in Neurosciences* 6:97–101. [aRN]
Kirschfeld, K. & Kammer, T. (1999) The Fröhlich effect: A consequence of the interaction of visual focal attention and metacontrast. *Vision Research* 39(22):3702–709. [aRN]
Koenderink, J. J. (1984a) The concept of local sign. In: *Limits in perception*, ed. A. J. van Doorn, W.A. van de Grind & J. J. Koenderink, pp. 495–547. VNU Science Press. [WvdG]
(1984b) Simultaneous order in nervous nets from a functional standpoint. *Biological Cybernetics* 50:35–41. [WvdG]
Koken, P. W. & Erkelens, C. J. (1992) Influences of hand movements on eye movements in tracking tasks in man. *Experimental Brain Research* 88:657–64. [JBJS]
Komatsu, H. & Wurtz, R. H. (1988) Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *Journal of Neurophysiology* 60:580–603. [GP]
Kosslyn, S. M. & Sussman, A. (1994) Roles of imagery in perception: Or, there is no such thing as immaculate perception. In: *The cognitive neurosciences*, ed. M. S. Gazzaniga. MIT Press. [rRN]



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Kosslyn, S. M., Thompson, W. L., Kim, I. J. & Alpert, N. M. (1995) Topographical representations of mental images in primary visual cortex. *Nature* 378(6556):496–98. [rRN]

Krauzlis, R. J. (2004) Recasting the Smooth Pursuit Eye Movement System. *Journal of Neurophysiology* 91(2):591–603. [JPM]

Krekelberg, B. & Albright, T. D. (2005) Motion mechanisms in macaque MT. *Journal of Neurophysiology* 93(5):2908–21. [BKr]

Krekelberg, B., Dannenberg, S., Hoffmann, K. P., Bremmer, F. & Ross, J. (2003) Neural correlates of implied motion. *Nature* 424(6949):674–77. [BKr]

Krekelberg, B. & Lappe, M. (1999) Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research* 39:2669–79. [GP]

(2001) Neuronal latencies and the position of moving objects. *Trends in Neurosciences* 24:335–39. [ZL, arRN, BRS, MZ]

Krekelberg, B., van Wezel, R. J. & Albright, T. D. (2006) Adaptation in macaque MT reduces perceived speed and improves speed discrimination. *Journal of Neurophysiology* 95(1):255–70. [BKr]

Lacquaniti, F. & Maioli, C. (1989) The role of preparation in tuning anticipatory and reflex responses during catching. *Journal of Neuroscience* 9(1):134–48. [aRN, MZ]

Lamarre, Y., Busby, L. & Spidalieri, G. (1983) Fast ballistic arm movements triggered by visual, auditory, and somesthetic stimuli in the monkey. I. Activity of precentral cortical neurons. *Journal of Neurophysiology* 50(6):1343–58. [aRN]

Lamme, V. A. & Roelfsema, P. R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience* 23(11):571–z79. [ZL, aRN]

Land, M. F. & Furneaux, S. (1997) The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London B* 352:1231–39. [GP]

Land, M. F. & McLeod, P. (2000) From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience* 3(12):1340–45. [JEC, aRN, GP, JBJS]

Lankheet, M. J. M. & van de Grind, W. (in press) Simultaneity versus asynchrony of visual motion and luminance changes. In: *Space and time in perception and action*, ed. R. Nijhawan & B. Khurana, Ch. 18. Cambridge University Press. [WvdG]

Le Runigo, C., Benguigui, N. & Bardy, B. G. (2005) Perception-action coupling and expertise in interceptive action. *Human Movement Science* 24:429–45. [NB]

Lee, C., Rohrer, W. H. & Sparks, D. L. (1988) Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332(6162):357–60. [aRN]

Lee, D. N. (1976) A theory of visual control of braking based on information about time-to-collision. *Perception* 5(4):437–59. [aRN, MZ]

Lee, D. N. & Reddish, P. E. (1981) Plummeting gannets: A paradigm of ecological optics. *Nature* 293:293–94. [aRN]

Lee, D. N., Young, D. S., Reddish, P., Lough, S. & Clayton, T. (1983) Visual timing in hitting an accelerating ball. *Quarterly Journal of Experimental Psychology* 35:333–46. [NB]

Lee, J., Williford, T. & Maunsell, J. H. (2007) Spatial attention and the latency of neuronal responses in macaque area V4. *Journal of Neuroscience* 27(36):9632–37. [DME]

Leon, M. I. & Shadlen, M. N. (2003) Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38(2):317–27. [MZ]

Levitt, J. B. & Lund, J. S. (1997) Contrast dependence of contextual effects in primate visual cortex. *Nature* 387(6628):73–76. [BRS]

Liberman, A. M. & Mattingly, I. G. (1985) The motor theory of speech perception revised. *Cognition* 21(1):1–36. [aRN]

Lin, Z. & He, S. (in press) Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology* [ZL]

Lisberger, S. G. & Movshon, J. A. (1999) Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. *Journal of Neuroscience* 19:2224–46. [GP, MZ]

Liu, J. & Newsome, W. T. (2005) Correlation between speed perception and neural activity in the middle temporal visual area. *Journal of Neuroscience* 25:711–22. [GP]

Livingstone, M. S. (1998) Mechanisms of direction selectivity in macaque V1. *Neuron* 20:509–26. [JB]

Lobjois, R., Benguigui, N. & Bertsch J. (2006) The effect of aging and tennis playing on coincidence-timing accuracy. *Journal of Aging and Physical Activity* 14(1):75–98. [NB, JEC]

MacAvoy, M. G., Gottlieb, J. P. & Bruce, C. J. (1991) Smooth pursuit eye movement representation in the primate frontal eye field. *Cerebral Cortex* 1:95–102. [GP]

Machamer, P., Darden, L. & Craver, C. F. (2000) Thinking about mechanisms. *Philosophy of Science* 67:1–25. [GT]

Macknik, S. L., Martinez-Conde, S. & Haglund, M. M. (2000) The role of spatiotemporal edges in visibility and visual masking. *Proceedings of the National Academy of Sciences USA* 97(13):7556–60. [aRN]

Maimon, G. & Assad, J. A. (2006) A cognitive signal for the proactive timing of action in macaque LIP. *Nature Neuroscience* 9(7):948–55. [MZ]

Marr, D. (1982) *Vision*. W. H. Freeman. [aRN]

Mateeff, S. & Hohnsbein, J. (1988) Perceptual latencies are shorter for motion towards the fovea than for motion away. *Vision Research* 28(6):711–19. [aRN]

Mates, J., Muller, U., Radil, T. & Poppel, E. (1994) Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience* 6:332–40. [rRN]

Matin, L., Boff, K. & Pola, J. (1976) Vernier offset produced by rotary target motion. *Perception & Psychophysics* 20(2):138–42. [DW]

Maunsell, J. H., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E. & Noerager, B. D. (1999) Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual Neuroscience* 16(1):1–14. [DME]

Maunsell, J. H. & Gibson, J. R. (1992) Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology* 68(4):1332–44. [aRN]

Maus, G. W. & Nijhawan, R. (2006) Forward displacements of fading objects in motion: The role of transient signals in perceiving position. *Vision Research* 46(26):4375–81. [PDH, DK, arRN, HO, GP]

(in press) Going going gone: Localizing abrupt offsets of moving objects. *Journal of Experimental Psychology: Human Perception and Performance*. [rRN]

Mayo, J. P. & Sommer, M. A. (submitted) Neuronal adaptation due to sequential visual stimulation in the frontal eye field. [JPM]

McGraw, P. V., Walsh, V. & Barrett, B. T. (2004) Motion-sensitive neurones in V5/MT modulate perceived spatial position. *Current Biology* 14(12):1090–1093. [JB]

McIntyre, J., Zago, M., Berthoz, A. & Lacquaniti, F. (2001) Does the brain model Newton's laws? *Nature Neuroscience* 4:693–94. [MZ]

McLeod, P. (1987) Visual reaction time and high-speed ball games. *Perception* 16 49–59. [NB]

Mehta, B. & Schaal, S. (2002) Forward models in visuomotor control. *Journal of Neurophysiology* 88(2):942–53. [aRN]

Meister, M., Wong, R. O., Baylor, D. A. & Shatz, C. J. (1991) Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. *Science* 252(5008):939–43. [rRN]

Merchant, H., Battaglia-Mayer, A. & Georgopoulos, A. P. (2004) Neural responses during interception of real and apparent circularly moving stimuli in motor cortex and area 7a. *Cerebral Cortex* 14:314–31. [MZ]

Merfeld, D. M., Zupan, L. & Peterka, R. J. (1999) Humans use internal models to estimate gravity and linear acceleration. *Nature* 398(6728):615–18. [aRN]

Metzger, W. (1932) Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychologische Forschung* 16:176–200. [aRN, WvdG]

Meyer, D. E., Osman, A. M., Irwin, D. E. & Yantis, S. (1988) Modern mental chronometry. *Biological Psychology* 26(1–3):3–67. [aRN]

Miall R. C. & Jackson J. K. (2006) Adaptation to visual feedback delays in manual tracking: evidence against the Smith Predictor model of human visually guided action. *Experimental Brain Research* 172:77–84. [DWC]

Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action*. Oxford University Press. [aRN]

Mishkin, M. & Ungerleider, L. G. (1983) Object vision and spatial vision: Two cortical pathways. *Trends in Neuroscience* 6:414–35. [aRN]

Morgan, M. J. & Thompson, P. (1975) Apparent motion and the Pulfrich effect. *Perception* 4(1):3–18. [aRN]

Mountcastle, V. B., Motter, B. C., Steinmetz, M. A. & Duffy, C. J. (1984) Looking and seeing: the visual functions of the parietal lobe. In: *Dynamic aspects of neocortical function*, ed. G. M. Edelman, W. E. Gall & W. M. Cowan, pp. 159–93. Wiley. [rRN]

Mrotek, L. A. & Soechting, J. F. (2007) Target interception: Hand-eye coordination and strategies. *Journal of Neuroscience* 27(27):7297–309. [JBJS]

Murakami, I. (2001) A flash-lag effect in random motion *Vision Research* 41:3101–19. [HO]

Müsseler, J. & Kerzel, D. (2004) The trial context determines adjusted localization of stimuli: Reconciling the Fröhlich and onset repulsion effects. *Vision Research* 44(19):2201–206. [DK]

Müsseler, J. & Prinz, W. (1996) Action planning during the presentation of stimulus sequences: effects of compatible and incompatible stimuli. *Psychological Research* 59(1):48–63. [aRN]

Müsseler, J., Stork, S. & Kerzel, D. (in press) Localising the onset of moving stimuli by pointing or relative judgment: Variations in the size of the Fröhlich effect. *Vision Research*. [DK]

Nakamura, K. & Colby, C. L. (2002) Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences of the USA* 99:4026–31. [GP]



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Nakamura, K., Matsumoto, K., Mikami, A. & Kubota, K. (1994) Visual response properties of single neurons in the temporal pole of behaving monkeys. *Journal of Neurophysiology* 71(3):1206–21. [aRN]

Nakayama, K. (1985) Biological image motion processing: a review. *Vision Research* 25(5):625–60. [rRN]

Namba, J. & Baldo, M. V. C. (2004) The modulation of the flash-lag effect by voluntary attention. *Perception* 34:621–31. [MVCB, BRS]

Navarra, J., Soto-Faraco, S. & Spence, C. (2007) Adaptation to audiotactile asynchrony. *Neuroscience Letters* 413:72–76. [DWC]

Neisser, U. (1976) *Cognition and reality: Principles and implications of cognitive psychology*. W.H. Freeman. [rRN]

Neuenschwander, S., Castelo-Branco, M., Baron, J. & Singer, W. (2002) Feed-forward synchronization: propagation of temporal patterns along the retinothalamocortical pathway. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 357(1428):1869–76. [rRN]

Newsome, W. T. & Paré, E. B. (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience* 8:2201–11. [GP]

Newsome, W. T., Wurtz, R. H. & Komatsu, H. (1988) Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *Journal of Neurophysiology* 60:604–20. [GP]

Nichols, M. J. & Newsome, W. T. (2002) Middle temporal visual area microstimulation influences veridical judgments of motion direction. *Journal of Neuroscience* 22:9530–40. [GP]

Nijhawan, R. (1992) Misalignment of contours through the interaction of the apparent and real motion systems. *Investigative Ophthalmology and Visual Science* 33(Suppl. 4):1415. [DME, aRN]

(1994) Motion extrapolation in catching. *Nature* 370(6487):256–57. [JB, MVCB, DME, BKh, arRN, HO, GP]

(1997) Visual decomposition of colour through motion extrapolation. *Nature* 386(6620):66–69. [MVCB, arRN]

(2001) The flash-lag phenomenon: object-motion and eye-movements. *Perception* 30:263–82. [arRN, GP]

(2002) Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences* 6:387–93. [arRN]

Nijhawan, R. & Khurana, B. (2002) Motion, space and mental imagery. *Behavioral and Brain Sciences* 25:203–204. [rRN]

Nijhawan, R. & Kirschfeld, K. (2003) Analogous mechanisms compensate for neural delays in the sensory and the motor pathways: Evidence from motor flash-lag. *Current Biology* 13(9):749–53. [aRN]

Nishida, S. & Johnston, A. (1999) Influence of motion signals on the perceived position of spatial pattern. *Nature* 397(6720):610–12. [DW]

Noguchi, Y. & Kakigi, R. (2008) Knowledge-based correction of flash-lag illusion. *Journal of Cognitive Neuroscience* 20:1–13. [BKh]

Nunes, G. (2003) Comment on "Eyesight and the solar Wein peak," by James M, Overduin. *American Journal of Physics* 71:519. [rRN]

Öğmen, H., Patel, S. S., Bedell, H. E. & Camuz, K. (2004) Differential latencies and the dynamics of the position computation process for moving targets, assessed with the flash-lag effect. *Vision Research* 44:2109–28. [HO, GP, DW]

O'Regan, J. K. & Noe, A. (2001) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24(5):939–73. [JBJS]

Palmer, S. E. (1999) *Vision science: Photons to phenomenology.* MIT Press. [PDH]

Parkinson, J. & Khurana, B. (2007) Temporal order of strokes primes letter recognition. *Quarterly Journal of Experimental Psychology* 60:1265–74. [aRN]

Patel, S. S., Öğmen, H., Bedell H. E. & Sampath V. (2000) Flash-lag effect: Differential latency, not postdiction. *Science* 290:1051. [HO, GP, BRS]

Petersen, S. E., Baker, J. F. & Allman, J. M. (1985) Direction-specific adaptation in area MT of the owl monkey. *Brain Research* 346(1):146–50. [BKr]

Poritsky, R. (1969) Two- and three-dimensional ultrastructure of boutons and glial cells on the motoneuronal surface in the cat spinal cord. *Journal of Comparative Neurology* 135(4):423–52. [aRN]

Port, N. L., Kruse, W., Lee, D. & Georgopoulos, A. P. (2001) Motor cortical activity during interception of moving targets. *Journal of Cognitive Neuroscience* 13(3):306–18. [aRN]

Price, N. S., Ibbotson, M. R., Ono, S. & Mustari, M. J. (2005) Rapid processing of retinal slip during saccades in macaque area MT. *Journal of Neurophysiology* 94(1):235–46. [BKr]

Price, N. S., Ono, S., Mustari, M. J. & Ibbotson, M. R. (2005) Comparing acceleration and speed tuning in macaque MT: Physiology and modeling. *Journal of Neurophysiology* 94:3451–64. [MZ]

Purushothaman, G. & Bradley, D. C. (2005) Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience* 8:99–106. [GP]

Purushothaman, G., Patel, S. S., Bedell, H. E. & Öğmen, H. (1998) Moving ahead through differential visual latency. *Nature* 396(6710):424. [arRN, HO, GP, BRS, DW]

Raiguel, S. E., Lagae, L., Gulyas, B. & Orban, G. A. (1989) Response latencies of visual cells in macaque areas V1, V2 and V5. *Brain Research* 493(1):155–59. [aRN]

Ramachandran, V. S. & Anstis, S. M. (1990) Illusory displacement of equiluminous kinetic edges. *Perception* 19(5):611–16. [JB, arRN, GP, DW]

Ramachandran, V. S., Rao, V. M. & Vidyasagar, T. R. (1974) Sharpness constancy during movement perception. *Perception* 3(1):97–98. [aRN]

Ratliff, F. (1965) *Mach bands: Quantitative studies on neural networks in the retina*. Holden-Day. [rRN]

Ratliff, F. & Hartline, H. K. (1959) The responses of limulus optic nerve fibers to patterns of illumination on the receptor mosaic. *Journal of General Physiology* 42(6):1241–55. [aRN]

Reed, E. S. (1996) *Encountering the world: Toward an ecological psychology.* Oxford University Press. [JBW]

Regan, D. (1992) Visual judgements and misjudgements in cricket, and the art of flight. *Perception* 21(1):91–115. [aRN]

Regan, D. & Beverley, K. I. (1984) Figure-ground segregation by motion contrast and by luminance contrast. *Journal of the Optical Society of America A* 1:433–42. [DW]

Reichardt, W. (1961) Autocorrelation, A principle for the evaluation of sensory information by the central nervous system. In: *Sensory communication*, ed. W. A. Rosenblith, pp. 303–17. Wiley. [JB]

Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (1997) The space around us. *Science* 277(5323):190–91. [aRN]

Robinson, D. L. & Kertzman, C. (1995) Covert orienting of attention in macaques: III. Contributions of the superior colliculus. *Journal of Neurophysiology* 74(2):713–21. [JPM]

Roelfsema, P. R., Lamme, V. A. F. & Spekreijse, H. (2000) The implementation of visual routines. *Vision Research* 40:1385–1411. [MVCB]

Rojas-Anya, H., Thirkettle, M. & Nijhawan, R. (2005) Flash-lag anisotropy for movement in three domains. *Perception* 34:219–20. [aRN]

Roufs, J. A. J. (1963) Perception lag as a function of stimulus luminance. *Vision Research* 3:81–91. [aRN]

Saijo, N., Murakami, I., Nishida, S. & Gomi, H. (2005) Large-field visual motion directly induces an involuntary rapid manual following response. *Journal of Neuroscience* 25(20):4941–51. [DW]

Salzman, C. D., Britten, K. H. & Newsome, W. T. (1990) Cortical microstimulation influences perceptual judgments of motion direction. *Nature* 346:174–77. [GP]

Salzman, C. D., Murasugi, C. M., Britten, K. H. & Newsome, W. T. (1992) Microstimulation in visual area MT: Effects on direction discrimination performance. *Journal of Neuroscience* 12:2331–55. [GP]

Sarich, D., Chappell, M. & Burgess, C. (2007) Dividing attention in the flash-lag illusion. *Vision Research* 47:544–47. [MVCB]

Sarlegna, F. R., Gauthier, G. M., Bourdin, C., Vercher, J. L. & Blouin, J. (2006) Internally driven control of reaching movements: A study on a proprioceptively deafferented subject. *Brain Research Bulletin* 69(4):404–15. [rRN]

Sarnat, H. B. & Netsky, M. G. (1981) *Evolution of the nervous system*. Oxford Uinversity Press. [aRN]

Schall, J. D. (2004) Selection of targets for saccadic eye movements. In: *Visual neurosciences*, ed. L. M. Chalupa & J. S. Werner. MIT Press. [GP]

Schiller, P. H. (1984) The superior colliculus and visual function. In: *Handbook of physiology: The nervous system. Neurophysiology*, ed. I. Darien-Smith. American Physiological Society. [aRN]

Schiller, P. H. & Malpeli, J. G. (1978) Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *Journal of Neurophysiology* 41(3):788–97. [aRN]

Schlack, A. & Albright, T. (2007) Remembering visual motion: Neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron* 53(6):881–90. [BKr]

Schlack, A., Krekelberg, B. & Albright, T. D. (2007) Recent history of stimulus speeds affects the speed tuning of neurons in area MT. *Journal of Neuroscience* 27(41):11009–18. [MZ]

Schlag, J. & Schlag-Rey, M. (2002) Through the eye slowly: delays and localization errors in the visual system. *Nature Reviews Neuroscience* 3:191–200. [aRN]

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(6):3272–78. [aRN]

Schoppmann, A. & Hoffmann, K. P. (1976) Continuous mapping of direction selectivity in the cat's visual cortex. *Neuroscience Letters* 2:177–81. [BKr]

Schwartz, G., Harris, R., Shrom, D. & Berry, M. J. (2007a) Detection and prediction of periodic patterns by the retina. *Nature Neuroscience* 10(5):552–54. [JPM]

Schwartz, G., Taylor, S., Fisher, C., Harris, R. & Berry, M. J., 2nd. (2007b) Synchronized firing among retinal ganglion cells signals motion reversal. *Neuron* 55(6):958–69. [WvdG, rRN]



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Senot, P., Zago, M., Lacquaniti, F. & McIntyre, J. (2005) Anticipating the effects of gravity when intercepting moving objects: Differentiating up and down based on nonvisual cues. *Journal of Neurophysiology* 94(6):4471–80. [MZ]

Serre, T., Kouh, M., Cadieu, C., Knoblich, U., Kreiman, G. & Poggio, T. (2005) A theory of object recognition: Computations and circuits in the feedforward path of the ventral stream in primate visual cortex. CBCL Paper No. 259/AI Memo No. 2005-036, Massachusetts Institute of Technology, MA: Cambridge. Available at: http://cbcl.mit.edu/projects/cbcl/publications/ai-publications/2005/AIM-2005-036.pdf. [GT]

Sestokas, A. K. & Lehmkuhle, S. (1986) Visual response latency of x- and y- cells in the dorsal lateral geniculate nucleus of the cat. *Vision Research* 26:1041–54. [DWC]

Shapley, R. M. & Victor, J. D. (1978) The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology* 285:275–98. [aRN]

Shaw, R. E. & Turvey, M. T. (1999) Ecological foundations of cognition: II. Degrees of freedom and conserved quantities in animal-environment systems. *Journal of Consciousness Studies* 6:111–23. [JBW]

Shaw, R. E. & Wagman, J. B. (2001) Explanatory burdens and natural law: Invoking a field description of perception-action. *Behavioral and Brain Sciences* 24:905–906. [JBW]

Sheridan, T. B. & Ferrel, W. R. (1963) Remote manipulative control with transmission delay. *Perceptual and Motor Skills* 20:1070–72. [DWC]

Sheth, B. R., Nijhawan, R. & Shimojo, S. (2000) Changing objects lead briefly flashed ones. *Nature Neuroscience* 3(5):489–95. [JB, MVCB, aRN, BRS]

Shi, Z. & Nijhawan, R. (under review) Behavioral significance of motion direction causes anisotropic flash-lag, flash-mislocalization and movement-mislocalization effects. *Journal of Vision*. [rRN]

Sillito, A. M., Jones, H. E., Gerstein, G. L. & West, D. C. (1994) Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature* 369(6480):479–82. [arRN]

Smeets, J. B. J. & Bekkering, H. (2000) Prediction of saccadic amplitude during smooth pursuit eye movements. *Human Movement Science* 19(3):275–95. [JBJS]

Smith, K. U., Wargo, L., Jones, R. & Smith, W. M. (1963) Delayed and space displaced sensory feedback and learning. *Perceptual and Motor Skills* 16:781–96. [DWC]

Smith, W. M., McCrary, J. R. & Smith, K. U. (1962) Delayed visual feedback and behavior. *Science* 132:1013–14. [DWC]

Snowden, R. J. (1998) Shifts in perceived position following adaptation to visual motion. *Current Biology* 8(24):1343–45. [DW]

Snyder, L. (1999) This way up: Illusions and internal models in the vestibular system. *Nature Neuroscience* 2(5):396–98. [arRN]

Somers, D. C., Todorov, E. V., Siapas, A. G., Toth, L. J., Kim, D. S. & Sur, M. (1998) A local circuit approach to understanding integration of long-range inputs in primary visual cortex. *Cerebral Cortex* 8(3):204–17. [BRS]

Sommer, M. A. & Wurtz, R. H. (2002) A pathway in primate brain for internal monitoring of movements. *Science* 296(5572):1480–82. [JPM]

(2006) Influence of the thalamus on spatial visual processing in frontal cortex. *Nature* 444(7117):374–77. [JPM]

Sparks, D. L. & Jay, M. F. (1986) The functional organization of the primate superior colliculus: a motor perspective. *Progress in Brain Research* 64:235–41. [aRN]

Sparks, D. L., Lee, C. & Rohrer, W. H. (1990) Population coding of the direction, amplitude, and velocity of saccadic eye movements by neurons in the superior colliculus. *Cold Spring Harbor Symposium on Quantitative Biology* 55:805–11. [aRN]

Sperry, R. W. (1950) Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology* 43:482–89. [aRN]

(1952) Neurology and the mind-brain problem. *American Scientist* 40:291–312. [aRN]

Stemmler, M., Usher, M. & Niebur, E. (1995) Lateral interactions in primary visual cortex: A model bridging physiology and psychophysics. *Science* 269(5232):1877–80. [BRS]

Stepp, N. & Turvey, M. T. (2007) Strong anticipation, weak anticipation, and ecological psychology. Presented at the 14th International Conference on Perception & Action, Yokohama, Japan, July 2007. [NS]

Sterzer, P. & Kleinschmidt, A. (2007) A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences USA* 104(1):323–28.

Stetson, C., Cui, X., Montague, P. R. & Eagleman D. M. (2006) Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron* 51:651–59. [DWC, DME]

Stoerig, P. & Cowey, A. (1997) Blindsight in man and monkey. *Brain* 120 (Part 3):535–59. [aRN]

Stratton, G. M. (1896) Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review* 3:611–17. [aRN]

Stuart, A., Kalinowski, J., Rastatter, M. P. & Lynch, K. (2002) Effect of delayed auditory feedback on normal speakers at two speech rates. *Journal of the Acoustical Society of America* 111:2237–41. [JEC]

Stürmer, B., Aschersleben, G. & Prinz, W. (2000) Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology: Human Perception and Performance* 26(6):1746–59. [aRN]

Suddendorf, T. & Corballis, M. C. (2007) Mental time travel across the disciplines: The future looks bright. *Behavioral and Brain Sciences* 30(3):335–45. [DME]

Sugita, Y. & Suzuki, Y. (2003) Audiovisual perception: Implicit estimation of sound-arrival time. *Nature* 421(6926):911. [rRN]

Sun, H. & Frost, B. J. (1998) Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons. *Nature Neuroscience* 1(4):296–303. [aRN]

Sundberg, K. A., Fallah, M. & Reynolds, J. H. (2006) A motion-dependent distortion of retinotopy in area V4. *Neuron* 49(3):447–57. [JB, aRN, MZ]

Taira, M., Mine, S., Georgopoulos, A. P., Murata, A. & Sakata, H. (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research* 83(1):29–36. [aRN]

Tanaka, K. & Fukushima, K. (1998) Neuronal responses related to smooth pursuit eye movements in the periarcuate cortical area of monkeys. *Journal of Neurophysiology* 80:28–47. [GP]

Taube, J. S., Muller, R. U. & Ranck, J. B., Jr. (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *Journal of Neuroscience* 10(2): 420–35. [rRN]

Tessier-Lavigne, M. (2000) Visual processing by the retina. In: *Principles of neural science*, ed. E. R. Kandel, J. H. Schwartz & T. M. Jessell. McGraw Hill. [aRN]

Thiel, A., Greschner, M., Eurich, C. W., Ammermuller, J. & Kretzberg, J. (2007) Contribution of individual retinal ganglion cell responses to velocity and acceleration encoding. *Journal of Neurophysiology* 98(4):2285–96. [MZ]

Thier, P. & Ilg, U. J. (2005) The neural basis of smooth-pursuit eye movements. *Current Opinion in Neurobiology* 15(6):645–52. [aRN, JBJS]

Tootell, R. B., Silverman, M. S., Switkes, E. & De Valois, R. L. (1982) Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science* 218(4575):902–904. [aRN]

Treisman, A. (1996) The binding problem. *Current Opinions in Neurobiology* 6(2):171–78. [ZL]

Tresilian, J. R. (1993) Four questions of time to contact: A critical examination of research on interceptive timing. *Perception* 22(6):653–80. [NB, JEC, arRN]

(1999) Visually timed action: time-out for "tau"? *Trends in Cognitive Sciences* 3(8):301–10. [aRN]

Tresilian, J. R., Plooy, A. & Carroll, T. J. (2004) Constraints on the spatiotemporal accuracy of interceptive action: Effects of target size on hitting a moving target. *Experimental Brain Research* 155:509–26. [NB]

Treue, S. (2003) Climbing the cortical ladder from sensation to perception. *Trends in Cognitive Sciences* 7(11):469–71. [aRN]

Turvey, M. T. (1992) Affordances and prospective control: An outline of the ontology. *Ecological Psychology* 4:173–87. [JBW]

(2004) Impredicativity, dynamics, and the perception-action divide. In: *Coordination dynamics: Issues and trends. Vol. 1. Applied complex systems*, ed. V. K. Jirsa & J. A. S. Kelso, pp. 1–20. Springer Verlag. [JBW]

Turvey, M. T. & Shaw, R. E. (1999) Ecological foundations of cognition: I. Symmetry and specificity of animal-environment systems. *Journal of Consciousness Studies* 6:95–110. [JBW]

van de Grind, W. (2002) Physical, neural, and mental timing. *Consciousness and Cognition* 11(2):241–64; discussion 308–13. [aRN, WvdG]

van de Grind, W. A. (2006) Representing times of the past, present and future in the brain. In: *Timing the future. A case for time-based prospective memory*, ed. J. Glicksohn & M. S. Myslobodski. World Scientific Publishing Company. [WvdG]

van den Berg, A. V. (1988) Human smooth pursuit during transient perturbations of predictable and unpredictable target movement. *Experimental Brain Research* 72:95–108. [JBJS]

van Santen, J. P. & Sperling, G. (1985) Elaborated Reichardt detectors. *Journal of the Optical Society of America* 2:300–21. [JB]

Vatakis, A., Navarra, J., Soto-Faraco, S. & Spence, C. (2007) Temporal recalibration during asynchronous audiovisual speech perception. *Experimental Brain Research* 181:173–81. [DWC, DJM]

von Holst, E. & Mittelstaedt, H. (1950) Das Reafferenzprinzip. *Naturwissenschaften* 37:464–76. [aRN]

Voss, H. U. (2000) Anticipating chaotic synchronization. *Physical Review E* 61:5115–19. [NS]

Vroomen, J., Keetels, M., de Gelder, B. & Bertelson, P. (2004) Recalibration of temporal order perception by exposure to audio–visual asynchrony. *Cognitive Brain Research* 22:32–35. [DWC]

Wagman, J. B. & Carello, C. (2001) Affordances and inertial constraints on tool use. *Ecological Psychology* 13:173–95. [JBW]

(2003) Haptically creating affordances: The user-tool interface. *Journal of Experimental Psychology: Applied* 9:175–86. [JBW]



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press

Wagman, J. B. & Malek, E. A. (in press) Perception of affordances for walking under a barrier from distal and proximal points of observation. *Ecological Psychology*. [JBW]
Wagner, H. (1982) Flow-field variables trigger landing in flies. *Nature* 297:147–48. [aRN]
Wald, G. (1968) The molecular basis of visual excitation. *Nature* 219(5156):800–807. [aRN]
Walls, G. L. (1942) *The vertebrate eye and its adaptive radiation*. Cranbrook Press. [aRN]
Wang, Y. & Frost, B. J. (1992) Time to collision is signalled by neurons in the nucleus rotundus of pigeons. *Nature* 356(6366):236–38. [aRN]
Warren, R. M. & Warren, R. P. (1968) *Helmholtz on perception: Its physiology and development*. Wiley. [aRN]
Warren, W. H. (1988) Actions mode and laws of control for the visual guidance of action. In: *Complex movement behavior: The motor-action controversy*, ed. O. G. Meijer & K. Roth, pp. 339–80. North-Holland. [NB]
Watanabe, K., Nijhawan, R. & Shimojo, S. (2002) Shifts in perceived position of flashed stimuli by illusory object motion. *Vision Research* 42(24):2645–50. [DW]
Watanabe, K., Sato, T. R. & Shimojo, S. (2003) Perceived shifts of flashed stimuli by visible and invisible object motion. *Perception* 32(5):545–59. [DW]
Watts, R. G. & Bahill, A. T. (1990) *Keep your eye on the ball: The science and folklore of baseball*. W. H. Freeman. [JEC]
Weiskrantz, L. (1996) Blindsight revisited. *Current Opinion in Neurobiology* 6(2):215–20. [aRN]
Weiss, Y. & Adelson, E. H. (2000) Adventures with gelatinous ellipses – constraints on models of human motion analysis. *Perception* 29:543–66. [JB]
Welch, R. B. (1978) *Perceptual modification: Adapting to altered sensory environments*. Academic Press. [DWC]
Westheimer, G. & McKee, S. P. (1977) Perception of temporal order in adjacent visual stimuli. *Vision Research* 17(8):887–92. [aRN]
Wexler, M. & Klam, F. (2001) Movement prediction and movement production. *Journal of Experimental Psychology: Human Perception and Performance* 27(1):48–64. [rRN]
Whitaker, D., McGraw, P. V. & Pearson, S. (1999) Non-veridical size perception of expanding and contracting objects. *Vision Research* 39(18):2999–3009. [DW]
Whitaker, D., Perarson, S., McGraw, P. V. & Banford, M. (1998) Keeping a step ahead of moving objects. *Investigative Ophthalmology and Visual Science* 39(Suppl):S1078. [aRN]
Whiting, H. T. A., Gill, E. B. & Stephenson, J. M. (1970). Critical time intervals for taking in flight information in a ball-catching task. *Ergonomics* 13:265–72. [NB]
Whitney, D. (2002) The influence of visual motion on perceived position. *Trends in Cognitive Sciences* 6(5):211–16. [DME, rRN, DW]
Whitney, D. & Cavanagh, P. (2000) Motion distorts visual space: shifting the perceived position of remote stationary objects. *Nature Neuroscience* 3:954–59. [JB, rRN, DW]
(2003) Motion adaptation shifts apparent position without the motion aftereffect. *Perception & Psychophysics* 65(7):1011–18. [DW]
Whitney, D., Ellison, A., Rice, N. J., Arnold, D., Goodale, M., Walsh, V. & Milner, D. (2007) Visually guided reaching depends on motion area MT+. *Cerebral Cortex* 17(11):2644–49. [DW]
Whitney, D., Goltz, H. C., Thomas, C. G., Gati, J. S., Menon, R. S. & Goodale, M. A. (2003) Flexible retinotopy: Motion-dependent position coding in the visual cortex. *Science* 302(5646):878–81. [JB]
Whitney, D. & Goodale, M. A. (2005) Visual motion due to eye movements helps guide the hand. *Experimental Brain Research* 162(3):394–400. [DW]
Whitney, D. & Murakami, I. (1998) Latency difference, not spatial extrapolation. *Nature Neuroscience* 1(8):656–57. [DME, arRN, JBJS, DW]
Whitney, D., Westwood, D. A. & Goodale, M. A. (2003) The influence of visual motion on fast reaching movements to a stationary object. *Nature* 423(6942):869–73. [DW]
Wickelgren, L. W. (1969) The ocular response of human newborns to intermittent visual movement. *Journal of Experimental Child Psychology* 8(3):469–82. [rRN]
Wiesel, T. N. & Hubel, D. H. (1963) Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology* 26:1003–17. [rRN]
Williams, H. & Nottebohm, F. (1985) Auditory responses in avian vocal motor neurons: a motor theory for song perception in birds. *Science* 229(4710):279–82. [aRN]
Williams, J. M. & Lit, A. (1983) Luminance-dependent visual latency for the Hess effect, the Pulfrich effect, and simple reaction time. *Vision Research* 23(2):171–79. [aRN, HO]
Williams, Z. M., Elfar, J. C., Eskandar, E. N., Toth, L. J. & Assad, J. A. (2003) Parietal activity and the perceived direction of ambiguous apparent motion. *Nature Neuroscience* 6(6):616–23. [aRN]
Wilson, J. A. & Anstis, S. M. (1969) Visual delay as a function of luminance. *American Journal of Psychology* 82:350–58. [HO]
Wilson, M. & Knoblich, G. (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131(3):460–73. [arRN]
Witten, I. B., Bergan, J. F. & Knudsen, E. I. (2006) Dynamic shifts in the owl's auditory space map predict moving sound location. *Nature Neuroscience* 9(11):1439–45. [aRN]
Wolpert, D. M. & Flanagan, J. R. (2001) Motor prediction. *Current Biology* 11(18):R729–32. [aRN]
Wolpert, D. M., Ghahramani, Z. & Jordan, M. I. (1995) An internal model for sensorimotor integration. *Science* 269(5232):1880–82. [aRN, GT]
Wolpert, D. M., Miall, R. C. & Kawato, M. (1998) Internal models in the cerebellum. *Trends in Cognitive Sciences* 2:338–47. [aRN]
Woodworth, R. S. (1899) The accuracy of voluntary movement. *Psychological Review* 3(2), Whole No. 13. [aRN]
Woodworth, R. S. & Schlosberg, H. (1954) *Experimental psychology*. Methuen. [aRN]
Wurtz, R. H., Richmond, B. J. & Judge, S. J. (1980) Vision during saccadic eye movements. III. Visual interactions in monkey superior colliculus. *Journal of Neurophysiology* 43(4):1168–81. [JPM]
Young, A. W., Hellawell, D. & Hay, D. C. (1987) Configural information in face perception. *Perception* 16:747–59. [BKh]
Zago, M., Bosco, G., Maffei, V., Iosa, M., Ivanenko, Y. P. & Lacquaniti, F. (2004) Internal models of target motion: Expected dynamics overrides measured kinematics in timing manual interceptions. *Journal of Neurophysiology* 91:1620–34. [MZ]
(2005) Fast adaptation of the internal model of gravity for manual interceptions: Evidence for event-dependent learning. *Journal of Neurophysiology* 93:1055–68. [MZ]
Zago, M. & Lacquaniti, F. (2005) Cognitive, perceptual and action-oriented representations of falling objects. *Neuropsychologia* 43:178–88. [MZ]
Zanker, J. M., Quenzer, T. & Fahle, M. (2001) Perceptual deformation induced by visual motion. *Naturwissenschaften* 88(3):129–32. [aRN]



https://doi.org/10.1017/S0140525X08004044 Published online by Cambridge University Press