Viewpoints

Motion Extrapolation in Visual Processing: Lessons from 25 Years of Flash-Lag Debate

Hinze Hogendoorn

Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, 3010 Victoria, Australia

Because of the delays inherent in neural transmission, the brain needs time to process incoming visual information. If these delays were not somehow compensated, we would consistently mislocalize moving objects behind their physical positions. Twenty-five years ago, Nijhawan used a perceptual illusion he called the flash-lag effect (FLE) to argue that the brain's visual system solves this computational challenge by extrapolating the position of moving objects (Nijhawan, 1994). Although motion extrapolation had been proposed a decade earlier (e.g., Finke et al., 1986), the proposal that it caused the FLE and functioned to compensate for computational delays was hotly debated in the years that followed, with several alternative interpretations put forth to explain the effect. Here, I argue, 25 years later, that evidence from behavioral, computational, and particularly recent functional neuroimaging studies converges to support the existence of motion extrapolation mechanisms in the visual system, as well as their causal involvement in the FLE. First, findings that were initially argued to challenge the motion extrapolation model of the FLE have since been explained, and those explanations have been tested and corroborated by more recent findings. Second, motion extrapolation explains the spatial shifts observed in several FLE conditions that cannot be explained by alternative (temporal) models of the FLE. Finally, neural mechanisms that actually perform motion extrapolation have been identified at multiple levels of the visual system, in multiple species, and with multiple different methods. I outline key questions that remain, and discuss possible directions for future research.

Introduction

The brain needs time to process incoming visual information from the eyes. This presents a challenge for the accurate localization of a moving object because the object will continue to move while visual information about its recent trajectory is flowing through the brain's visual system. One way the brain might solve this problem is by motion extrapolation: using the previous trajectory of a moving object to predict the object's position in the present moment. In this way, the brain might compensate for the delays incurred during neural transmission and enable accurate localization of moving objects.

Motion extrapolation was first proposed by Finke et al. (1986) in the context of an effect they named representational momentum (Freyd and Finke, 1984). In their interpretation, visual memories are distorted by motion such that the remembered position of a moving object is shifted along that object's trajectory. Roughly a decade later, Nijhawan (1994) argued for the existence of motion extrapolation in perceptual (as opposed to memory) mechanisms on the basis of a visual illusion, which has become known as the flash-lag effect (FLE). In the FLE, a stationary object is flashed briefly alongside a moving object (Fig. 1).

The result is that the flashed object is perceived in a position behind the moving object. This effect was first reported by Metzger (1932), and independently discovered again by MacKay (1958), who demonstrated, among many other versions of the effect, that the glowing tip of a moving cigarette appears to float ahead of its base when illuminated by a strobe light. However, it was the presentation of the effect in terms of motion extrapolation by Nijhawan (1994) that sparked an explosion of interest in the illusion, and in the neural mechanisms that might underlie it.

Nijhawan (1994) proposed that the FLE results from motion extrapolation mechanisms that serve to compensate for delays incurred during the transmission of visual information from the retina to the visual cortex, a delay on the order of ~70 ms (e.g., Lamme and Roelfsema, 2000). By extrapolating the position of a moving object into the future by the same amount of time as was lost during neural transmission, the visual system would be able to accurately represent the object's real-time position. In turn, this would enable us to accurately interact with a moving object despite the transmission delays inherent in the visual system, such as when catching a ball. Other authors had previously noted the potential utility of extrapolation in guiding motor control (e.g., Finke et al., 1986), but Nijhawan was the first to emphasize extrapolation as a mechanism to compensate for perceptual, rather than motor, delays (but for a discussion of the role of the motor system in such extrapolation, see Kerzel and Gegenfurtner, 2003). In this interpretation of the FLE, because the moving bar is continuously presented along a predictable motion trajectory, its position can be extrapolated into the future. The flashes on either side, in contrast, are stationary and unpredictable, and so cannot be extrapolated. As a result,
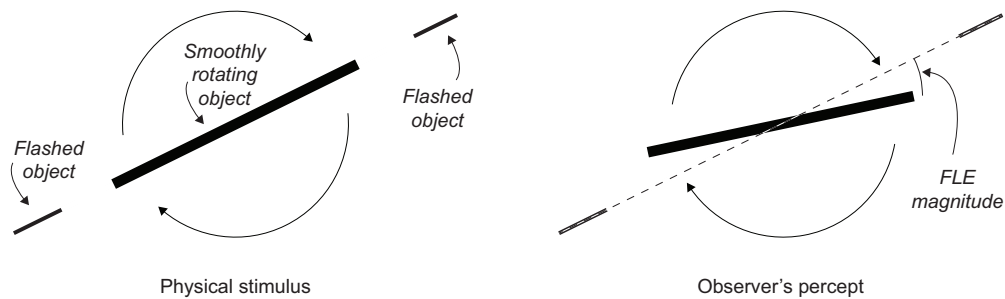
Figure 1. In the FLE, a stationary object is briefly flashed in alignment with a continuously moving object. The result is that the moving object is perceived to lead the flashed object. In the version of the illusion presented by Nijhawan (1994), the stimulus consisted of a single rotating solid rod, presented in darkness with the central section continuously illuminated. The ends of the solid rod were briefly illuminated by a strobe light to generate the stationary flashes. The difference between the perceived relative positions of the stationary and moving objects is taken as the magnitude of the effect.

although the moving bar and the flashed segments are physically aligned, only the position of the moving bar can be extrapolated, and the bar is perceived to lead the two flashed segments.

In this paper, I use the term "motion extrapolation" to refer specifically to neural mechanisms in the visual system that enable it to represent the physical position of a moving object ahead of its most recently detected position. It entails a spatial displacement forward along a trajectory, which has a predictive effect in that it activates neural populations coding for the (likely) future position of the object before those populations would typically receive their input. As a result, those populations effectively represent the position of the object with a lower latency. In this way, extrapolation mechanisms could compensate for (part of) the effects of neural transmission delays on the visual localization of moving objects (depending on the speed of the moving object and the accumulated neural transmission delay by that point in the visual system).

In the years following the publication of Nijhawan's motion extrapolation account of the FLE, numerous authors have investigated a wide range of experimental variables, motivating several alternative accounts of the illusion. The considerable body of empirical work on the FLE has been excellently reviewed by others (Krekelberg and Lappe, 2001; Nijhawan, 2008; Maus et al., 2010; Hubbard, 2014), and Hubbard (2014) provides a comprehensive catalog of empirical variables that have been investigated in the context of the FLE. Furthermore, Hubbard (2013) noted a number of empirical and theoretical parallels between the FLE and the representational momentum literature (for review, see Hubbard, 2014, 2018a).

Here, I do not aim to add an additional empirical review of behavioral studies to the already substantial literature. I also do not want to argue that motion extrapolation (or any other single mechanism) can explain the full breadth of the FLE literature, which is notorious for its contradictions and inconsistencies. Instead, I will briefly summarize the different accounts that have been put forth, and then argue that convergent evidence from behavioral, neurophysiological, computational, and neuroimaging studies together convincingly demonstrates the existence of neural motion extrapolation mechanisms in the visual system, and that the available evidence indicates that these mechanisms play a role (but not necessarily the only role) in motion-position illusions, such as the FLE.

Alternative accounts of the FLE
Following the Nijhawan (1994) study, several research groups presented alternative interpretations of the FLE. The following are six of the most prominent explanations:

Attentional shifting
Baldo and Klein (1995) proposed that the perceived offset between the moving bar and the flash in the FLE was a consequence of the allocation of visual attention to the moving bar. Their interpretation assumes that, when viewing the stimulus, attention is focused on the moving target, and that an attentional shift is initiated when the flash is presented. While the shift is underway, the target continues to move. By the time attention "arrives" at the flashed stimulus and it reaches awareness, the target has already traveled some additional distance, which causes the perceived misalignment. This theory was motivated by a number of experiments that varied the distance between the moving object and flash, which showed that the FLE increased when the moving bar was further away from the flash (and under the implicit assumption that the attentional shift therefore took more time). Consistent with this explanation, when the position of the flash is cued so that attention is directed to it before the flash is presented, the illusion is reduced (Baldo et al., 2002; Namba and Baldo, 2004; Chappell et al., 2006; Sarich et al., 2007). However, the time-cost of an attentional shift has been found to be independent of its spatial extent (Sperling and Weichselgartner, 1995). Additionally, without additional assumptions, the attentional shifting explanation cannot explain why the FLE is observed in flash-initiated displays, in which the flash is presented concurrently with the appearance of the moving object and no attentional shift is required (Khurana et al., 2000).

Differential latencies
A second account of the FLE interprets the effect as evidence for differential latencies for the processing of position and motion signals in the brain. According to this interpretation, the brain processes moving objects more rapidly than static objects, such that, at any given moment, the neural representation of the position of a moving object is more up-to-date than the position of the flash (Whitney and Murakami, 1998; Whitney and Cavanagh, 2000a). The differential latency explanation is consistent with the lack of overshoot when the moving stimulus abruptly reverses direction (Whitney and Murakami, 1998) and with the observation that the FLE scales with the relative luminance of the moving object (Purushothaman et al., 1998). However, studies of temporal order judgment on flashed and moving stimuli have shown that flashed stimuli are processed at least as fast as moving stimuli (Eagleman and Sejnowski, 2000), and may indeed have a processing advantage, rather than a disadvantage (Nijhawan et al., 2004).

Temporal averaging

Krekelberg and Lappe (1999) proposed a third explanation for the FLE, in which the instantaneous position of a moving object is calculated based on its average position over a certain period of time. For the flash, this equates to its physical position, but for the moving object this produces some (possibly weighted) spatial average, notably including positions that the object occupied after the flash. This interpretation is consistent with the presence of the FLE in the flash-initiated condition, as well as with the absence of the FLE in the flash-terminated condition in which the moving object disappears simultaneously with the flash. However, estimates of the width of this window vary widely, from a single point sample taken with some latency after the flash (Brenner and Smeets, 2000) to integration windows spanning ~50 ms (Whitney et al., 2000) and even up to ~500 ms (Krekelberg and Lappe, 1999), although 500 ms is clearly longer than the time taken for the observer to consciously perceive (and act on) the stimulus.

Postdiction

Eagleman and Sejnowski (2000) further developed the temporal averaging account into a theory they dubbed "postdiction." Motivated by the observation that, when the moving object reverses direction concurrently with the presentation of the flash, the FLE is observed in the new direction of the object, postdiction extends the temporal averaging explanation by proposing that the flash resets the integration window, wiping out any earlier position information. Critically, the position information acquired after the flash is then "postdictively" assigned to the moment of the flash, as a counterpart to how earlier information might predictively be assigned to a future moment. It is interesting to note that, although postdiction is often presented as the conceptual opposite of prediction (i.e., motion extrapolation), the two actually invoke similar mechanisms: the perceived position of a moving object at a given time is modulated by motion signals integrated over a certain window. The only difference is the relative temporal position of that window with respect to the flash: whether motion is integrated before or after the flash.

Representational momentum

Hubbard (2013) noted that the perceived forward displacement of the moving object in the FLE (relative to the flash) is conceptually similar to a phenomenon called representational momentum. Hubbard (2013) argued that the FLE could be considered a special case of this phenomenon, in which the judged position of the moving target was compared with the position of a nearby stationary object rather than to the actual position of the target. In representational momentum, the remembered final position of a moving object is displaced forward along the object's trajectory (Freyd and Finke, 1984). Although it seems likely that, in many cases, both effects play a role in (mis)localizing moving objects, there are important differences that indicate that the two are mechanistically dissociable. First, representational momentum is sensitive to higher-order cognitive interpretations in a way that has not been reported for the FLE. For example, representational momentum is smaller when an object moves upward compared with when it "falls" downward, whereas no such directional effects have been reported for the FLE (e.g., Ichikawa and Masakura, 2006, 2010). Similarly, representational momentum is greater for a triangle labeled as a "rocket" than for the same triangle labeled as a "church" (Hubbard, 1995a; Reed and Vinson, 1996), whereas, to my knowledge, no comparable effect has been reported for the FLE. Furthermore, the FLE is maximal for

retinal motion (when observers maintain fixation) (Nijhawan, 2001). Conversely, several studies have reported that representational momentum is reduced or eliminated when observers are required to fixate (e.g., Kerzel, 2000, 2006; De Sá Teixeira et al., 2013). However, the fact that representational momentum is observed for apparent motion stimuli even during fixation (Kerzel, 2003), for auditory targets (Hubbard, 1995b), as well as for static stimuli that only imply motion (Freyd, 1983) means that eye movements cannot be the sole cause of the effect (for a review of this discussion, see Hubbard, 2018b). Although there are therefore important similarities between the FLE and representational momentum, and the two likely often overlap or operate in parallel, these and other differences make it unlikely that the FLE can be fully explained as a subset of representational momentum.

Discrete processing

The sixth explanation of the FLE is a relative newcomer to the debate. Schneider (2018) argues that the FLE and related illusions are the result of the discrete nature of perceptual processing. In this interpretation, the apparently continuous stream of conscious perception is actually the product of a sequence of discrete "perceptual moments," as if the visual world were rhythmically sampled in the same way that a video camera converts a moving image to a series of still frames (VanRullen and Koch, 2003; VanRullen, 2016; Fiebelkorn and Kastner, 2019). Schneider (2018) proposed that information about the position of a moving object that is acquired in a given sample is referred back in time to occupy the "gap" between that sample and the previous one. In this way, the perceived position of a moving object is shifted backward in time. In essence, this explanation is similar to postdiction, with the added constraint that the process operates periodically, referring input back to the most recent gap, rather than referring it back to a fixed temporal extent. Although Schneider (2018) presents an elegant reanalysis of a large existing FLE dataset (Murakami, 2001), and other studies have implicated rhythmic processes in the FLE (Chakravarthi and VanRullen, 2012; Chota and VanRullen, 2019), empirical literature directly testing predictions from this model is not yet available.

Evidence for motion extrapolation in the visual system

Over the past 25 years, an overwhelming amount of evidence has accumulated in support of the idea that the visual system implements motion extrapolation mechanisms that might contribute to compensating for neural delays. Below, I outline what I believe are the key lines of evidence for motion extrapolation.

Why extrapolation does not usually lead to overshoot in the FLE

Possibly the primary challenge to the motion extrapolation model of the FLE has been its apparent inconsistency with the finding that no FLE is observed when the moving object disappears concurrently with the flash (the flash-terminated condition) (Eagleman and Sejnowski, 2000). If the position of the moving object is extrapolated, why then do observers not report perceiving the object beyond the endpoint of its trajectory? It is important to note here that, in most FLE studies, observers do not directly report the position of the moving object in absolute coordinates; rather, they compare the position of the moving object relative to a nearby flash. However, motion itself has been

found to distort nearby visual space (the flash-drag effect) (Whitney and Cavanagh, 2000b), causing the perceived location of the flash to shift in the direction of motion. Similar effects have been observed in typical FLE (Watanabe et al., 2002) and representational momentum displays (Hubbard and Ruppel, 2018). This means that the flash-drag effect effectively reduces the perceived FLE (Eagleman and Sejnowski, 2007; Shi and de'Sperati, 2008). Importantly, although no FLE is typically observed in the flash-terminated condition when observers are asked to localize the final position of the moving object relative to the flash, the perceived final position does indeed overshoot the true endpoint in absolute coordinates (Hubbard, 2008; Shi and de'Sperati, 2008).

In addition, Nijhawan (2008) proposed that visual transients (e.g., the sudden disappearance of the object) trigger a correction-for-extrapolation mechanism that overwrites the extrapolated position of the object with the position information provided by the transient. Although some have felt that this explanation seemed *post hoc*, Maus and Nijhawan (2006) elegantly showed that, when the object gradually fades rather than abruptly disappearing, the forward shift apparent in the FLE is once again observed. Furthermore, a recent study provided a direct demonstration of a correction-for-extrapolation mechanism (Blom et al., 2019) using the related flash-grab effect (Cavanagh and Anstis, 2013). In the flash-grab effect, an object's position is misperceived when it is flashed on a moving background that changes direction. Blom et al. (2019) observed that this mislocalization is a vector sum of the direction of the new background motion (reminiscent of the flash-initiated FLE) and, importantly, the direction opposite the original motion direction. In other words, the transient pulls the perceived position of the object back along the extrapolated motion trajectory, effectively working to correct a prediction that failed to eventuate (Blom et al., 2019). Interestingly, the idea that an early overshoot might be corrected by a later corrective mechanisms is consistent with the time course of representational momentum, which initially increases, peaks at a few hundred milliseconds, and then decreases again (Hubbard, 2018b).

Together, these experiments highlight two reasons why the FLE is typically not observed when the moving object disappears at the same time as the flash. First, even if the perceived endpoint of the object's trajectory is shifted in absolute coordinates (as reported in representational momentum), the object's motion itself subtly distorts the space around it. This causes nearby visual landmarks (e.g., the flash to which the moving object is compared) to shift in the same direction, such that a comparison against those landmarks will reveal little or no shift. Second, the position information provided by the abrupt disappearance of the object works to override the extrapolation signal, thereby correcting for overextrapolation.

Moving objects do overshoot when their disappearance is concealed
The phenomenon of representational momentum reveals that moving objects are typically mislocalized slightly beyond the end of their trajectory (Hubbard, 2005). However, this overshoot usually decays within a few hundred milliseconds after the disappearance of the object (Hubbard, 2018b), suggesting that, when an object does not continue to move along its extrapolated trajectory as expected, the erroneous prediction fades or is overwritten. As a result, we do not typically experience disappearing objects ahead of where they disappear. However, several studies have demonstrated that, in laboratory conditions carefully chosen to

conceal the transient associated with the object's disappearance, observers do indeed perceive objects in regions where they are never presented (or cannot physically be detected). For instance, Maus and Nijhawan (2008) showed observers a monocularly presented object moving into the blind spot, asking them to report where the object disappeared. They found that observers reported seeing the object disappear well within the blind spot, in an area of the retina physically devoid of photoreceptors. In a similar experiment, Shi and Nijhawan (2012) presented blue and green objects moving into the foveal blue-light scotoma, an area of the central visual field that is unable to detect blue light. They observed that blue (but not green) objects disappearing within the scotoma were perceived to move beyond the end of their trajectory. Both studies rendered the transient evoked by the disappearing object invisible, thereby preventing position information from that transient from overriding the extrapolated position. As a result, the observer ultimately perceived the object in the extrapolated position, ahead of where it was last detected on the retina.

It is important to note that none of the alternative explanations of the FLE is able to explain these observations because they invoke temporal, rather than spatial, mechanisms (for discussion, see Hubbard, 2014). Temporal shifts, whether caused by latency differences, averaging, postdiction, or temporal sampling, can shift when an object is perceived to be in a certain position, but cannot explain how a stimulus is perceived in a region of the visual field where no physical stimulus energy can be detected. These findings therefore provide strong evidence for the existence of extrapolation mechanisms (White, 2018).

Neural mechanisms subserving motion extrapolation have been shown to exist
Perhaps the most important argument for the relevance of motion extrapolation in the FLE is that the actual neural mechanisms that perform motion extrapolation have been discovered in multiple species and at multiple levels of the visual system. Berry et al. (1999) demonstrated in salamanders and rabbits that motion extrapolation starts as early as the retina. They showed that presenting a moving bar generated a wave of spiking activity in retinal ganglion cell populations, which traveled near or even in front of the leading edge of the bar, despite the delays incurred during neural transduction. This effectively allows the retina to anticipate the future positions of moving objects. Schwartz et al. (2007) went on to show in salamanders and mice that the retina even implements synchronized bursts of firing when objects deviate from their trajectory, effectively signaling to downstream neurons that the previously extrapolated position was violated. Moving objects are represented ahead of synchronously flashed objects in both cat (Jancke et al., 2004b) and macaque V1 (Subramaniyan et al., 2018). Long-range horizontal connections in cat V1 propagate waves of subthreshold activation, preactivating nearby positions and even inducing illusory motion percepts (Jancke et al., 2004a). Most recently, Benvenuti et al. (2019) showed anticipatory activation along motion trajectories in the primary visual cortex of awake macaques, likewise arguing for an important role for long-range horizontal connections. Finally, by modeling such network connections *in silico*, Jancke and Erlhagen (2010; see also Jancke et al., 2009) showed that, at stimulus speeds relevant to the FLE, local cortical feedback loops not only produce the classical FLE but also predict its behavior in the flash-initiated and flash-terminated conditions (Nijhawan, 2002).

In humans, Hogendoorn and Burkitt (2018) demonstrated comparable effects using an EEG decoding paradigm, showing that the positions of predictably moving objects were decodable from the pattern of scalp activity earlier than the positions of unpredictably moving objects. This revealed that anticipatory mechanisms preactivate the future positions of predictably moving objects, allowing the brain to represent those objects' positions with lower latency. Schellekens et al. (2016) used fMRI to demonstrate that predictive motion extrapolation reduced the BOLD response to predictably moving dots. Their argument was that according to predictive coding models (e.g., Rao and Ballard, 1999), a predicted stimulus should generate a weaker prediction error than an unpredicted stimulus, and thus a weaker BOLD response. By comparing the BOLD response evoked by dots (predictably) moving into a target region to dots (unpredictably) appearing in that region and moving out, they showed that indeed, motion prediction attenuated the neural response to dots moving into a target region. Maus et al. (2013) used fMRI to show that motion shifts position representations in motion area MT, a finding that was recently corroborated by Schneider et al. (2019) using 7T fMRI, showing that population receptive fields are shifted in the direction opposite to visual motion across a range of visual areas.

Most recently, Blom et al. (2020) used an EEG decoding paradigm to study the neural pattern of activation at the end of apparent motion sequences. Consistent with motion extrapolation, they showed that, when apparent motion sequences ended unexpectedly, the pattern of EEG activity revealed that the neural representation of the expected (but not presented) subsequent stimulus was nevertheless activated. In other words, classification analyses showed that motion extrapolation alone (in the absence of sensory input at the expected position) was sufficient to activate position representations that would ordinarily be activated by afferent sensory signals. Furthermore, and consistent with the idea that motion extrapolation might be implemented to compensate for neural processing delays (Nijhawan, 1994, 2008), this predictive activation was observed earlier than it would have been on the basis of sensory information alone. This provides direct neural evidence that the visual system extrapolates the position of moving objects to preactivate the anticipated future positions of those objects, and ties in neatly with a recent fMRI finding that similarly showed predictive activation of the expected future positions of a moving object (Ekman et al., 2017). Together, there is a wealth of neural evidence, acquired using a range of techniques in multiple species, demonstrating the existence of neural motion extrapolation mechanisms at multiple locations in the visual system.

Extrapolation starts early and cascades through the visual system
The FLE is undiminished when the flash is presented concurrently with the appearance of the moving object (the flash-initiated condition) (Khurana and Nijhawan, 1995). This observation was initially difficult for the motion extrapolation theory to explain, since there is no motion preceding the flash to generate an extrapolation signal. However, Nijhawan (2008) noted that early neurons in the visual system can calculate the velocity estimate necessary to enable motion extrapolation within <1 ms (Westheimer and McKee, 1977). Because neurophysiological recordings have shown that motion extrapolation already starts in the retina (Berry et al., 1999), this means that a motion extrapolation signal is available to downstream neurons throughout the visual system essentially concurrently with other visual features

extracted from the visual scene. Motion extrapolation therefore does not require a substantial amount of previous motion history (Nijhawan, 2008), explaining why the FLE is observed even when there is no motion signal preceding the flash. Using an EEG decoding paradigm, we observed the same to be true for the flash-grab effect: it is possible to decode the illusory position of the flash already in the first visual information to reach visual cortex (Hogendoorn et al., 2015). Recent fMRI evidence suggests that the effect of this early extrapolation signal is to shift the receptive fields of downstream neurons in the opposite direction, such that they are processing afferent sensory signals as if they originated from the extrapolated position (Harvey and Dumoulin, 2016; Schneider et al., 2019).

Khoei et al. (2017) used computational modeling to show that a neural population code that includes both position and velocity information in this way produces a simulated FLE. Moreover, the FLE generated in their simulations elegantly reproduces observations from human psychophysics experiments, including the presence and absence of the FLE in the flash-initiated and flash-terminated conditions, respectively. It is interesting to note that this computational model closely resembles the descriptive model presented by Eagleman and Sejnowski (2007), which proposed that motion signals bias localization judgments. Although they argue that their model contradicts the motion extrapolation hypothesis, van Heusden et al. (2019) note that their suggestion that "motion biasing will normally push objects closer to their true location in the world ... by a clever method of updating signals that have become stale due to processing time" is actually entirely consistent with motion extrapolation. Indeed, their proposal that velocity signals bias position judgments precisely describes the central premise of the computational model developed by Khoei et al. (2017), as well as of the more general theoretical model proposed recently by Hogendoorn and Burkitt (2019).

Motion extrapolation also explains other perceptual effects
In addition to the FLE, motion extrapolation is also consistent with observations from other perception experiments. For example, we previously used an apparent motion paradigm to show that the extrapolated position of an object in apparent motion generated interference ahead of the object (as measured by reaction time on a detection task) (Hogendoorn et al., 2008). Using a different paradigm, Roach et al. (2011) showed observers patches of moving bars, and asked them to detect sinusoidally modulated targets at the leading and trailing edges of the motion patches. They observed that detection thresholds for these targets depended on the phase alignment between the targets and the motion patch, and that this was observed only at the leading edge of the motion patch. This is consistent with the interpretation that the moving pattern was extrapolated beyond the edge of the patch, and additively combined with new sensory input from that location.

Finally, motion extrapolation is also consistent with a curious feature of the "high-ϕ" illusion, a motion illusion recently reported by Wexler et al. (2013). In this illusion, a rotating random texture (the inducer) is abruptly replaced by a new random texture. In most conditions, the effect of this sudden transition is that the observer perceives the texture to jerk backward. Strikingly, however, when the inducer is presented only very briefly (e.g., ~16 ms), the jump is perceived forward, not backward. Furthermore, when a static object is briefly flashed at the same time as the transition, it is similarly mislocalized either behind or ahead of its position, depending on the inducer

duration (Hogendoorn et al., 2019). The forward shift for brief inducer durations very closely aligns with a model of motion extrapolation whereby afferent position signals are processed together with rapidly extracted velocity signals to shift the perceived position of objects (and textures, in this case) in their direction of motion.

Conclusions and future directions

The rediscovery of the flash-lag effect by Nijhawan (1994), and his proposal that it resulted from motion extrapolation mechanisms in the visual system, sparked a broad scientific debate about the origins of the illusion and what it can tell us about the processing of time, space, and motion in the visual system. Twenty-five years later, I argue that the accumulated evidence from neurophysiological, computational, behavioral, and neuroimaging studies converges to support the existence of motion extrapolation mechanisms in the visual system, as well as their role in generating the FLE.

Important questions remain. For example: how do motion extrapolation mechanisms at different levels in the visual system interact? We recently proposed that a hierarchical neural processing architecture that implements extrapolation at each step would allow the visual system to align its multiple hierarchical processing stages in real time (Hogendoorn and Burkitt, 2019). Without compensation, transmission delays in the visual system progressively accumulate along the visual hierarchy, such that, at any given moment, higher areas represent older information than lower areas. In our proposal, extrapolation mechanisms at each level of processing compensate for the incremental delays incurred at that level. For predictably changing stimuli (such as smooth motion), the result is that different levels of the visual system become aligned in time. Consistent with this hypothesis, early EEG decoding results (Johnson et al., 2019) indicate that successive hierarchical position representations that become activated sequentially for stationary objects are activated almost concurrently for smoothly moving objects moving on predictable trajectories. However, this same mechanism would not be expected to generalize to situations where the motion signal is more abstract, such as for the representational momentum observed for implied motion in frozen action photographs (Freyd, 1983). Whether and how these visual motion extrapolation mechanisms might interact with other instances of extrapolation (e.g., in memory) remains to be investigated.

Another outstanding question is the degree to which the brain might implement similar extrapolation mechanisms for other stimulus dimensions. On the one hand, effects analogous to the FLE have been demonstrated for a range of stimulus dimensions, including size (Cai and Schlag, 2001), luminance (Sheth et al., 2000), color (Au and Watanabe, 2013; but see Arnold et al., 2009), and entropy (Sheth et al., 2000), and even for other sensory modalities, including audition (Hubbard, 1995b; Alais and Burr, 2003), touch (Lihan, 2013; Cellini et al., 2016), and proprioception (Nijhawan and Kirschfeld, 2003). The same is true for representational momentum (for review, see Hubbard, 2005). We have also reported comparable effects in the oculomotor system, with the magnitude of a flash-grab stimulus depending on the latency of saccades that target it (van Heusden et al., 2018). Together, these studies could be taken as evidence for feature- and modality-independent prediction mechanisms. Such feature-independent prediction mechanisms might be mediated by attention, as in the attentional "drag" theory proposed by Callahan-Flintoft et al. (2020). On the other hand,

some mechanisms involved in motion extrapolation are clearly feature- and modality-specific, such as motion extrapolation mechanisms in the retina (Berry et al., 1999; Schwartz et al., 2007) and preactivation of V1 (Ekman et al., 2017). I would argue that the extrapolation of position on the basis of motion information is therefore probably at least partly distinct from the prediction of other stimulus dimensions in that it is enabled by feature-specific, dedicated neural mechanisms in the (early) visual system. The degree to which the prediction of other sensory properties is enabled by shared or analogous mechanisms therefore remains to be elucidated.

Finally, important questions remain regarding the relationship between the neural representations of moving objects and the observer's conscious perception of those objects. For predictably moving objects, motion extrapolation allows an observer to consciously perceive objects closer to their instantaneous position in space by compensating for neural transmission delays (e.g., extrapolation to the present) (White, 2018). But what happens when objects unexpectedly change direction? As outlined above, observers do not report perceiving the object beyond its change point (Eagleman and Sejnowski, 2000). Instead, observers tend to report displacements in the new direction of motion (Hubbard, 2005; Chappell and Hinchy, 2014), although this depends on the observer's prior expectations about the object and its trajectory (Verfaillie and d'Ydewalle, 1991; Reed and Vinson, 1996). Logically, however, the same transmission delay must apply to the visual information coding for the change in direction, such that (in the absence of higher-order expectations) the observer must (briefly) perceive the object moving beyond the end of its original trajectory, even if that percept is later erased from memory. Consistent with this idea, Blom et al. (2020) recently used EEG decoding to demonstrate that, when an object in apparent motion reverses direction, the brain does indeed briefly preactivate neural representations coding for the object in the expected position, until sensory information coding for where the object actually went becomes available. Whether this erroneous activation is sufficient to lead to a (brief) conscious percept remains an interesting open question.

References

Alais D, Burr D (2003) The "flash-lag" effect occurs in audition and cross-modally. Curr Biol 13:59–63.

Arnold DH, Ong Y, Roseboom W (2009) Simple differential latencies modulate, but do not cause the flash-lag effect. J Vis 9:4.1–4.8.

Au RK, Watanabe K (2013) Object motion continuity and the flash-lag effect. Vision Res 92:19–25.

Baldo MV, Klein SA (1995) Extrapolation or attention shift? Nature 378:565–566.

Baldo MV, Kihara AH, Namba J, Klein SA (2002) Evidence for an attentional component of the perceptual misalignment between moving and flashing stimuli. Perception 31:17–30.

Benvenuti G, Chemla S, Boonman A, Perrinet L, Masson G, Chavane F (2019) Anticipatory responses along motion trajectories in awake monkey area V1. bioRxiv 010017. doi: 10.1101/2020.03.26.010017.

Berry MJ, Brivanlou IH, Jordan TA, Meister M (1999) Anticipation of moving stimuli by the retina. Nature 398:334–338.

Blom T, Liang Q, Hogendoorn H (2019) When predictions fail: correction for extrapolation in the flash-grab effect. J Vis 19:3–11.

Blom T, Feuerriegel D, Johnson P, Bode S, Hogendoorn H (2020) Predictions drive neural representations of visual events ahead of incoming sensory information. Proc Natl Acad Sci USA 117:7510–7515.

Brenner E, Smeets JB (2000) Motion extrapolation is not responsible for the flash-lag effect. Vision Res 40:1645–1648.

Cai RH, Schlag J (2001) Asynchronous feature binding and the flash-lag illusion. Invest Ophthalmol Vis Sci 42:S711.

Callahan-Flintoft C, Holcombe AO, Wyble B (2020) A delay in sampling information from temporally autocorrelated visual stimuli. Nat Commun 11:1852.

Cavanagh P, Anstis SM (2013) The flash grab effect. Vision Res 91:8–20.

Cellini C, Scocchia L, Drewing K (2016) The buzz-lag effect. Exp Brain Res 234:2849–2857.

Chakravarthi R, VanRullen R (2012) Conscious updating is a rhythmic process. Proc Natl Acad Sci USA 109:10599–10604.

Chappell M, Hinchy J (2014) Turning the corner with the flash–lag illusion. Vision Res 96:39–44.

Chappell M, Hine TJ, Acworth C, Hardwick DR (2006) Attention 'capture' by the flash-lag flash. Vision Res 46:3205–3213.

Chota S, VanRullen R (2019) Visual entrainment at 10 Hz causes periodic modulation of the flash lag illusion. Front Neurosci 13:232.

De Sá Teixeira NA, Hecht H, Oliveira AM (2013) The representational dynamics of remembered projectile locations. J Exp Psychol Hum Percept Perform 39:1690–1699.

Eagleman DM, Sejnowski TJ (2000) Motion integration and postdiction in visual awareness. Science 287:2036–2038.

Eagleman DM, Sejnowski TJ (2007) Motion signals bias localization judgments: a unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. J Vis 7:3.

Ekman M, Kok P, de Lange FP (2017) Time-compressed preplay of anticipated events in human primary visual cortex. Nat Commun 8:15276.

Fiebelkorn IC, Kastner S (2019) A rhythmic theory of attention. Trends Cogn Sci 23:87–101.

Finke RA, Freyd JJ, Shyi GC (1986) Implied velocity and acceleration induce transformations of visual memory. J Exp Psychol Gen 115:175–188.

Freyd JJ (1983) The mental representation of movement when static stimuli are viewed. Percept Psychophys 33:575–581.

Freyd JJ, Finke RA (1984) Representational momentum. J Exp Psychol Learn Mem Cogn 10:126–132.

Harvey BM, Dumoulin SO (2016) Visual motion transforms visual space representations similarly throughout the human visual hierarchy. Neuroimage 127:173–185.

Hogendoorn H, Burkitt AN (2018) Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. Neuroimage 171:55–61.

Hogendoorn H, Burkitt AN (2019) Predictive coding with neural transmission delays: a real-time temporal alignment hypothesis. eNeuro 6: ENEURO.0412-18.2019.

Hogendoorn H, Carlson TA, Verstraten FA (2008) Interpolation and extrapolation on the path of apparent motion. Vision Res 48:872–881.

Hogendoorn H, Verstraten FAJ, Cavanagh P (2015) Strikingly rapid neural basis of motion-induced position shifts revealed by high temporal-resolution EEG pattern classification. Vision Res 113:1–10.

Hogendoorn H, Davies S, Johnson P (2019) The high-phi illusion has analogous but dissociable effects on perceived position and perceived motion. Proceedings of the Australasian Cognitive Neuroscience Society. Launceston, Australia.

Hubbard TL (1995a) Environmental invariants in the representation of motion: implied dynamics and representational momentum, gravity, friction, and centripetal force. Psychon Bull Rev 2:322–338.

Hubbard TL (1995b) Auditory representational momentum: surface form, direction, and velocity effects. Am J Psychol 108:255.

Hubbard TL (2005) Representational momentum and related displacements in spatial memory: a review of the findings. Psychon Bull Rev 12:822–851.

Hubbard TL (2008) Representational momentum contributes to motion induced mislocalization of stationary objects. Vis cogn 16:44–67.

Hubbard TL (2013) Do the flash-lag effect and representational momentum involve similar extrapolations? Front Psychol 4:1–6.

Hubbard TL (2014) The flash-lag effect and related mislocalizations: findings, properties, and theories. Psychol Bull 140:308–338.

Hubbard TL (2018a) The flash-lag effect. In: Spatial biases in perception and cognition (Hubbard TL, ed), pp 139–155. Cambridge: Cambridge University Press.

Hubbard TL (2018b) Influences on representational momentum. In: Spatial biases in perception and cognition (Hubbard TL, ed), pp 121–138. Cambridge: Cambridge University Press.

Hubbard TL, Ruppel SE (2018) Representational momentum and anisotropies in nearby visual space. Atten Percept Psychophys 80:94–105.

Ichikawa M, Masakura Y (2006) Manual control of the visual stimulus reduces the flash-lag effect. Vision Res 46:2192–2203.

Ichikawa M, Masakura Y (2010) Reduction of the flash-lag effect in terms of active observation. Atten Percept Psychophys 72:1032–1044.

Jancke D, Erlhagen W (2010) Bridging the gap: a model of common neural mechanisms underlying the Fröhlich effect, the flash-lag effect, and the representational momentum effect. In: Space and time in perception and action (Nijhawan R, Khurana B, eds), pp 422–440. Cambridge: Cambridge UP.

Jancke D, Chavane F, Naaman S, Grinvald A (2004a) Imaging cortical correlates of illusion in early visual cortex. Nature 428:423–426.

Jancke D, Erlhagen W, Schöner G, Dinse HR (2004b) Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex. J Physiol 556:971–982.

Jancke D, Chavane F, Grinvald A (2009) Stimulus localization by neuronal populations in early visual cortex: linking functional architecture to perception. In: Dynamics of visual motion processing (Masson G, Ilg U, eds), pp 95–116. New York: Springer.

Johnson P, Blom T, Feuerriegel D, Bode S, Hogendoorn H (2019) The visual system compensates for accumulated neural delays to represent moving objects closer to their real-time positions. Proceedings of the Australasian Cognitive Neuroscience Society. Launceston, Australia.

Kerzel D (2000) Eye movements and visible persistence explain the mislocalization of the final position of a moving target. Vision Res 40:3703–3715.

Kerzel D (2003) Attention maintains mental extrapolation of target position: irrelevant distractors eliminate forward displacement after implied motion. Cognition 88:109–131.

Kerzel D (2006) Why eye movements and perceptual factors have to be controlled in studies on "representational momentum." Psychon Bull Rev 13:166–173.

Kerzel D, Gegenfurtner KR (2003) Neuronal processing delays are compensated in the sensorimotor branch of the visual system. Curr Biol 13:1975–1978.

Khoei MA, Masson GS, Perrinet LU (2017) The flash-lag effect as a motion-based predictive shift. PLoS Comput Biol 13:e1005068.

Khurana B, Nijhawan R (1995) Extrapolation or attention shift? Nature 378:566.

Khurana B, Watanabe K, Nijhawan R (2000) The role of attention in motion extrapolation: are moving objects 'corrected' or flashed objects attentionally delayed? Perception 29:675–692.

Krekelberg B, Lappe M (1999) Temporal recruitment along the trajectory of moving objects and the perception of position. Vision Res 39:2669–2679.

Krekelberg B, Lappe M (2001) Neuronal latencies and the position of moving objects. Trends Neurosci 24:335–339.

Lamme VA, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci 23:571–579.

Lihan C (2013) Tactile flash lag effect: taps with changing intensities lead briefly flashed taps. World Haptics Conference, pp 253–258. IEEE.

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

Maus GW, Nijhawan R (2006) Forward displacements of fading objects in motion: the role of transient signals in perceiving position. Vision Res 46:4375–4381.

Maus GW, Nijhawan R (2008) Motion extrapolation into the blind spot. Psychol Sci 19:1087–1091.

Maus GW, Khurana B, Nijhawan R (2010) History and theory of flash-lag: past, present, and future. In: Space and time in perception and action (Nijhawan R, Khurana B, eds), pp 477–500. Cambridge: Cambridge UP.

Maus GW, Fischer J, Whitney D (2013) Motion-dependent representation of space in area MT$^+$. Neuron 78:554–562.

Metzger W (1932) Versuch einer gemeinsamen Theorie der Phänomene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. Psychol Forsch 16:176–200.

Murakami I (2001) A flash-lag effect in random motion. Vision Res 41:3101–3119.

Namba J, Baldo MV (2004) The modulation of the flash-lag effect by voluntary attention. Perception 33:621–631.

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

Nijhawan R (2001) The flash-lag phenomenon: object motion and eye movements. Perception 30:263–282.

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

Nijhawan R (2008) Visual prediction: psychophysics and neurophysiology of compensation for time delays. Behav Brain Sci 31:179–239.

Nijhawan R, Kirschfeld K (2003) Analogous mechanisms compensate for neural delays in the sensory and the motor pathways. Curr Biol 13:749–753.

Nijhawan R, Watanabe K, Khurana B, Shimojo S (2004) Compensation of neural delays in visual-motor behaviour: no evidence for shorter afferent delays for visual motion. Vis Cogn 11:275–298.

Purushothaman G, Patel SS, Bedell HE, Ogmen H (1998) Moving ahead through differential visual latency. Nature 396:424–424.

Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87.

Reed CL, Vinson NG (1996) Conceptual effects on representational momentum. J Exp Psychol Hum Percept Perform 22:839–850.

Roach NW, McGraw PV, Johnston A (2011) Visual motion induces a forward prediction of spatial pattern. Curr Biol 21:740–745.

Sarich D, Chappell M, Burgess C (2007) Dividing attention in the flash-lag illusion. Vision Res 47:544–547.

Schellekens W, van Wezel RJ, Petridou N, Ramsey NF, Raemaekers M (2016) Predictive coding for motion stimuli in human early visual cortex. Brain Struct Funct 221:879–890.

Schneider KA (2018) The flash-lag, Fröhlich and related motion illusions are natural consequences of discrete sampling in the visual system. Front Psychol 9:1227.

Schneider M, Marquardt I, Sengupta S, Martino F, De Goebel R (2019) Motion displaces population receptive fields in the direction opposite to motion. bioRxiv 759183. doi: 10.1101/759183.

Schwartz G, Taylor S, Fisher C, Harris R, Berry MJ (2007) Synchronized firing among retinal ganglion cells signals motion reversal. Neuron 55:958–969.

Sheth BR, Nijhawan R, Shimojo S (2000) Changing objects lead briefly flashed ones. Nat Neurosci 3:489–495.

Shi Z, de'Sperati C (2008) Motion-induced positional biases in the flash-lag configuration. Cogn Neuropsychol 25:1027–1038.

Shi Z, Nijhawan R (2012) Motion extrapolation in the central fovea. PLoS One 7:e33651.

Sperling G, Weichselgartner E (1995) Episodic theory of the dynamics of spatial attention. Psychol Rev 102:503–532.

Subramaniyan M, Ecker AS, Patel SS, Cotton RJ, Bethge M, Pitkow X, Berens P, Tolias AS (2018) Faster processing of moving compared with flashed bars in awake macaque V1 provides a neural correlate of the flash lag illusion. J Neurophysiol 120:2430–2452.

van Heusden E, Rolfs M, Cavanagh P, Hogendoorn H (2018) Motion extrapolation for eye movements predicts perceived motion-induced position shifts. J Neurosci 38:8243–8250.

van Heusden E, Harris AM, Garrido MI, Hogendoorn H (2019) Predictive coding of visual motion in both monocular and binocular human visual processing. J Vis 19:3.

VanRullen R (2016) Perceptual cycles. Trends Cogn Sci 20:723–735.

VanRullen R, Koch C (2003) Is perception discrete or continuous? Trends Cogn Sci 7:207–213.

Verfaillie K, d'Ydewalle G (1991) Representational momentum and event course anticipation in the perception of implied periodical motions. J Exp Psychol Learn Mem Cogn 17:302–313.

Watanabe K, Nijhawan R, Shimojo S (2002) Shifts in perceived position of flashed stimuli by illusory object motion. Vision Res 42:2645–2650.

Westheimer G, McKee SP (1977) Spatial configurations for visual hyperacuity. Vision Res 17:941–947.

Wexler M, Glennerster A, Cavanagh P, Ito H, Seno T (2013) Default perception of high-speed motion. Proc Natl Acad Sci USA 110:7080–7085.

White PA (2018) Is the perceived present a predictive model of the objective present? Vis Cogn 26:624–654.

Whitney D, Cavanagh P (2000a) The position of moving objects. Science 289:1107.

Whitney D, Cavanagh P (2000b) Motion distorts visual space: shifting the perceived position of remote stationary objects. Nat Neurosci 3:954–959.

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

Whitney D, Cavanagh P, Murakami I (2000) Temporal facilitation for moving stimuli is independent of changes in direction. Vision Res 40:3829–3839.