Is the perceived present a predictive model of the objective present?

Peter A. White

School of Psychology,

Cardiff University,

Tower Building,

Park Place,

Cardiff CF10 3YG,

Wales,

U. K.

email: whitepa@cardiff.ac.uk

Abstract

Processing latencies for coherent, high level percepts in vision are at least 100 ms and possibly as much as 500 ms. Processing latencies are less in other modalities, but still significant. This seems to imply that perception lags behind reality by an amount equal to the processing latency. It has been proposed that the brain can compensate for perceptual processing latencies by using the most recent available information to extrapolate forward, thereby constructing a model of what the world beyond the senses is like now. The present paper reviews several lines of evidence relating to this hypothesis, including the flash-lag effect, motion-induced position shifts, representational momentum, static visual illusions, and motion extrapolation at the retina. There are alternative explanations for most of the results but there are some findings for which no competing explanation has yet been proposed. Collectively, the evidence for extrapolation to the present is suggestive but not yet conclusive. An alternative account of compensation for processing latencies, based on the hypothesis of rapid emergence of percepts, is proposed.

Keywords: perceived present moment; extrapolation to the present; predictive coding; flash-lag effect; perceptual processing latencies.

Word count: Main text: 16,921

Is the perceived present a predictive model of the objective present?

All of us have the feeling that we live in the present moment; that is, in terms of perception, that we perceive the world outside our brains as it is now. I shall hereafter use the term "perceived present" to refer to the collective set of perceptual information that seems to us to represent the outside world as it is now. In fact, operation of peripheral sensors, neural transduction and perceptual processing of input information take time, so it can be argued that our conscious percepts are out of date by an amount equal to the combined latency of those things (hereafter just "processing latency", for convenience).[1] That is, we are in effect perceiving a moment that is in the past by however long it takes for perceptual processing to generate the percept of that moment. An alternative possibility is that the brain uses the processed information, not to represent the recent past, but to predict the present moment. In that case, conscious percepts would be the contents of the predictive model of the present moment, generated from the most recent available information. The aim of this paper is to review the evidence concerning that possibility. I shall start by briefly describing the processing latency problem, and I shall then proceed to a review of several lines of research that have been taken as evidence for the predictive model hypothesis, which I shall call the *extrapolation to the present* hypothesis. At the end I shall propose an alternative account, based on the hypothesis that conscious percepts begin to emerge no more than 100 - 150 ms after stimulus onset (ASO). Several kinds of adjustment to the temporal location of things perceived are mentioned in this review, of which extrapolation to the present is one, so a list is provided in box 1 (see also Bachmann, Breitmeyer, & Öğmen, 2011).

Terms such as "perceived present" and "present moment" are used frequently in this paper. The "present" in perception is a somewhat ambiguous term. At one extreme, some authors have proposed that we exist in a perceived present with a time span of several seconds (Clay, 1882; James, 1890). An example often taken, and apparently dating back to

St. Augustine (Rovelli, 2018), is auditory perception of a melodic phrase (Clay, 1882; Lloyd, 2012; Wittmann, 2011): all of the notes that have just been heard seem to be contained in the present; certainly the current note is perceived in a context of the previous notes making up the phrase of which it is part. At the other extreme, the present moment would be defined in physics in terms of the Planck time, 10^{-44} s, which is far shorter than the time scale of any neurophysiological activity. For present purposes the term is being used to refer to perceptual information on the millisecond time scale: functionally speaking, the informational products of perceptual processing before they are transferred (if they are) to subsequent stores. The exact time scale of this cannot be identified at present and may vary between processes and modalities. Two brief sounds separated by 2 ms can be distinguished as nonsimultaneous in perceptual processing (Heinrich & Schneider, 2006; Wiegrebe & Krumbholz, 1999; Zera & Green, 1993) and it could be argued that only one of those is in the present at any given moment. In other processes and other modalities, stimuli must be separated by 50 - 100 ms before they are perceived as nonsimultaneous (Brown & Sainsbury, 2002; Craig & Busey, 2003; Geffen, Rosa, & Luciano, 2000). It can be said, however, that, for purposes of this paper, the "present" refers to products of perceptual processing that are represented as contemporaneous on the millisecond time scale, and not to longer spans of time.

How long does perceptual processing take?

There is no simple answer to the question. Processing latencies vary between modalities; latencies within modalities vary depending on the specific process; and a coherent perceptual representation depends in part on synchronisation across modalities so that, for example, visual information about lip movements is synchronised with auditory information about speech. Relevant evidence will be briefly reviewed.

In the case of vision, the earliest time of cortical response to a visual stimulus occurs in visual cortex area V1 approximately 40 - 70 ms ASO, and this represents the combined time for photoreceptor response plus response time of ganglion cells and neural transduction to V1 (Gollisch & Meister, 2010; Johnston & Lagnado, 2015; Lamme & Roelfsema, 2000; Martínez, Anllo-Vento, Sereno, Frank, Buxton, Dubowitz, Wong, Hinrichs, Heinze, & Hillyard, 1999; Pitcher, Goldhaber, Duchaine, Walsh, & Kanwisher, 2012; Tapia & Beck, 2014; Vroomen & Keetels, 2010). Audition is faster, with a peak primary cortical evoked potential response latency of about 8 - 15 ms (Lakatos, Pincze, Fu, Javitt, Karmos, & Schroeder, 2005; Liegeois-Chauvel, Musolino, & Chauvel, 1991; Vroomen & Keetels, 2010; Zeki, 2001).

Several authors have distinguished two levels of processing in vision, called level 1 and level 2 or low-level and high-level processing (Battelli, Pascual-Leone, & Cavanagh, 2007; Burr & Thompson, 2011; Cavanagh, 2011; Di Lollo, Enns, & Rensink, 2000; Holcombe, 2009; Itti & Koch, 2001; Rensink, 2000; Seiffert & Cavanagh, 1998; Wutz & Melcher, 2014). Level 1 processing is generally regarded as fast, automatic, non-attentive, bottom-up or feedforward, and local. Level 2 processing tends to be slower, to involve attentive processing, to focus on a more global analysis (such as, at the level of individuated perceptual objects rather than local feature detection), and to involve pre-existing structures, semantic processing, and featural continuity analysis (Holcombe, 2009).

The subjective completeness and coherence of the perceived present implies that it is a product of level 2 processing. Level 2 processing starts about 100 ms ASO (Holcombe, 2009), but it continues for a long time and involves multiple stages (Fahrenfort, Scholte, & Lamme, 2008). Early level 2 processing may involve the construction of perceptual objects but is still dominated by registration of low-level physical features of stimuli (Rossion & Caharel, 2011). Object identification probably begins around 150 - 200 ms ASO, and other kinds of semantic or categorical processing have similar latencies (Keyes, Brady, Reilly, &

Foxe, 2010; Li, Zhong, Chen, & Mo, 2013; Rossion & Caharel, 2011; Thorpe, Fize, & Marlot, 1996; Tortosa, Lupiánez, & Ruz, 2013; van der Lugt, Banfield, Osinsky, & Münte, 2012; van Heijnsbergen, Meeren, Grèzes, & de Gelder, 2007). This includes distinguishing faces from other kinds of stimuli (Rossion & Caharel, 2011), distinguishing between animals and non-animals (Li et al., 2013), and identifying some kinds of emotional facial expression (Tortosa et al., 2013; van Heijnsbergen et al., 2007). Categorisation of gist information about scenes or situations also occurs around 150 ms ASO (Goffaux, Jacques, Mouraux, Oliva, Schyns, & Rossion, 2005; Groen, Ghebreab, Prins, Lamme, & Scholte, 2013; Rousselet, Fabre-Thorpe, & Thorpe, 2002; Sun, Simon-Dack, Gordon, & Teder, 2011; Thorpe et al., 1996).

Those results suggest that around 150 - 200 ms ASO is a critical time for attentive processing and semantic analysis of visual stimuli. But processing does not end there, and several studies have found stimulus discriminations at much longer latencies, including categorisation of objects (animals) around 230 ms ASO (Li et al., 2013), differentiation between fearful and neutral faces around 400 - 600 ms ASO (Kiss & Eimer, 2008), and differentiation between direct and averted gaze around 420 - 580 ms ASO (Itier, Alain, Kovacevic, & McIntosh, 2007). A map of the spatial layout of the visual world is constructed over at least 600 ms and possibly as much as 1,000 ms (Yoshimoto, Uchida-Ota, & Takeuchi, 2014; Zimmermann, Morrone, & Burr, 2014). Temporal integration for both biological and non-biological motion perception can occur on a time scale up to about 3,000 ms (Burr & Santoro, 2001; Neri, Morrone, & Burr, 1998).

In addition, cross-modal synchronisation is required to deal with differences in processing latencies between modalities. As we have already seen, auditory processing is faster than visual processing (Vroomen & Keetels, 2010). For touch, neural conduction delays vary with the distance of the point of stimulation from the brain, and there is a difference of about 30 ms between stimuli at the nose and at the toes (Halliday & Mingay, 1964; Vroomen & Keetels, 2010). In everyday life we are rarely if ever aware of asynchrony in perception between different modalities. Many studies have shown that there are multiple ways in which the brain deals with asynchrony and generates a perceived world in which objectively corresponding events are synchronised despite the problems caused by spatial distance and neural processing latencies (see Vroomen & Keetels, 2010, and Chen & Vroomen, 2013, for reviews). An event in one modality can be registered as synchronous with an event in another modality even if they were objectively separated by 200 - 250 ms (Conrey & Pisoni, 2006; Dixon & Spitz, 1980; Mégevand, Molholm, Nayak, & Foxe, 2013), which suggests a substantial processing latency between onset of the first stimulus and occurrence of a synchronised conscious percept.

A large body of research has attempted to discern the latency of emergence of conscious percepts in ERP waveform differences between visual stimuli that are and are not consciously perceived (see Koivisto & Revonsuo, 2010, for a review). Although definitive conclusions cannot be drawn, most authors agree that the latency to emergence of conscious percepts is not less than 200 ms (Koivisto & Revonsuo, 2010; Pitts, Metzler, & Hillyard, 2014; Railo, Koivisto, & Revonsuo, 2011), and some have argued for a latency of around 300 ms (Dehaene & Changeux, 2011; Del Cul, Baillet, & Dehaene, 2007; Sergent, Baillet, & Dehaene, 2005). A latency of 400 ms has been proposed by Herzog, Kammer, and Scharnowski (2016) on the grounds of evidence that visual feature integration can be affected by transcranial magnetic stimulation applied up to 400 ms after presentation of the stimuli (Scharnowski, Rüter, Jolij, Hermens, Kammer, & Herzog, 2009)

This is only a brief survey of a large body of research, but it serves to make two important points. One is that, on the face of it, perception lags some way behind reality. In the case of vision, some components of perceptual information lag reality by more than half a second, because that is how long it takes for them to be fully processed (e.g. Zimmermann et al., 2014). The other point is that it has not (yet) been possible to identify a single latency that defines the emergence of a conscious percept. The range of processing latencies, even just within vision, is great, and processing latencies also vary substantially between modalities, necessitating further delay while cross-modal synchronisation is effected. Any hypothesis that the perceived present lags behind the objective present must stipulate the latency, but perceptual processing research indicates that there is no single latency, and no obvious location for the perceived present in the stream of perceptual processing.[2]

The research discussed in this section tends to concern presentation of stimuli with abrupt onset that are essentially unpredictable. Processing latencies may be shorter for stimuli that are predictable (Bachmann, 1989; Bachmann, Luiga, Põder, & Kalev, 2003), such as new input information about the motion of an object that has already been perceived as in motion. However, this only adds to the difficulty in identifying a single latency at which a product of perceptual processing emerges. There will be further discussion of this issue after the review of evidence.

If the perceived present lags behind objective reality by a substantial fraction of a second, why are we not aware of this asynchrony, and why does it not cause severe practical problems such as co-ordinating actions on objects? One answer to the latter question is that there are two processing routes in vision, one (dorsal or magnocellular) specialised for processing spatial and motion information and the other (ventral or parvocellular) specialised for processing object and colour information (Goodale & Milner, 1992; Mishkin, Ungerleider, & Macko, 1983). It could be argued that the dorsal stream provides information used in the timing and execution of actions, with extrapolation mechanisms to guide timing, while visual perception of the ball (and everything else in the scene) is a product of processing in the ventral stream that is not tied to action and lags behind the objective present. Dorsal stream processing supports rapid reactions to environmental events: the fastest possible reaction time to a visually perceived event is about 100 - 110 ms (Brenner & Smeets, 1997; Castiello, Paulignon, & Jeannerod, 1991; Lisberger, 2010; Prablanc & Martin, 1992), and there is evidence that co-ordination of action to environment is accomplished with the aid of predictive models of object motion (Brenner & Smeets, 2015; Zago & Lacquaniti, 2005; Zago, McIntyre, Senot, & Lacquaniti, 2009). Thus, well-timed and rapid reactive interventions can be executed before conscious percepts emerge.

That is unsatisfactory, however, because it would seem to generate obvious perceptual asynchrony between action and perception. Consider the problem faced by a sportsperson who has to intercept and act on a ball moving rapidly through the air, as in cricket, baseball, or tennis. Nijhawan (2008) calculated that, with a ballistic projectile moving at 90 mph, if we assume that it takes about 100 ms for a percept of the projectile to be constructed, the perceived location of the projectile lags about 13 feet behind its actual location at any moment (see also Land & Mcleod, 2000).[3] With a processing latency of 400 ms, as proposed by Herzog et al. (2016), the perceived location of the projectile would lag about 50 feet behind its objective location. In badminton, some professionals can hit a shuttlecock at 200 mph. A badminton court is only 44 feet long, but a shuttlecock with an average speed of 200 mph across its path from racquet to racquet (or floor) would cover about 28 feet in 100 ms, and about 100 feet in 400 ms. If it took 400 ms to construct a conscious percept, the visual percept of the shuttlecock would be so far behind its objective location that the recipient would have intercepted it before he or she saw the shuttlecock being hit by their opponent. As a badminton player, I can testify that that degree of temporal dislocation between action and perception does not occur, nor even the amount of temporal dislocation implied by a shorter processing latency and a more realistic estimate of the projectile's speed for a club standard player. Subjectively, sound, vision, and action are synchronised.

One possible explanation for that is that the synchronisation is itself a product of perceptual processing, so that the entire perceptual experience of acting on moving objects (and everything else in perception) is synchronised but running behind objective reality: the action itself is a rapid response and appropriately timed, but perception of the execution of the action, the shuttlecock, the noise of the impact of racquet on shuttlecock, the behaviour of one's opponent, and everything else, is synchronised but delayed. Such perception would have to be entirely dislocated from the perception for action system because it occurs too late to be of any use to action, so under that hypothesis the information used to guide actions on projectiles would not form part of the perceived present at all.

The latency of the perceived present under that hypothesis would be substantial: it would have to be synchronised to the latency of the slowest process, because synchronisation cannot be accomplished until all the perceptual information to be synchronised has been processed, and the duration of the synchronisation process would have to be added to that. The temporal window of integration across auditory and visual input is in the region of 200 - 250 ms (Chen & Vroomen, 2013; Freeman, Ipser, Palmbaha, Paunoiu, Brown, Lambert, Leff, & Driver, 2013; van Wassenhove, Grant, & Poeppel, 2007). That does not include the time taken for the synchronisation process itself, so the minimum latency for the entire perceived present under the hypothesis would be about 250 ms, and probably longer. Therefore, if the perceived present, which encompasses perceptual information about perceptual objects and contexts, action output and perceptual feedback of action and its effects on perceived objects, and cross-modal integration and synchronisation, was all synchronised to the same latency, all of it would be running at least 250 ms behind the objective present and probably more. Meanwhile, the dorsal system is still generating reactions on a time scale of ~100 ms. So the two systems would have to be virtually independent.

The perceived present as a model of the objective present

There is another possibility, however. Perhaps the brain can use out-of-date information to construct an up-to-date model of the world beyond the senses. Thus, although the most recent processed information may be out of date by a substantial fraction of a second, it provides informational support for a process that generates a coherent set of predictions for the state of the world in the present, and that set of predictions is the perceived present. Such a model could not be guaranteed to be accurate but, so long as it was reasonably accurate most of the time, it would have obvious practical advantages over a model that was entirely inaccurate by virtue of being out of date.

Under the hypothesis of extrapolation to the present, different processing latencies for different processes are not a problem because, in principle, they can each generate input into the perceived present at their own rate. The perceived present is an ongoing model of the ouside world at the present moment. It is constructed by extrapolation from the most recent available perceptual information. As the model is a set of predictions, it may be altered only when newly available information effectively disconfirms a prediction in the model (Galletti & Fattori, 2003). In that case the disconfirmatory information forms the basis for a new extrapolation which is then inserted into the model and integrated with what is already there. A mechanism of that kind automatically compensates, not only for processing latencies per se, but for differences in processing latencies between modalities and between different processes within modalities.

Such automatic compensation for processing latencies does not mean that the model of the present is accurate from the start. Suppose that a stimulus comprising a static image of a human face is presented to an observer. Then, as the research reviewed above shows, different features are processed with different latencies. Thus, the model of the present, in respect of the face stimulus, is constructed piece by piece over a few hundred milliseconds and, at any given time during that process, it is liable to be inaccurate in respect of any feature that has not yet been processed. But all the time, the information content of the model, no matter how incomplete, is an extrapolation to the present.

That is more easily understood with reference to stimuli that are changing over time: for the static image of a face, substitute a video of a face that is doing something, such as talking. From the start of the video, the model of the face in the present is constructed over a short period of time, as before, and the content of the model is an extrapolation of the current state of the face based on the available information. The extrapolation is based on available perceptual information and the use of models of facial movement to generate the extrapolation. For example, the movement of the mouth might fit with an existing model of pronunication of a particular syllable, and that model is used to generate an extrapolation to the state of the mouth in the present. The extrapolation process takes time, so from the initiation of the stimulus the accuracy of the extrapolation to the present will improve over some hundreds of milliseconds as incoming information is processed and extrapolations are generated. Once the initial phase of construction of the extrapolation to the present for the face is completed, further adjustments are made as and when newly processed information disconfirms predictions in the extrapolation.

So, for a new and complex stimulus, the extrapolation may be slow to get going but, after a while, it is a synchronised and integrated model updated by adjustments generated by new information that disconfirms components of the existing model. If this does not sound like much of an improvement on perception that is running behind objective reality by a few hundred milliseconds, it is worth noting that sudden, unpredicted appearances of novel complex stimuli are rather more common in the laboratory than they are in the world in general. Most of our sensory environment is predictable with a reasonable level of accuracy on the millisecond time scale, so extrapolation is generally accurate enough to be reliable. As an example, around me as I write this are numerous static objects, such as my desk and the walls of my office. The persistence of these objects and their properties can be extrapolated to the present with little risk of inaccuracy. Many other events involve change, but many of these are also predictable because the change fits a model that can be applied to the data, such as a model of a particular kind of object motion (White, 2012). Of course, some events, such as a sudden noise from outside my office, cannot be predicted: these are processed, entered into and integrated with the perceived present as quickly as possible.

One more issue will be briefly mentioned here and discussed further later in the paper. This is postdiction. Postdiction is the use of new perceptual information to reinterpret the recent history of a stimulus (Shimojo, 2014). If a novel and unpredicted stimulus occurs, it may fit with a model of change from a previous state. In that case, not just the extrapolation to the present but also the recent history leading up to the extrapolation to the present may be adjusted in accordance with the prediction based on the fit between the novel stimulus, the current information, and the model of change. As an example, the observed motion of a previously stationary object may result in a postdictive adjustment to the perceived trajectory of an object that is perceived as making the stationary object move by colliding with it (Kim, Feldman, & Singh, 2013). That is, the pre-collision motion of the causal object is reinterpreted to fit with a model of how the motion direction of the formerly stationary object could have been generated. Once the postdictive reinterpretation has been made, the previous interpretation is effectively obliterated, so from the observer's point of view it will seem as though the adjusted motion percept is the only percept there has ever been. The extrapolation to the present is supported by a postdictively adjusted recent history. Postdiction is a natural complement to extrapolation to the present: if the perceived present is a set of predictions, the recent perceived past can be a set of postdictions, within certain time and processing constraints. Postdiction may be critical to the interpretation of research evidence bearing on the hypothesis of extrapolation to the present, as will be shown later.

The hypothesis of extrapolation to the present implies that postdiction on a short time scale is, in effect, prediction. Suppose, for the sake of simplicity, that the visual processing latency is always 200 ms. Then extrapolation to the present is a prediction forward 200 ms from the most recent available information. Now suppose that the recent history of the stimulus is postdictively reinterpreted on a time scale of 100 ms; that is, the last 100 ms leading up to the current percept is postdictively adjusted. This means, in fact, that the postdictive reinterpretation is also a prediction forward by 100 - 200 ms: the limit of the postdiction may be 100 ms earlier than the extrapolated present but that places it 100 ms to the future from the emergence of the information after 200 ms of processing. So, under the hypothesis of extrapolation to the present, postdiction within the time scale of extrapolation to the present is in fact prediction.

Several sets of research findings are relevent to the hypothesis of extrapolation to the present, and these will be reviewed in subsequent sections. The review begins, however, with a simple but compelling observation.

Review part 1: Gaze direction

Where do sportspeople look when a projectile is approaching them? Coaches advise them to keep their eye on the ball at least until it is just in front of the bat or racquet or hand. Several studies have shown that eye movements do accurately track the current position of the projectile over large parts of its trajectory, in cricket (Land & McLeod, 2000), squash (Hayhoe, McKinney, Chajka, & Pelz, 2012), table tennis (Rodrigues, Vickers, & Williams, 2002), baseball (Bahill & LaRitz, 1984), and catching a ball (Cesqui, Mezzetti, Lacquaniti, & d'Avella, 2002). Players do not track the ball all the time and, in particular, show anticipatory saccades to locations where the ball's behaviour is likely to change unpredictably, as in the bounce of a cricket ball (Land & McLeod, 2000). However, when they do track it, as all the authors cited say, their eye is on the ball, with reasonable accuracy, and gaze direction does not lag behind the current location of the projectile at any time.

The problem with those findings is that the most recent perceptual information the player has about the ball concerns a location several feet behind the location at which their gaze is directed, depending on how fast the ball is moving. Nijhawan and Wu (2009) published a photograph of a professional tennis player taken just before a moving ball was struck. The ball is visible just a foot or so in front of the racquet, and the player's gaze is on the ball. This seems entirely natural. In fact, however, a percept of the ball constructed with a latency of 100 ms or more would not be located there. Nijhawan and Wu added a small circle indicating the most recently available location of the ball, assuming a speed of 60 mph and a processing latency of 100 ms, and it is clearly, not just well behind the real ball, but also well behind the direction of the player's gaze. Information about the real ball in the present is not yet available to the player: if the player has a conscious percept of the ball at the focus of his gaze, it must be an extrapolation to the present.

The player is not always looking at where the ball is: as I have already said, there is evidence for anticipatory saccades to a critical location where the ball is expected to be (Hayhoe et al., 2012; Land & McLeod, 2000). However, if information about the ball is to be acquired, gaze must be directed at where it is now. There is no point gazing at where it was 100 ms ago because it isn't there any more. So there must be some kind of extrapolation of the projectile's trajectory to guide gaze direction when gaze is directed at the projectile. The point, however, is that the percept of the ball appears at (or very near) the focus of the player's gaze. Since the gaze is directed at a predicted location of the ball, the percept of a ball at that predicted location must itself be a prediction. That prediction is an extrapolation to the present.

One possible counterargument is that perceived gaze direction is subject to processing delay as well. In that case, the percept of the ball at the focus of gaze can be explained by the fact that perceived gaze direction and perception of the ball are both subject to the same amount of delay. Let us first examine the facts of the situation. The player's gaze is directed at the current location of the ball: that is clear in the photograph. The player has a percept of the ball at that location: players of ball games would attest that they keep their eyes on the ball (with the exceptions described above), meaning that they see the ball at the focus of their gaze. Processing input information about the ball takes at least 100 ms - for the purposes of the argument, let us assume that it is exactly that. Therefore it will be another 100 ms before information about the ball at the location shown in the photograph generates a percept in the brain of the player. Under the counterargument, it also takes 100 ms to process input information about gaze direction. Therefore the player's percept of their gaze direction is as it actually was when the ball was at the location indicated by the white circle. Perceived gaze direction and perceived location of the ball coincide, and both are running 100 ms behind the actual ball and the actual direction of gaze.

It is first necessary to ask what is meant by "perceived gaze direction". Information that might determine a person's experience of their own gaze direction might be visual, in terms of static features of the visual scene, and also involves feedback from kinaesthetic sensors in the eye movement muscles (Matthews, 1982; McCloskey, 1981). Cortical responses latencies of <10 ms to stretching of an eye muscle have been found (Donaldson, 1979, 2000). That is much faster than vision, where information does not leave the retina until about 70 ms after the stimulus has impinged on it. Processing latency thereafter is not known. However, it is likely to be rapid. The area of the cortex that responds to eye muscle movement appears to be involved in the maintenance of stability in the visual world (Donaldson, 2000). Duhamel, Colby, and Goldberg (1992) found that neurons that should fire about 70 ms after the saccade brings a stimulus into their receptive field actually started firing 80 ms <u>before</u> the start of the saccade. Subsequent research has shown that this presaccadic remapping maintains stability of the visual world across saccades (He, Mo, & Fang, 2017; Rao, Mayo, & Sommer, 2016). It has been argued that this involves a comparator process in which a model of the predicted eye movement is matched against input information of the actual eye movement (Blakemore, Wolpert, & Frith, 2002; Donaldson, 2000; Grush, 2004; Miall & Wolpert, 1996; Sperry, 1950; Von Holst and

Mittelstaedt, 1950). For this to work effectively, processing of eye movement feedback would have to be rapid. The actual processing latency, therefore, is probably much less than the processing latency for visual information. In that case, input information about gaze direction would not lag as far behind real gaze direction as perceptual information about the ball lags behind the actual ball. If that is the case, then the delay in perceived gaze direction would not be sufficient to account for the percept of the ball at the focus of the player's gaze. However, further research would be necessary to establish the amount of disparity. In the absence of that, the fate of the counterargument cannot be decided.

Evidence for the use of a predictive model of object motion as a guide to interception has already been discussed, and prediction helps to account for the accuracy of timing achieved in fast interceptions (Borghuis & Leonardo, 2015; McLeod & Jenkins, 1991; Perrinet, Adams, & Friston, 2014; Regan, 1992, 1997; Zago et al., 2009). However, the extrapolation that guides action is not to the present but to the future, to the spatial location or time at which the interception is predicted to occur. That moment obviously converges on the present moment as the moment of contact approaches. Before contact occurs, however, the extrapolation to the future that guides action must be updated and adjusted when error is detected, and that is an important reason for directing gaze at the ball's expected current location. The ball's present location is judged by extrapolation from the most recent available information about its past motion, and gaze direction is guided by that.

There are, therefore, two extrapolations: one to the present, so that information about the ball's motion at that time can be picked up, and one to the future, to the anticipated time of contact, as a guide to the interceptive action. It is tempting to suggest that extrapolation to the present occurs in the ventral stream and extrapolation to the future anticipated moment of interception occurs in the dorsal stream. However, it is not clear that the two streams are completely independent, nor is it certain that conscious (visual) percepts are confined to the ventral stream (Briscoe & Schwenkler, 2015; Nijhawan, 2008). Therefore I shall say only that extrapolation to the present can be distinguished from extrapolation to the future anticipated moment of interception. Predicting the trajectory or future locations of the ball for purposes of interception, and directing the gaze to a particular predicted location of the ball, are both problematic (Brenner & Smeets, 2015; Perrinet et al., 2014). But the particular problem that is relevant here is that, when the player gazes at that predicted location, there is a conscious visual percept of the ball there, when the most recent available perceptual information concerns a location of the ball several feet away from the direction of gaze. It is hard to understand what that percept could be, other than an extrapolation to the present.

Review part 2: perception of moving objects

The flash-lag effect

Nijhawan (1994) presented a stimulus comprising a rigid rod rotating about its centre. The central third of the rod was under constant illumination. The outer thirds were in darkness except for a 5 ms period of illumination, which is the flash. The outer thirds were perceived as lagging behind the central part of the rod by an amount that increased as speed of rotation increased. This is the flash-lag illusion. Nijhawan argued that the motion of the constantly visible section of the rod is predictable and can therefore be extrapolated to overcome the latency problem. It is less easy to extrapolate from the visual input about the outer thirds, however, because their momentary appearance (and disappearance) is not predicted, so the apparent lag represents the outcome of processing that is not extrapolated. This was interpreted by Nijhawan (1994, 2008) as supporting the hypothesis that visual percepts are extrapolations to the extent that extrapolation from available information is possible. Several other kinds of stimuli have been used in tests of this effect, usually computer-generated (Hubbard, 2014; Krekelberg & Lappe, 2001), but there is something striking and appealing about an illusion that occurs in perception of a solid, rigid object.

The now substantial literature on the flash-lag effect has been reviewed elsewhere (Arstila, 2015; Hubbard, 2014; Krekelberg & Lappe, 2001; Maus, Khurana, & Nijhawan, 2010; Nijhawan, 2008). For present purposes, the issue is whether research on the flash-lag effect has yielded evidence that unambiguously supports the hypothesis of extrapolation to the present. In brief, there are several competing explanatory accounts (Hubbard, 2014) and several findings that the extrapolation hypothesis does not seem to be able to explain (e.g. Purushothaman, Patel, Bedell, & Öğmen, 1998), although some of these have now been shown to be predicted by a Bayesian predictive coding version of the extrapolation to the present hypothesis (Khoei, Masson, & Perrinet, 2017).

For illustrative purposes, one alternative account of the flash-lag effect will be briefly described. This is the postdiction hypothesis developed by Eagleman and Sejnowski (2000). Eagleman and Sejnowski (2000) found that the occurrence of the flash-lag effect depended on the motion of the moving object after the flash but not before. The effect was reversed if the moving object reversed direction after the flash. By varying the duration of motion reversal, they found that the perceived location of the flash was affected by motion up to about 80 ms after the flash. They argued that the flash resets the temporal integration of information about the moving object over time, so that the location of the flash is perceived in relation to the motion of the moving object determined by integration over the 80 ms following the flash. The interpretation is supported by other evidence for postdictive reinterpretation of perceptual information on short time scales (Kolers & von Grünau, 1976; Shimojo, 2014). Postdiction will be further discussed in the commentaries on the evidence below. Other studies have found problems for the postdiction interpretation of the flash-lag effect, however (see Hubbard, 2014), and it is likely that pre-flash motion information is not altogther neglected.

It is likely that several factors and aspects of visual processing are involved in determing the occurrence and extent of the flash-lag effect. Some factors increase lag and others decrease it, to the extent that a reversal of the flash-lag effect can occur under some circumstances (Eagleman & Sejnowski, 2000; Howard, Masom, & Holcombe, 2011). The critical question is whether there are any findings that can (at present, at least) be explained only by the extrapolation hypothesis. The following appear to be the strongest candidates.

Shi and Nijhawan (2012) took advantage of the fact that there are two foveal scotomas, foveal insensitivity to dim light and to blue light. They argued that a blue object moving towards the foveal blue light scotoma and stopping before it gets there should be extrapolated into the area of the scotoma, whereas a green object that stops before it gets there should not, because the fovea is more sensitive to it. A similar argument applies to dim versus bright moving objects. The results of their study supported these predictions. But why does the extrapolation occur for the blue stimulus and not for the green stimulus? The authors argued that the extrapolation would in fact occur for both stimuli, but as soon as the stopping of the object is processed, the extrapolation is suppressed and the recent history of the extrapolation is effectively obliterated by a backward masking effect from the relatively strong signal that the object has stopped. In their experiment, the stop signal appears rapidly for the green object because of foveal sensitivity to green, but less rapidly for the blue object because foveal insensitivity to blue means that the cessation of motion is not picked up so quickly. Therefore, and as they found, extrapolation occurs for the blue object but is suppressed for the green object. The results are difficult for any other account of the flash-lag effect to explain, because the difference in perceived motion is associated purely with a difference in colour of the stimulus, in the case of the blue versus green stimuli, and no other hypothesis has addressed effects of colour.

Nijhawan (1997) presented a moving green bar. A red line was briefly flashed within the bar. Theoretically, this should be perceived as a yellow line superimposed on the green bar. In fact, however, observers reported a red line displaced behind the green bar in relation to the green bar's direction of motion. Nijhawan argued that the displacement occurred because of extrapolation to the present: the motion of the green bar was continuous and predictable, and therefore extrapolated to the present. The flashed red line was not predictable, and therefore was not extrapolated to the present, resulting in displacement behind the green bar as perceived. Although the retinal stimulus was the red line imposed on the green bar, Nijhawan argued that the two were separated cortically so that the local colour of the red line was recovered from the stimulus information. This was, in effect, a side-effect of the displacement due to extrapolation of the green bar to the present. As far as I have been able to discover, no alternative explanation of this percept has been proposed, so it stands as evidence favouring the extrapolation to the present hypothesis (see Cavanagh, 1997; Hubbard, 2014). In his commentary on this research, Cavanagh (1997) argued that extrapolation to the present "can only occur for objects that are undergoing smooth change" (p. 19). That is not correct: it can occur for anything in perception that is predictable, which includes non-changing things, any kind of smooth change, any kind of implied change (e.g. Finke, Freyd, & Shyi, 1986), and any kind of change that conforms to an internal model that can generate predictions, such as a model of acceleration under gravity (Zago et al., 2009) or a model of momentum (Hubbard, 2015). That is to say, it can occur for almost everything in perception. Isolated, unpredicted flashes of light are unusual, not typical of visual experience.

If construction of a model of the objective present moment is a general function of perception, it should be possible to find evidence for illusions similar to the flash-lag illusion in other modalities. Indeed, an auditory version of the flash-lag effect has been reported (Alais & Burr, 2003; Arrighi, Alais, & Burr, 2005), and also a haptic version (Cellini, Scocchia, & Drewing, 2016; Nijhawan & Kirschfeld, 2003). In the study by Cellini et al. (2016), participants moved a finger along a runway while a finger on the other hand was stationary at a different location to provide a reference point. At some point in the motion of the finger a vibrotactile stimulus was applied to the moving finger, and participants judged where the finger was at the time the stimulus was delivered. The moving finger was judged to be ahead of its objective position, which is evidence for what the authors called a "buzz-lag" effect. Thus, the position of the moving finger is extrapolated to the present on the basis of the most recently available information. The tactile buzz is unexpected and so cannot be extrapolated to the present, but is perceived late, in accordance with the processing latency. By the time the processing of the buzz results in a conscious percept, the extrapolated position of the finger has moved on beyond where it was when the vibrotactile stimulus was applied to it. That results in the buzz-lag effect. The importance of these findings is that explanations for the flash-lag effect that refer to specific characteristics of the visual system cannot apply to them. Extrapolation to the present should occur across all modalities and could therefore account for the auditory and haptic versions of the flash-lag effect.

However, it is also possible that there are features of haptic processing that would generate a judgment error that happens to be in the same direction as that of the flash-lag effect, but for different reasons, and other modality-general features could also account for it. Cellini et al. argued that their results could be accounted for by a temporal sampling hypothesis (Brenner & Smeets, 2000). The location of a moving object may be sampled at intervals rather than being continuously monitored. In the flash-lag effect, a sampling process is triggered by the detection of the flash, but there is a latency in the sampling process that is not taken into account when the location estimate is generated. Consistent with this, Cellini et al. found that the size of the buzz-lag effect was greater at greater speed of motion: at faster speeds, the finger moves further between sampling events that occur at fixed intervals, and that would increase the size of the buzz-lag effect as speed increased. Brenner and Smeets (2000) showed that the visual flash-lag effect was significantly reduced if participants were presented with a cue that predicted the occurrence of the flash, but not if the cue was contemporaneous with or subsequent to the flash. The cue would presumably serve to trigger the sampling process, and that would reduce or eliminate the flash-lag effect if the cue preceded the flash. That effect would also be consistent with the hypothesis of extrapolation to the present, however: the cue renders the flash predictable so its occurrence can be predicted in the same manner as an extrapolation to the present. At present, therefore, neither explanation can be ruled out.

If the flash-lag effect represents extrapolation to the present, one would expect the temporal amount of the illusion to match the processing latency which, as we have seen, is a minimum of 100 ms (Holcombe, 2009). The temporal amount of the illusion was estimated at 27 ms by Chakravarthi and VanRullen (2012), and between 45 and 90 ms by other authors (Arnold, Ong, & Roseboom, 2009; Eagleman & Sejnowski, 2000, 2007; van de Grind, 2002). This is not enough to compensate fully for the latency in processing of the flash. It is likely that extrapolation to the present would be more accurate than that, so the discrepancy is a problem. One possibility is that the temporal amount of the illusion is attenuated in part by lag reducing factors that are independent of extrapolation to the present (Howard et al., 2011). Another possible way out, in the form of a hypothesized compensation mechanism, will be discussed later.

Many factors may contribute to the flash-lag effect, to the extent that the opposite of a flash-lag effect can occur under some conditions, and it is very difficult to discriminate between extrapolation to the present and alternative explanations such as temporal sampling (Brenner & Smeets, 2000), postdiction (Eagleman & Sejnowski, 2000), and differential processing latencies for static and moving stimuli (Öğmen, Patel, Bedell, & Camuz, 2004; Purushothaman et al., 1998; Whitney & Murakami, 1998; Whitney, Murakami, & Cavanagh, 2000). At present, the studies by Shi and Nijhawan (2012) and Nijhawan (1997) provide the strongest evidence for extrapolation to the present from the literature on the flash-lag effect because no competing explanations have been proposed, but it is still possible that the results reflect properties of the visual system that have yet to be elucidated.

Motion induced position shifts

The flash-lag effect is one of many phenomena in which perceived locations of objects are affected by motion. For example, a kinetic edge, which is a border defined in terms of motion, such as one group of dots moving uniformly while another group is stationary, is itself perceived as moving when in fact it is stationary (Ramachandran & Anstis, 1990). The same has been found for a stationary aperture containing a drifting Gabor patch (Caniard, Bülthoff, & Thornton, 2015) or in relation to another, vertically aligned Gabor patch (De Valois & De Valois, 1991). The perceived location of a moving object tends to be shifted in its direction of motion (Matin, Boff, & Pola, 1976; Whitney, 2002; Whitney, Wurnitsch, Hontiveros, & Louie, 2008); a transient change in colour of a moving object is perceived as further along the object's trajectory than it was (Cai & Schlag, 2001); the judged final location of a rotating stimulus is affected by rotation of a frame within which the stimulus is presented (Hubbard, 1993, and see next section); and the perceived location of a flash can be shifted in the direction of motion of a nearby stimulus (Whitney & Cavanagh, 2000).

There are many possible explanations for motion induced position shifts, and extrapolation is not applicable to all such shifts. Some of them have, however, been explained as the effect of some kind of visual extrapolation. De Valois and De Valois (1991) proposed that the position shift they found could be explained as compensation for neural delays in the service of accurate interception, although they pointed out that the adjustment was not optimal for this purpose because it varied with retinal eccentricity. Caniard et al. (2015) and Jordan and Hunsinger (2008), the latter a study of representational momentum (see next section), found that errors in perceived location were greater if participants actively controlled the motion of the stimuli than if they passively observed them. This supports an extrapolation hypothesis: extrapolation should be greater in the context of planned action than in the context of passive observation because, in the former case, the extrapolation must be made to the time of the planned action in the future, whereas, in the latter case, extrapolation need only be made to the present moment. This is consistent with the evidence of several studies that motion induced position shifts are, at least in part, phenomena of high-level perceptual processing involving attentive processing (Kohler, Cavanagh, & Tse, 2015; Tse, Whitney, Anstis, & Cavanagh, 2011).

However, as we have already seen, extrapolation to the future in the service of action is different from extrapolation to the present. Thus, if extrapolation to the future is occurring when participants are controlling or expecting to act on the stimuli, that does not imply that extrapolation to the present is occurring when they are not. Caniard et al. (2015) pointed out that the difference between the active and passive viewing conditions could result from operational features of the visual guidance of action (i.e. other than extrapolation to the future), which would of course not apply in the passive viewing condition. They also suggested that active involvement could shift the frame of reference for coding the location of an object to a body-centred system, but that this would not happen with passive viewing. This illustrates the difficulty of disentangling the hypothesized extrapolation to the present from the multitude of other mechanisms that could generate motion-induced position shifts.

Some kinds of motion induced position shifts cannot be explained by the extrapolation hypothesis. A striking example is the "flash-grab" effect. If a moving stimulus reverses direction, the perceived location of the reversal is short of the actual location. If a flash is presented contemporaneously with the reversal, the perceived location of the flash is displaced in the same direction (Cavanagh & Anstis, 2013). In this case, both the motion reversal and the flash are unpredicted, so neither can be extrapolated to the present. Additionally, the flash seems to be perceptually associated with the moving stimulus: Cavanagh and Anstis found that, for the effect to happen, the flash must occur close to the reversal in space as well as in time. There is evidence that the origins of the flash-grab effect lie in early visual processing, beginning around 80 ms ASO (Hogendoorn, Verstraten, & Cavanagh, 2015; Kohler et al., 2015), but it is also affected by attention, which is involved later in the processing stream (Cavanagh & Anstis, 2013).

The flash-grab effect poses a problem for the extrapolation to the present hypothesis. If the flash is presented contemporaneously with an unpredicted reversal of direction of a moving stimulus, then there should be an extrapolation of the moving stimulus in its pre-reversal direction. In that case, the flash should be perceived as lagging the moving stimulus. Instead, the flash and the moving stimulus are both displaced in the post-reversal direction. Accounts in which both effects are products of temporal integration over about 80 ms, such as the postdiction account by Eagleman and Sejnowski (2000, 2007), are better able to explain the results. On the other hand, there could be postdictive over-writing of an earlier percept on a time scale up to about 200 ms (Shimojo, 2014), in which case an initial extrapolation to the present would be postdictively overwritten by the process that generates the flash-grab effect, and only the latter would be available for reporting. Thus, up to the moment when the direction reversal is detected, there is an extrapolation to the present in the pre-reversal direction. When the direction reversal is detected, which involves integration of motion information over about 80 ms, the extrapolation in the pre-reversal direction is postdictively abolished and replaced with a new percept of the post-reversal direction, and the flash is "grabbed" by that.

Different features and components of stimuli are processed separately, especially in early visual processing, and the construction of a coherent percept requires the components to be integrated. There are many factors that can affect this integration process (Burr & Thompson, 2011; Hubbard, 2014; Krekelberg & Lappe, 2001), and stimuli that vary in motion and predictability can expose the problems the visual system faces in integrating information over time. This means that it is difficult if not impossible to distinguish effects of unpredicted stimuli and extrapolation to the present from early separate feature processing and subsequent integration. Properties of visual motion such as velocity perceived by integrating information over time (McKee & Welch, 1985; Simpson, 1994; Snowden & Braddick, 1991), so some uncertainty about the spatial relation between a moving object and a stationary one may result from that. For these reasons, it is not possible at present to identify a motion induced position shift that can be explained only as an outcome of extrapolation to the present and not in any other way. Given the many factors that may affect the relative times at which different events are perceived to occur (Brenner & Smeets, 2000; Hubbard, 2014; Krekelberg & Lappe, 2001; Maus et al., 2010), and the many other problems the visual system has in binding changing features of objects (e.g. Hubbard, 2014; Kang & Shevell, 2012; Moutoussis & Zeki, 1997a, 1997b; Sheth, Nijhawan, & Shimojo, 2000), there appears to be little prospect of identifying a phenomenon in visual perception of moving stimuli that can be explained only by the extrapolation to the present hypothesis.

Representational momentum (RM)

To begin with a brief note on terminology, strictly speaking, RM refers to just the horizontal component of the extrapolation of an object's motion. The term "displacement" is used to cover both RM and other components of motion extrapolation, such as the downward component that may reflect an implicit belief about effects of gravity (Hubbard, 2005). That convention will be used here.

In studies of displacement, participants are asked to indicate the location in a visual display at which a moving object disappeared. Most stimuli have presented objects in implied motion, with successively presented static frames showing an object in different locations or orientations, or in apparent motion across a computer screen. There is a general tendency to remember the site of the object’s disappearance as being displaced in the direction of the implied or apparent motion (Freyd & Finke, 1984; Freyd, 1987; Hubbard, 2005). RM represents an extrapolation of object motion beyond the final location of the stimulus in the direction of motion, and Hubbard (2013) argued that RM and the flash-lag effect were different manifestations of the same basic phenomenon. The aim of this section, therefore, is to assess whether displacement is extrapolation to the present or not.

Finke et al. (1986) proposed that RM occurs because "there is a natural tendency to mentally extrapolate implied motions into the future" (p. 176). They argued that it is useful to anticipate the future behaviour of objects to the extent that it is predictable, and that it contributes to action monitoring and control. Hubbard (2005) proposed that RM represents an extrapolation of a moving object's location to the point at which a planned interception would intersect the object's trajectory. To the extent that target motion is predictable, such extrapolation would facilitate effective interaction with the object. This will be called the "interception" hypothesis.

One implication of the interception hypothesis is that extrapolations should be as accurate as possible, and should therefore be based on relevant real world knowledge. Numerous studies have shown that displacement is based on knowledge of object motion, incorporating inferences about dynamics as well as kinematics (Finke et al., 1986; Hubbard, 2005; White, 2012). For example, Finke et al. (1986) showed that the extrapolation exhibited in RM is not instantly stopped. Instead, the extrapolation models the application of a countervailing force that slows the object down, which they called "cognitive resistance" (p. 177). Also, the extrapolated trajectory is not a simple horizontal extension of horizontal motion exhibited by the target object but, if the object appears unsupported, shows a downwards component which may represent an assumption about the effects of gravity (Hubbard, 1990; Hubbard & Bharucha, 1988; Motes, Hubbard, Courtney, & Rypma, 2008). Not all effects observed in research on displacement reflect knowledge of object motion, however. For example, the amount of displacement observed is affected by the presence and relative location of another, stationary object with which the moving object does not interact (Hubbard & Ruppel, 1999).

Under the extrapolation hypothesis, the perceived location of the moving object is its predicted location at the present moment, based on the most recent available information. The interception hypothesis about RM, however, is that it is a prediction to the moment in a planned action at which the moving object will be intercepted. That is a moment in the future, not the present moment. This implies that, if the interception hypothesis is correct, then the magnitude of RM should be greater than that of the illusory displacement of the moving object in the flash-lag effect. This is in fact the case. The lag in the flash-lag effect varies from about 27 ms to about 90 ms (Arnold et al., 2009; Chakravarthi & VanRullen, 2012; Eagleman & Sejnowski, 2000, 2007; van de Grind, 2002), but the amount of extrapolation in RM can be several hundred milliseconds (Nakamoto, Mori, Ikudome, Unenaka, & Imanaka (2015).

Additional relevant evidence comes from studies of auditory displacement. There have been several of these; for example, Hubbard (1995) presented a series of pitch changes or glissandi, and found evidence for displacement in the direction on the pitch dimension implied by the series. However, the studies of most relevance here are of spatial perception, particularly where direct comparisons can be made between visual and auditory spatial perception (Feinkohl, Locke, Leung, & Carlile, 2014; Getzmann & Lewald, 2009; Getzmann, Lewald, & Guski, 2004; Schmiedchen, Freigang, Rübsamen, & Richter, 2013). The studies have found evidence for displacement with moving auditory stimuli. Feinkohl et al. (2014) found two independent effects: auditory stimuli were mislocalised towards a visual fixation point, an effect apparently not found with visual stimuli and, once that was controlled for, a displacement effect. They found that the amount of auditory displacement increased with increasing velocity of the stimulus, an effect also found with visual displacement (Hubbard, 2005). However, the finding of most relevance for present purposes is that the magnitude of the displacement effect was similar for auditory and visual stimuli, as it was also in Schmiedchen et al. (2013). This favours the interception hypothesis over the extrapolation to the present hypothesis. Auditory processing is faster than visual processing (Vroomen & Keetels, 2010), so extrapolation to the present would be shorter for auditory than for visual stimuli, other things being equal. However, extrapolation to interception should be the same for both modalities because it is to the same anticipated future moment, that when interception is expected to occur.

In summary, the evidence currently favours the hypothesis that displacement is a form of extrapolation to the moment of an anticipated interception, not extrapolation to the present. The objection that the participants are passive observers not expecting to act on the stimulus is countered in two ways. First, Schubotz and von Cramon (2001, 2002) found that perceptual predictions led to activation in areas of the brain associated with action planning, suggesting that the motion stimuli in displacement studies activate anticipatory action planning even when no action needs to be made. Second, Kerzel and Gegenfurter (2003) compared conditions in which observers indicated the location at which the object disappeared by either reaching to and contacting the display with a finger or judging a subsequently presented probe stimulus. More RM was observed in the former condition than in the latter, consistent with the hypothesis of a connection with the action system. Perhaps a certain amount of extrapolation occurs when the observer is merely passive (e.g. at the retina; see below) and more occurs at a late stage of processing when integrating perceptual analysis with motor plans. This can be contrasted with the flash-lag effect, where it has been found that active control of the moving stimulus by the participants reduced, rather than enhanced, the extent of the flash-lag effect (Ichikawa & Masakura, 2006), at least when the coupling between mouse movement and stimulus movement had a familiar directional relation (Ichikawa & Masakura, 2010). It is therefore most likely that extrapolation to the present and displacement are different phenomena, and that displacement is a case of extrapolation to guide interceptions (Nakamoto et al., 2015).

Judging time of occurrence of an unpredicted stimulus

Libet, Gleason, Wright, and Pearl (1983) ran a study in which participants were asked to make a voluntary movement at a time of their own choosing. They watched a clock with a spot of light rotating around its perimeter and reported the position of the spot of light (hereafter "clock time") when they first felt the urge to move. The relevant condition for present purposes is not the voluntary movement condition but a control condition in which participants reported the clock time when a brief somatosensory stimulus was presented. The motion of the spot of light is regular and predictable, so under the extrapolation to the present hypothesis it should be extrapolated forward to compensate for processing latencies. The somatosensory stimulus is not predictable, however, and is in fact equivalent to the flash in a flash-lag display, so it should not be extrapolated forward. This leads to a prediction that the somatosensory stimulus should be reported as occurring later than it did, by an amount equal to the processing latency. In fact Libet et al. (1983) found that the stimulus was reported as occurring a mean of 47 ms earlier than it did, contrary to the prediction based on the extrapolation to the present hypothesis.

Research on cross-modal temporal order judgment has shown that tactile stimuli have to be presented about 30 ms before visual stimuli if they are to be perceived as simultaneous (Efron, 1963; Spence, Shore, & Klein, 2001). That is not necessarily an indication of the different processing latencies of the two modalities and could just reflect the way in which the mechanism of cross-modal temporal comparison operates. However, the difference implies that the skin stimulus should be reported as occurring about 30 ms earlier than it did, which is consistent with the mean of 47 ms reported by Libet et al. (1983). Thus, the results can be explained in terms of the mechanism of cross-modal temporal order judgment. However, evidence was reported earlier that the temporal amount of the flash-lag illusion is in the region 27 - 90 ms. Subtracting 30 ms from that still implies that the skin stimulus should be perceived as occurring either when it did or later than it did. Therefore the finding is not consistent with the hypothesis of extrapolation to the present. However, the commentary section on postdiction will show that that might not be the end of the story.

Commentary 1: deliberative compensation

Joordens, Spalek, Razmy, and van Duijn (2004) found evidence consistent with a hypothesis that the extent of occurrence of RM is affected by a compensation process that is subject to disruption by cognitive load manipulations and therefore presumably is controlled rather than automatic processing. Participants in one of the experiments by Joordens et al. reported in debriefing that they were making a deliberative compensation: they noticed a possible biasing factor and deliberately tried to compensate for it in their judgments. Such a compensation mechanism could be a common feature of post-perceptual processing of temporal information: the research discussed so far has relied on explicit reports of percepts or judgments, and therefore all of it is subject to influence from deliberative compensation mechanisms. Thus, deliberative compensation could account for the evidence that the extent of the flash-lag effect is not as great as would be expected under the extrapolation to the present hypothesis; it could be operative in motion induced position shifts as well; and it could affect reports of clock time in the study by Libet et al. (1983). If compensation is a controlled process then it is liable to be affected by many features of experimental methods and manipulations. There is evidence for effects of deliberative compensation on other kinds of processes (Joordens et al., 2004; Merikle, Joordens, & Stolz, 1997) but so far there has been no further investigation of its effects on the phenomena of relevance to this paper. This adds to the difficulty of drawing conclusions from any of the relevant evidence. It is perhaps worth noting that deliberative compensation is a possibility in any experiment where data are derived from voluntary responses or judgments, and that the extent of its occurrence is currently unknown.

Commentary 2: the relevance of postdiction

Postdiction is pervasive in perception (Choi & Scholl, 2006; Eagleman, 2010; Eagleman & Sejnowski, 2000; Kilgard & Merzenich, 1995; Kim, Feldman, & Singh, 2013; Müsseler & Tigglebeck, 2013; Parsons, Novich, & Eagleman, 2013; Shimojo, 2014; Yarrow, Haggard, Heal, Brown, & Rothwell, 2001). For example, postdiction is necessarily involved in apparent motion (Gepshtein & Kubovy, 2007; Kolers & von Grünau, 1976; Ramachandran & Anstis, 1986): where a flash occurs successively at locations A and B, motion from A to B cannot be inferred until the flash at location B has occurred, which means that the apparent motion must be a postdictive construction. There is evidence that the reach of postdictive reinterpretation extends at least 200 ms into the past (Choi & Scholl, 2006; Shimojo, 2014). This does not mean literal alteration to the past, of course: it means that perceptual processing of a stimulus continues for long enough that subsequent input as much as 200 ms after the stimulus can affect how it is perceived. Under some circumstances the temporal reach of postdictive reinterpretation could be as much as 1,000 ms (Geldard & Sherrick, 1972; Khuu, Kidd, & Errington, 2010). However, a general amount of 200 ms is sufficient for the argument made here.

Postdiction creates a problem for any research on extrapolation to the present on a time scale of 200 ms or less, which covers all of the research reviewed so far. The occurrence of an isolated flash (Nijhawan, 1994) or a single somatosensory stimulus (Libet et al., 1983) cannot be predicted unless it is a member of a regular series, but it can be postdicted. That is, the perceived time of occurrence of the stimulus can be postdictively altered within a scale of about 200 ms. This means that any unpredicted stimulus can be extrapolated to the present by postdictive adjustment.[4] Since this hypothetical adjustment occurs on a short time scale, almost any overt report, judgment, or behavioural measure would be informed by the postdictively adjusted time of occurrence and not by the original percept, which would have been obliterated by the postdictive adjustment. The only exception would be responses made before the postdictive judgment had occurred or had time to influence judgment. This might be the case for studies with speeded responses. For example, in studies of the temporal development of metacontrast masking (Lachter & Durgin, 1999; Lachter, Durgin, & Washington, 2000), participants were required to respond no more than 480 ms after the stimulus presentation. Given the time required to generate and deliver the response, it is likely that such a rapid response would not be influenced by postdictive adjustment.

If this kind of postdiction is occurring, then none of the evidence considered so far constitutes a valid test of the extrapolation to the present hypothesis because all stimuli, predictable or unpredictable, are extrapolated to the present. In the case of unpredicted stimuli, the mechanism of extrapolation is postdiction. Thus, no predictions can be made about discrepancies between perceived times of occurrence of different events. It is possible that some of the evidence shows imperfections in extrapolation mechanisms: this might apply, for example, to the study by Shi and Nijhawan (2012), since it is unlikely that a perceptual adjustment mechanism would be sensitive to effects of foveal sensitivity to light of different wavelengths. It is far from clear that this must be the case, however, so the evidence from that study, as from all the others, cannot be regarded as either confirmatory or disconfirmatory for the extrapolation to the present hypothesis unless the hypothesis of postdictive adjustment can be ruled out.

Review part 3: Static visual illusions

It has been claimed that many static visual illusions can be explained by the hypothesis that the visual system uses available information to extrapolate to the present moment (Changizi, 2001; Changizi et al., 2008; Changizi & Widders, 2002; De Valois & De Valois, 1991). As Changizi (2001, p. 195) put it, "given the proximal stimulus, the scene an observer perceives is *the probable scene present at the time of the percept*". This has become known as the "perceiving the present" hypothesis, although that is, in my view, a misleading label, because it gives the impression that zero delay perception occurs. The proposal should more properly be called extrapolation from static visual features.

The essence of the argument is that static two-dimensional geometric figures possess cues to the kind of three-dimensional structure represented in them, and cues to the observer's direction of motion relative to the structure represented in the figure. Thus, the illusion represents a perceptual extrapolation from the available information to the present, based on the assumption that the observer is moving with respect to the structures depicted, in the direction indicated by the motion cues. This in turn assumes that the observer is gazing in the direction of motion for a substantial proportion of the time. This argument motivated a detailed analysis of regularities in the optic flow which ultimately resulted in a systematic categorisation of visual illusions into a 7 x 4 grid of possibilities (Changizi et al., 2008). The analysis generated predictions for new visual illusions, or variations on ones already known, which were supported by experimental findings (Changizi et al., 2008; Changizi & Widders, 2002). The stimuli that give rise to the illusions present cues to the observer's motion direction that generate extrapolations to the present, but the extrapolations are incorrect because in fact the stimuli are static and the observer is not moving with respect to them (or the structure depicted in them). The illusion is a product of the incorrect extrapolation.

A key component of the argument made by Changizi and colleagues is that static figures may contain clues to or static representations of optic flow, implying that the observer is moving forward with respect to the figure. One of the illusions interpreted in this way was the Hering illusion, in which two parallel lines are presented in a context of radial lines in a sunburst pattern. The illusion is that the parallel lines appear to be bowed outwards (Figure 1). Changizi et al. (2008) argued that the sunburst pattern resembled the optic flow pattern that would occur if the observer were walking toward the parallel lines. The illusion represents an extrapolation of the static image of the parallel lines to the image that would be presented if the observer approached the lines: the bowing occurs because the centre of the parallel lines would be closer to the observer than the extremes, thereby distorting the image at the retina, and this disparity would increase as the observer approached the lines.

Vaughn and Eagleman (2013) argued that, under that interpretation, presenting moving optic flow patterns instead of the static radial lines should also induce the illusion. To test this, they presented static parallel lines in three different optic flow contexts using fields of dots: dots moving outwards, implying observer motion toward the vertical lines, dots moving inwards, implying observer motion away from the lines, and static dots. The illusion did not occur with the static dots, but did occur with both optic flow stimuli. This is a problem because the inward moving dots should have implied that the observer was moving away from the object and should therefore have given rise to an illusion in the opposite direction. The authors suggested that the perceptual mechanism is directionally insensitive and interprets any kind of optic flow as implying forward motion (cf. Lewis & McBeath, 2004). This seems like weak support for the hypothesis of extrapolation from static visual features, but other features of the results strengthen it. As presentation time of the parallel lines increased, so the illusion decreased. The authors argued that prolonged presentation shows that the bars are not moving with the optic flow, so they are perceptually decoupled from it. At present no other satisfactory interpretation of these results has been proposed.

The strength of the extrapolation from static visual features hypothesis is that it generates a systematic categorisation of a large number of visual illusions, locating them in a common explanatory framework (Changizi et al., 2008). The main weakness lies in the assumptions about the observer's motion. A core proposition of the account is that the extrapolation made is that which is empirically most likely given the available information. The information in the static figure implies both a three-dimensional structure and a direction of observer motion, depending on the assumption that the observer is usually gazing in the direction of motion. That assumption remains in need of empirical confirmation and has been questioned (Briscoe, 2010). In addition, although the contents of the figure may provide cues to observer motion, there are other cues to observer non-motion, in particular the evident fact that the scene depicted in the stimulus does not change over time, and proprioceptive cues that the observer is in fact stationary. The latter objection may not be fatal. Much visual experience occurs under conditions of passive motion, especially during infancy when much perceptual learning takes place, and even for adults when they are in moving vehicles, so the visual system must be accustomed to motion-induced changes in visual input in the absence of proprioceptive cues to movement.

However, there is also some evidence for illusions that disconfirm predictions generated by the hypothesis of extrapolation from static visual features (Briscoe, 2010; Prinzmetal, Shimamura, & Mikolinski, 2001). To illustrate, Prinzmetal et al. (2001) investigated the conditions under which the Ponzo illusion occurs: this is the illusion in which two vertical lines of equal length are perceived as of unequal length when presented in a context of radial lines (Figure 2). A key component of the explanation of the Ponzo illusion under the hypothesis of extrapolation from static visual features is that the radial lines are interpreted as perspective indicators, so that one line is misperceived as longer than the other because it is interpreted as being further away. (This interpretation is not exclusive to the hypothesis of extrapolation from static visual features - e.g. Gregory, 1963 - but is an important part of it.) Prinzmetal et al. obtained findings that are not predicted by this account. For example, they presented the two lines in different spatial relations to a figure that should have been interpreted as a perspective indicator. They found that occurrences, nonoccurrences, and reversals of the Ponzo illusion occurred in a pattern that could not be accounted for by the perspective interpretation. Instead, the results were consistent with a tilt constancy theory, according to which some illusions are generated by distortions to the observer's sense of what is vertical by contexts of tilted lines. Prinzmetal et al. argued that this theory could account for the Hering illusion as well.

At present, the hypothesis of extrapolation from static visual features provides a powerful and comprehensive explanatory account of many static visual illusions, but the problems for the hypothesis indicate that the issue is far from settled. It remains possible that visual illusions can be explained without the assumption of observer motion that is critical to the hypothesis of extrapolation from static visual features. It is not likely that a single theory can account for all visual illusions. The issue is whether the hypothesis of extrapolation from static visual features is the correct explanation for some visual illusions, and that remains unclear at present.

Having said that, there is an interesting connection to another body of research that has presented static visual stimuli with cues to either motion or dynamics. In studies by Freyd (1983) and Futterweit and Beilin (1994), static images taken from a recorded action sequence were presented. Participants were then asked to judge whether a second image was the same as or different from the first. The evidence showed longer reaction times (Freyd, 1983) or higher error rates (Futterweit & Beilin, 1994) when the image to be judged was after the first image in time than when it was before. Freyd, Pantzer, and Cheng (1988) found a similar pattern of error reporting for static images that implied a direction of motion in an inanimate object. Freyd (1987) argued that these effects show encoding in memory in terms of dynamic representations, meaning that implied forces and masses are encoded along with observable features of the image. This does not fully explain the results, however, since force vectors can be encoded at a moment: that is, one could just remember the observable features of the stimulus with implied force vectors attached. Remembering the stimulus as something a little further forward in time suggests that something more than just encoding of force vectors is going on. That something could be extrapolation. It is not clear that it must be extrapolation to the present, rather than extrapolation to an anticipated time of contact. Indeed, a study by Didierjean and Marmeche (2005) that found differences between judgments of experts and novices with stimuli concerning events in basketball indicates that extrapolation may occur even further ahead than that: a static scene can, for an expert, evoke an entire action sequence that is likely to follow it. However, the research combines with that by Changizi et al. (2008) to suggest the possibility of extrapolation to the present in viewing all kinds of static images that imply something going on. This is worthy of further investigation.

Kawabe, Yamada, and Miura (2007) presented static stimuli comprising a circle accompanied by three parallel straight lines that resembled motion lines, lines used in static images to convey an impression of motion in a particular direction. In several studies they found that the remembered location of the circle was biased in the direction of motion implied by the motion lines. Having ruled out several other interpretations, they argued that the findings supported the hypothesis that motion lines trigger motion extrapolation, which results in the memory errors shown in the experiments. This can also be interpreted as extrapolation to the present from a static image. Kawabe et al. reviewed evidence showing that motion lines activate early motion processing (e.g. Burr & Ross, 2002), and they argued that this gives rise to "motion processing without motion perception" (p. 318), generating anticipation of future positions of the objects.

The phenomena described in the preceding two paragraphs are not visual illusions. It could be argued that they are memory illusions (Roediger, 1996), although the stage of processing at which the error is introduced is not clear. Nevertheless, it is striking that all of the findings discussed here with stimuli comprising static images can be accommodated within the same hypothesis of extrapolation from static visual features. This is consistent with the general hypothesis that perceptual processing is an attempt to model the present by extrapolating from the most recent available information.

Review part 4: Motion extrapolation at the retina

Berry, Brivanlou, Jordan, and Meister (1999) recorded responses of retinal ganglion cells in rabbits and salamanders to moving object stimuli. The usual response latency in retinal ganglion cells (from light entering the eye) is about 70 ms (Baylor & Hodgkin, 1974; Johnston & Lagnado, 2015). When there is a moving stimulus, peak firing rate occurs earlier than that in cells ahead of the moving stimulus. Berry et al. suggested that the cells begin to fire "when the bar begins to invade their receptive-field centre" (p. 335). This is perhaps better described as an early response than as an anticipatory response because it does depend on the stimulus beginning to invade the receptive field, but nevertheless the effect is to abolish the 70 ms processing delay. The early response did not occur if speed of stimulus motion at the retina exceeded ~1 mm s^{-1}, and Berry et al. noted that this corresponded to the limits on retinal speeds at which extrapolation of perceived motion occurs (Nijhawan, 1994). Thus, in a situation where most other information is subject to a 70 ms processing delay at the retina, information about motion below a certain speed is subject to zero delay. This is functionally equivalent to extrapolation to the present. Palmer, Marre, Berry, and Bialek (2015) argued for a predictive coding approach to retinal processing, such that predictable input is essentially disregarded and only unpredicted input is responded to. In keeping with this, "The retina actively responds to predictable features of the visual stimulus... and, in the case of smooth motion, can anticipate an object's location in a manner that corrects for its own processing delay" (p. 6912). Their research indicated that the retina is close to optimal in that respect.

The implications of that for cortical processing remain to be ascertained. It would be odd if extrapolation occurred at the retina only for the functional advantages of that to be lost by subsequent processing delays and, if retinal mechanisms are capable of extrapolation, then it would be odd if cortical mechanisms were not capable of it. Müsseler, Stork, and Kerzel (2002) proposed a functionally equivalent mechanism at cortical level, in the form of an asymmetrical spreading activation model (see also Whitney & Murakami, 1998). A visual stimulus elicits an activation pattern that spreads to adjacent parts of the visual field. In the case of a moving stimulus, the previously activated parts of the visual field in the direction in which the stimulus is moving contribute to the spread of activity, in effect running ahead of the current location of activation associated with the stimulus. This compensates for processing latencies in the direction of motion. Müsseler et al. argued that this could account for the flash-lag effect, and different aspects of the process could account for other illusory position shifts, including RM. The model's predictions have yet to be tested, and it would be difficult to disentangle the effects of the mechanism from the many other factors that affect perceived locations of moving objects (see above).

One limitation of both the retinal and cortical extrapolation mechanisms is that they seem to apply only to the case of motion perception: different mechanisms would seem to be required for other kinds of predictable change, such as luminance change at a constant rate, and changes in other modalities. Also, it is not clear whether the mechanism is sensitive to differences in velocity (within the evident 1 mm s^{-1} speed limit) or whether it operates at a constant rate regardless of stimulus velocity. Insensitivity to velocity would reduce the accuracy of the extrapolation.

Review part 5: Other modalities

Almost all of the research discussed so far concerns the visual modality. There has been research on RM in audition (Feinkohl et al., 2014; Getzmann & Lewald, 2009; Getzmann et al., 2004; Hubbard, 1995; Johnston & Jones, 2006; Schmiedchen et al., 2013), and a study of RM in the haptic modality (Brouwer, Thornton, & Franz, 2005), but the review has shown that extrapolation to the present is not the most likely explanation for RM. There has been research on auditory and haptic equivalents to the flash-lag effect (Alais & Burr, 2003; Arrighi et al., 2005; Cellini et al., 2016; Nijhawan & Kirschfeld, 2003), but these may be explained by modality-specific processing features. There are two indirect lines of evidence concerning other modalities, however, that may be relevant.

As I discussed briefly early on in the paper, there are cross-modal synchronisation processes that integrate information from different modalities to establish percepts of contemporaneous events. Thus, if visual percepts are extrapolations to the present, it would seem that percepts in other modalities must be extrapolations to the present as well, otherwise cross-modal synchronisation would break down and there would be evident asynchronies between stimuli that would be expected to be synchronous, such as visual information about actions on a musical instrument and auditory information about the sound produced.

The second line of evidence is that there are predictive processes in other modalities. Taking the example of speech perception, it is well established that auditory processing involves the generation of predictions about imminent events in speech, such as word endings or the likely continuation of a sentence when one or another form of the indefinite article is presented (D'Ausilio, Jarmolowska, Busan, Bufalari, & Craighero, 2011; Delong, Urbach, & Kutas, 2005; Kuperberg & Jaeger, 2016; Maess, Mamashli, Obleser, Helle, & Friederici, 2016). This kind of prediction is equivalent to the prediction in the perception for action system that guides interceptive actions, in that it predicts beyond the present moment. The amount of extrapolation to the present is determined by the processing latency. The amount of prediction in speech perception is determined by, among other things, the latency to the input that is the subject of the prediction. This may vary independently of processing latency. The duration of individual spoken words, for example, ranges approximately from 300 to 1,000 ms (Stephens, Honey, & Hasson, 2013; Vollrath, Kazenwadel, and Krüger, 1992). The amount of extrapolation in speech perception is, therefore, both variable and often longer than the processing latency. A simple example is prediction of the end of a conversational turn, which supports timing of the initiation of a conversational turn by a new speaker with, often, considerable accuracy (Garrod & Pickering, 2015; Sacks, Schegloff, & Jefferson, 1974). This form of prediction, therefore, resembles the form of extrapolation that is a guide to action in the interception of projectiles, and is different from extrapolation to the present. However, if prediction to the future is possible in audition, then prediction to the present does not seem to be out of the question. There is an evident need for further research on this.

An alternative hypothesis: the short lag

It is clear from the foregoing review that there is both a quantity and a good variety of research evidence that is consistent with the hypothesis of extrapolation to the present. There are results for which no plausible alternative to extrapolation to the present has yet been proposed, principally extrapolation at the retina (Berry et al., 1997), and the colour versions of the flash-lag illusion (Nijhawan, 1997; Shi & Nijhawan, 2012). There are, however, alternative possible explanations for most of the evidence, and there is active contention about the best interpretation of the evidence in most areas. The validity of the extrapolation to the present hypothesis, therefore, remains uncertain.

But is there a viable alternative to extrapolation to the present? I argued earlier in this paper that processing latencies in visual perception (in the ventral pathway) would result in subjectively evident asynchronies with action. In the case of badminton, I argued that a player could have a visual percept of their opponent hitting the shuttlecock at them after they had already intercepted the shot. That kind of asynchrony does not occur. Extrapolation to the present is one way of precluding such asynchronies, by compensating for processing latencies. Another possibility is that the entire perceived world, including all modalities and perception through the vision for action pathway and the experience of acting itself, lags the objective present by a substantial fraction of a second. Under that possibility, either processing through the dorsal pathway must be delayed far beyond what the reaction time evidence shows it to be, or perception for action is entirely divorced from the perceived present. The only other possibility is that conscious percepts emerge with a latency that is sufficiently short as to engender no experienced asynchrony with action. That is the possibility that will be explored in this section.

As far as I have been able to discover, there is no research evidence on the detection of asynchrony between information processing in the dorsal and ventral streams. The best proxy evidence would be detection of cross-modal asynchrony. The minimum threshold for detection of cross-modal asynchrony involving visual and either auditory or somatosensory stimuli is about 22 - 30 ms (Spence, Baddeley, Zampini, James, & Shore, 2003; Spence et al., 2001; Zampini, Shore, & Spence, 2003). However, under some circumstances, asynchrony between visual and auditory input must be as great as 200 ms before it becomes noticeable (Dixon & Spitz, 1980), and unnoticed but substantial temporal mismatches between auditory and visual stimuli seem to be common (Freeman et al., 2013; Ipser, Agolli, Bajraktari, Al-Alawi, Djaafara, & Freeman, 2017). If that has any validity as a guide to asynchrony detection in general, then the perceived present could lag behind information processing in the dorsal stream by as much as 200 ms before it would be noticeable, though the evidence suggests that shorter asynchronies can be detected at least some of the time. How short could the latency to conscious visual perception be, and could it be short enough to account for the lack of perceived asynchrony in the badminton example?

If motion extrapolation occurs at the retina (Berry et al., 1999), then the earliest possible latency for a conscious visual percept is a little more than 30 ms, if we assume a minimum for level 1 cortical processing of 30 ms (Holcombe, 2009). There is also evidence for a fast visual processing route involving the amygdala, specifially for threat-related stimuli, on a time scale of about 30 - 60 ms (Luo, Holroyd, Jones, Hendler, & Blair, 2007; Morel, Beaucousin, Perrin, & George, 2012), although it is not clear that this is associated with conscious percepts at that latency. Could there be a conscious visual percept with a latency of only 30 ms?

According to Holcombe (2009), higher level visual processing begins about 100 ms ASO. A recent study by Bieniek, Bennett, Sekuler, and Rousselet (2016) yielded a latency of 90 ms. It remains questionable whether any semantic processing occurs as early as that: a more realistic estimate for that would be around 130 - 150 ms (Rossion & Caharel, 2011) which, as was also shown earlier in the paper, is also the earliest latency for representation of categorical information about scene gist. It is also the latency of the earliest response in the ERP waveform that differentiates between conscious and nonconscious stimuli (Bagattini, Mazzi, & Savazzi, 2015; Koivisto & Revonsuo, 2010; Shafto & Pitts, 2015). But those latencies are for novel, unpredicted stimuli. There is evidence that processing latencies are shorter for predicted stimuli (Bachmann, 1989; Bachmann et al., 2003), and a process of adjusting a percept that already exists may be faster than a process that generates a novel percept from unanticipated input (Galletti & Fattori, 2003). There is a case, then, that some conscious visual percepts could emerge with latencies considerably less than 150 ms.

On the other hand, the evidence reviewed earlier in the paper indicates that conscious percepts for novel stimuli are unlikely to have a latency less than ~150 ms, and some perceptual products are generated with latencies much longer than that. This suggests that either conscious percepts do not emerge until perceptual processing is essentially complete, which would be after 300 - 400 ms, or that conscious percepts emerge after about 150 ms but are somewhat incomplete at that latency, and are gradually filled in as more information becomes available. There are, however, ways in which perceptual processing can compensate for the incompleteness of perceptual information at early stages.

1. Incomplete percepts. One possibility is that even incomplete perceptual information could be emerge sufficiently early to avoid evident asynchrony with action. Thus, a perceptual object may be individuated and motion may be computed on a time scale of about 100 ms. Even though much featural and semantic information might be unavailable at that early stage, what is there might suffice for an individuated object percept that is fuzzy in some respects but clear in terms of its judged motion trajectory. That is to say, percepts could emerge long before feature analysis is complete, but still containing enough information to be useful.

2. Re-entrant processing. Re-entrant processing is a way in which pre-existing perceptual structures can be involved in the interpretation of sensory input. It involves a rapid wave of stimulation from low levels through to high levels followed by descending signals constituting an iterative loop (Bar, 2003; Bar, Kassam, Ghuman, Boshyan, Schmid, Dale, Hämäläinen, Marinkovic, Schacter, Rosen, & Halgren, 2006; Di Lollo, 2012; Di Lollo, Enns, & Rensink, 2000; Hochstein & Ahissar, 2002; Kahan & Enns, 2014; Tapia & Beck, 2014). Enns and di Lollo (2000) proposed that existing perceptual structures functioned as hypotheses that were matched against incoming information: the incoming information is then filled out by the information in the hypothesis that provides the best match to the data. Thus, early processing of a moving object that conveys information about object boundaries and size might be sufficient for a rapid match to a hypothesis that fills in other information about the identity of the object. The latency with which re-entrant processing occurs is not certain, but Kahan and Enns (2014) have reported evidence that it occurs in early visual

processing, even before object individuation has occurred. If that is the case, then filling out incoming data with semantically rich perceptual hypotheses by matching could generate a relatively complete and categorical percept at an early stage, consistent with emergence of a conscious percept by 150 ms.

3. Postdiction. As discussed above, the available evidence (Choi & Scholl, 2006; Shimojo, 2014) indicates that the reach of postdictive reinterpretation extends at least 200 ms into the past. In that case, a conscious percept that emerges after 150 ms may be subject to postdictive reinterpretation utilising information that emerges after about 350 ms, which would enable a more complete percept to be constructed. This is not literal alteration to the past, of course. At the 150 ms mark, the percept is still incomplete and lacks in particular what the postdictive interpretation will later provide. However, the unaltered percept survives for no more than 200 ms, after which only the postdictively interpreted version is available and reportable. The previous version could still have influenced subsequent processing to some degree. However, studies have shown that reportable stimuli are associated with what is often called global ignition, meaning amplification of information to high-level activity, whereas unreportable stimuli tend to have limited and short-lived influence (Dehaene & Changeux, 2011; Dehaene, Charles, King, & Marti, 2014). Dehaene et al. (2014) argued that this ignition occurs 300 ms or more after the stimulus in question is presented, a time by which the unaltered percept has been obliterated by postdiction. Because of that, it is unlikely that the early percept will enter subsequent processing: subsequent processing is more likely to take as input the later, postdictively reinterpreted perceptual information. In effect, from the perceiver's point of view (and particularly from the point of view of making any kind of report about what was seen), only the postdictively reinterpreted percept is available, and so it will seem as though that is the only version there has ever been.

4. Synchronisation. As we have already seen, synchronisation is ubiquitous in perceptual processing. Within vision, different features of a perceptual object have different

processing latencies. For example, colour is processed before orientation by about 63 ms (Moutoussis & Zeki, 1997b). This has important implications for perception of a moving object: if a projectile is moving at 90 mph, then the most recent available information about the ball's orientation is about 8 feet behind the most recent available information about its colour, so there should be a percept of orientation lagging about 8 feet behind a percept of colour. This is not the case: several studies have shown that percepts of objects moving across the visual field are sharp, clear, and unified at a single location, more so than would be expected from the temporal resolution of visual processing (Bedell, Tong, & Aydin, 2010; Burr, 1980; Marinovic & Arnold, 2013; Ramachandran, Madhusudhan, & Vidyasagar, 1974; Scharnowski, Hermens, Kammer, Öğmen, & Herzog, 2007; Tong, Patel, & Bedell, 2005; Westerink & Teunissen, 1995). This shows that feature information is bound into a synchronous, coherent percept. Although usually discussed as a problem of feature binding (e.g. Holcombe, 2009), it is also a problem of synchronisation, as the cricket ball example shows: asynchronous outputs have to be bound into a synchronous representation of a moving object where all of the features are in the same location in visual space at a given time.

It might seem that synchronisation is not much help to the hypothesis of emergence of conscious percepts after 150 ms. After all, synchronisation cannot be accomplished until the slower of the two sets of perceptual information that are to be synchronised has emerged, and the synchronisation process itself presumably takes time. Fully synchronised percepts could therefore emerge several hundred ms ASO. I shall return to that issue in a moment. First, however, the main point to be made about synchronisation here is that it can apply to information in the dorsal and ventral streams. Thus, the experience of acting on an object, with a latency of less than 100 ms for processing visual information in the dorsal stream, can be synchronised with an early conscious percept emerging after 150 ms of processing in the ventral stream: the temporal gap between them is no more than 100 ms, which is well within the reach of synchronisation mechanisms (Conrey & Pisoni, 2006; Mégevand et al., 2013; van Wassenhove et al., 2007). Synchronisation, therefore, may be an important contributor to perceived synchrony between acting on a moving object and the conscious percept of that object's motion.

As to the time taken for synchronisation, the same point that was made about postdiction applies to synchronisation as well. The point of view that obtains before synchronisation is accomplished is brief, and pre-synchronised information has little or no availability to subsequent processing, particularly post-perceptual processing. The point of view that obtains after synchronisation obtains for much longer: if the relevant information is attentively processed, it can be retained on a time scale of seconds. Thus, it is the products of synchronisation processes that are available to further processing, not their unsynchronised precursors.

5. Maintenance and adjustment. Most of visual perception involves maintenance and adjustment of percepts that have already been well processed. This is because most of our visual environment stays the same on short time scales, or changes in ways that can be fitted by predictive models. Much of what is around us is static objects. They have been there for some time, perceptual processing of them is complete, and they are simply maintained in a completely processed form from one moment to the next, with due adjustments for such things as saccades, the observer's own movement, and changes in lighting conditions (Donaldson, 2000; Galletti & Fattori, 2003). Things that are in motion have also been completely processed and the change in their position can be predicted from one moment to the next by means of a model of object motion (White, 2012) that is updated by testing against new input. To return to Nijhawan's cricket ball example, it might take 400 ms for a complete percept of the ball to be constructed from its first appearance but, once that process is complete, the complete percept can be mantained and adjusted from moment to moment in accordance with new input information about changes, e.g. in the ball's location or trajectory

(Bahill & Karnavas, 1993).[5] A key point in this respect is that processing changes to features of existing objects is faster than processing of new stimuli: if the feature change plays the role of the flash in a study of the flash-lag effect, the flash-lag effect is abolished (Kanai, Carlson, Verstraten, & Walsh, 2009). Kanai et al. found a flash-lag effect of about 45 ms in a standard flash-lag presentation, which is consistent with other estimates of the amount of the flash-lag effect (see above), and indicates that processing of feature changes is about 45 ms faster than processing of novel stimuli. Thus, not only maintaining stable or predictable percepts, but even adjusting individual features of perceptual objects, happens on a significantly shorter time scale than constructing a percept of a novel stimulus. Studies looking at processing latencies to novel stimuli overestimate mean processing latencies to stimuli in general.

In that respect, that component of the short lag hypothesis somewhat resembles the differential latency hypothesis. This is a proposal that the flash-lag effect occurs because moving stimuli have shorter processing latencies than stationary flashed stimuli (Öğmen et al., 2004; Purushothaman et al., 1998; Whitney & Murakami, 1998; Whitney et al., 2000). That can be regarded as a specific version of the general hypothesis that processing latencies differ for different kinds of stimuli, and for different features of stimuli, though it does not specifically say that the latency difference arises because the moving stimulus is predictable and the flashed stimulus is not. Hubbard (2014) pointed out some findings that are problematic for the differential latency hypothesis, but if that specific hypothesis turns out to be false, that would not imply that the general hypothesis of different processing latencies for predictable and unpredictable stimuli was false.

To summarise, then, the hypothesis under discussion here is that most processing latencies for visual percepts are significantly less than 150 ms, perhaps even as little as 30 ms under some circumstances; this may be called the short lag hypothesis. Perceptual processing is not complete at that 150 ms ASO, but: (i) incomplete percepts can still emerge and be useful; (ii) re-entrant processing yields early percepts that are more than incomplete sketches because matching hypotheses fill in information; (iii) early products of processing in the ventral stream that are conscious at 150 ms can be synchronised with the products of processing in the dorsal stream, despite the shorter latencies that obtain there; (iv) subsequent processes of postdiction and synchronisation contribute to the perceptual impression of synchrony, completeness, and coherence; (v) most percepts are maintained information about perceptual processing of static or otherwise predictable stimuli that were completely processed earlier; (vi) maintaining and updating existing percepts has a much shorter latency than constructing a percept of a novel stimulus. Although processes such as postdiction and synchronisation take longer than 150 ms to operate, they are essentially complete before perceptual information becomes available to further processing, and for that reason further processing receives perceptual information that is more complete than the earliest conscious percept.

The short lag hypothesis is only a hypothesis. Clear evidence for the early emergence of percepts, in whatever state of incompleteness, is lacking (see, for example, Koivisto & Revonsuo, 2010), and there has yet to be any research on whether the proposed mechanisms really contribute to the subjective impression of completeness in perception despite the short latency to conscious percepts. I merely suggest it as a viable alternative to extrapolation to the present that may be worthy of further investigation.

Conclusion

The case for

The present moment provides a reference point for calibrating and synchronising the contents of percepts that emerge with different latencies. Extrapolating to the present also synchronises perceptual information in the two visual perceptual routes, so that there is a conscious percept of a projectile being intercepted at the moment when the interception actually occurs. The case for the hypothesis of extrapolation to the present, therefore, can be summarised as follows. The gaze of sportspeople supposedly keeping their eye on the ball is for much of the time directed at where the ball is now and they have a conscious percept of it there, even though the processing latency problem means that information about it is not available for at least another 100 ms (Nijhawan & Wu, 2009). Several studies have provided evidence consistent with the hypothesis of extrapolation to the present, as reviewed above. The perceived present cannot lag far behind perception for action because the asynchrony would be detected, and extrapolation to the present would minimise that lag. Prediction is pervasive in perceptual processing, from retinal processing (Berry et al., 1999) through prediction of visual scene changes consequent on saccades (Donaldson, 2000), and moment-to-moment use of existing perceptual structures to predict the next input (Hohwy, 2013; Khoei et al., 2017) to the longer extrapolations to an anticipated point of contact that occur in the interception of a moving object (Brenner & Smeets, 2015) and even beyond that (Didierjean & Marmeche, 2005). It is not implausible, therefore, for prediction to the present to be involved in the construction of visual conscious percepts, complementing other forms of prediction in perceptual processing.

The case against

The case against the hypothesis of extrapolation to the present can be summarised as follows. None of the experimental evidence is compelling. In most cases there are alternative possible interpretations of the results, and what remains is suggestive but no more than that. The avoidance of obvious asynchrony between conscious percepts and perception for action can be accounted for by the short lag hypothesis, under which conscious percepts emerge with a latency of about 100 - 150 ms, and even less in some cases; various mechanisms discussed above may be involved in filling in or otherwise compensating for longer delays in perceptual processing.

Concluding remarks

If it takes 100 ms or more to process visual information, perception of changing things is bound to be inaccurate because it is out of date. A predictive model of the present is liable to be inaccurate in a different way, because some things are inherently unpredictable. However, many kinds of change are predictable because they happen at rates that are near-constant on the millisecond time scale. It would, therefore, be advantageous to extrapolate from available information to the likely present, and to detect and adjust for errors when they occur. The evidence for something functionally equivalent to extrapolation to the present in the retina is strong, but the possibility of extrapolation to the present in cortical processing has not yet been decisively confirmed or disconfirmed. If the predictions of the extrapolation hypothesis can be distinguished from those of other mechanisms and factors in visual information processing, then that might be the best prospect for obtaining strong confirmatory or disconfirmatory evidence. If extrapolation to the present is occurring in vision, it is almost certainly occurring in other modalities as well, so there is a need for research to investigate that possibility as well.

Footnotes

1. In addition, light and sound take time to reach the peripheral sensors and, for distant sources, that time is perceptibly greater for sound than for light: the speed of sound through air depends on temperature but is in the region of 343 m/s, whereas the speed of light in air is about 299,700 km/s (the speed of light in a vacuum divided by the refractive index for air, which is 1.0001). There is evidence for audio-visual synchronisation on a time scale of 200 ms or more, although this depends on many factors (see Chen & Vroomen, 2013, and Vroomen & Keetels, 2010, for useful reviews). There is also evidence that the brain compensates for differences in arrival time of stimuli if the visual stimulus renders the arrival time of the auditory stimulus predictable (Petrini, Russell, & Pollick, 2009). However, sensitivity to asynchronous cross-modal atimuli may be quite limited (Freeman, Ipser, Palmbaha, Paunoiu, Brown, Lambert, Leff, & Driver, 2013; Ipser et al., 2017).

2. In this respect, the consensus of research has some resemblance to the "multiple drafts" model of consciousness (Dennett, 1991; Dennett & Kinsbourne, 1992). It only has "some" resemblance to the multiple drafts model, in part because that is a model of consciousness that attempts to replace the notion of a single "Cartesian" viewpoint of consciousness for a distributed representation with local modification. The present paper is not concerned with the nature of consciousness: terms such as "conscious percepts" only where other authors have used them, as in the literature on time to emergence of visual awareness (Koivisto & Revonsuo, 2010).

3. This assumes that the 100 ms processing latency remains constant throughout tracking of the ball's motion. Could it be that, once the stimulus has initially been processed, subsequent processing of the ball becomes faster so that the percept of it effectively lags less far behind the present? If this happened, it would have consequences for perception: the ball would appear to be moving faster than it actually was, as the percept gradually caught up with the actual present location of the ball. In fact there is evidence consistent with that possibility. Runeson (1974) found that objectively constant speed of motion is not perceived

as such. If an object is perceived to start moving with instantaneous acceleration to a constant velocity, observers tend to report the impression that the object starts moving quickly and then abruptly slows down. The kind of motion that they perceive as constant velocity involves a rate of acceleration that is initially rapid and then slows. Runeson proposed that this indicated a propensity to perceive motion as biologically plausible, at least when it is not perceived as externally caused (e.g. by contact from another object). However, both effects are compatible with more rapid processing of stimulus information after the initial stimulus is presented (see also Bachmann, 1989; Bachmann et al., 2003). In the case of constant velocity, a long latency for the initial percept of the stimulus followed by a rapid reduction in processing latency to an asymptotic value would create the impression of rapid initial motion followed by slowing to constant velocity. In the case of accelerating motion, the motion would be perceived as at constant velocity if the rate of acceleration in the stimulus object ran parallel to the rate of reduction in processing latency. Both effects are also what would be expected if an extrapolation to the present was being constructed after an initial unpredicted stimulus became predictable.

4. It has been argued that postdiction is involved in at least some versions of the flash-lag effect (Arstila, 2015; Eagleman & Sejnowski, 2000, 2007). However, the specific postdictive effect proposed has been that the localization of the moving stimulus is influenced by its motion for as much as 80 ms after the flash has occurred, which can account for misperception of the spatial relation between the flash and the moving stimulus without invoking extrapolation to the present. It has not previously been considered that extrapolation to the present might itself involve postdiction, in such a way as to allow reconstruction of the spatial location of the flash.

5. One plausible mechanism for updating a percept would be a comparator process. Comparator processes operate in action monitoring, and function by taking in a model of the planned action and its consequences, and sensory feedback about the actual action and

consequences, and detecting discrepancies which alert the system to a need to modify behaviour (e.g. Blakemore et al., 2002; Fourneret & Jeannerod, 1998). A functionally similar model could operate to detect discrepancies between an ongoing perceptual model and new input (see also Galletti & Fattori, 2003; Hohwy, 2013).

References

Alais, D., & Burr, D. C. (2003). The "flash-lag" effect occurs in audition and cross-modally. *Current Biology, 13,* 59-63.

Arnold, D. H., Ong, Y., & Roseboom, W. (2009). Simple differential latencies modulate, but do not cause the flash-lag effect. *Journal of Vision, 9,* (5), No. 4.

Arrighi, R., Alais, D., & Burr, D. C. (2005). Neural latencies do not explain the auditory and audio-visual flash-lag effect. *Vision Research, 45,* 2917-2925.

Arstila, V. (2015). Keeping postdiction simple. *Consciousness and Cognition, 38,* 205-216.

Bachmann, T. (1989). Microgenesis as traced by the transient paired-forms paradigm. *Acta Psychologica, 70,* 3-17.

Bachmann, T., Breitmeyer, B. G., & Öğmen, H., (2011). *The Experimental Phenomena of Consciousness: A Brief Dictionary*. Oxford: Oxford University Press.

Bachmann, T., Luiga, I, Põder, E., & Kalev, K. (2003). Perceptual acceleration of objects in stream: evidence from flash-lag displays. *Consciousness and Cognition, 12,* 279-297.

Bagattini, C., Mazzi, C., & Savazzi, S. (2015). Waves of awareness for occipital and parietal phosphenes perception. *Neuropsychologia, 70,* 114-125.

Bahill, A. T., & Karnavas, W. J. (1993). The perceptual illusion of baseball's rising fastball and breaking curveball. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 3-14.

Bahill, A. T., & LaRitz, T. (1984). Why can't batters keep their eyes on the ball? *American Scientist, 72,* 249-253.

Bar, M., (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15,* 600-609.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämälainen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006).

Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of America, 103,* 449-454.

Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The '*when*' pathway of the right parietal lobe. *Trends in Cognitive Sciences, 11,* 204-210.

Baylor, D.A., & Hodgkin, A. L. (1974). Changes in time scale and sensitivity in turtle photoreceptors. *Journal of Physiology, 242,* 729-758.

Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature, 398,* 334-338.

Bedell, H. E., Tong, J., & Aydin, M. (2010). The perception of motion smear during eye and head movements. *Vision Research, 50,* 2692-2701.

Bieniek, M. M., Bennett, P. J., Sekuler, A. B., & Rousselet, G. A. (2016). A robust and representative lower bound on object processing speed in humans. *European Journal of Neuroscience, 44,* 1804-1814.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences, 6,* 237-242.

Borghuis, B. G., & Leonardo, A. (2015). The role of motion extrapolation in amphibian prey capture. *Journal of Neuroscience, 35,* 15430-15441.

Brenner, E., & Smeets, J. B. J. (1997). Fast responses of the human hand to changes in target position. *Journal of Motor Behavior, 29,* 297-310.

Brenner, E., & Smeets, J. B. J. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Research, 40,* 1645-1648.

Brenner, E., & Smeets, J. B. J. (2015). How people achieve their amazing temporal precision in interception. *Journal of Vision, 15* (3), No. 8.

Briscoe, R. E. (2010). Perceiving the present: systematization of illusions or illusion of systematization? *Cognitive Science, 34,* 1530-1542.

Briscoe, R., & Schwenkler, J. (2015). Conscious vision in action. *Cognitive Science, 39,* 1435-1467.

Brouwer, A.-M., Thornton, I. M., & Franz, V. H. (2005). Forward displacement in grasping and visually judging pliers. *Visual Cognition, 12,* 800-816.

Brown, L. N., & Sainsbury, R. S. (2002). Age- and sex-related differences in temporal judgments to visual stimuli: support for hemispheric equivalence. *Perception and Psychophysics, 64,* 693-702.

Burr, D. C. (1980). Motion smear. *Nature, 284,* 164-165.

Burr, D. C., & Ross, J. (2002). Direct evidence that "speedlines" influence motion mechanisms. *Journal of Neuroscience, 22,* 8661-8664.

Burr, D. C., & Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Research, 41,* 1891-1899.

Burr, D. C., & Thompson, P. (2011). Motion psychophysics: 1985-2010. *Vision Research, 51,* 1431-1456.

Cai, R. H., & Schlag, J. (2001). A new form of illusory conjunction between color and shape. *Journal of Vision, 1,* No. 127.

Caniard, F., Bülthoff, H. H., & Thornton, I. M. (2015). Action can amplify motion-induced illusory displacement. *Frontiers in Human Neuroscience, 8,* No. 1058, 1-14.

Castiello, U., Paulignon, Y., & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness. *Brain, 114,* 2639-2655.

Cavanagh, P. (1997). Predicting the present. *Nature, 386,* 19-20.

Cavanagh, P. (2011). Visual Cognition. *Vision Research, 51,* 1538-1551.

Cavanagh, P., & Anstis, S. (2013). The flash grab effect. *Vision Research, 91,* 8-20.

Cellini, C., Scocchia, L., & Drewing, K. (2016). The buzz-lag effect. *Experimental Brain Research, 234,* 2849-2857.

Cesqui, B., Mezzetti, M., Lacquaniti, F., & d'Avella, A. (2015). Gaze behavior in one-handed catching and its relation with interceptive performance: what the eyes can't tell. *Plos One, 10,* e0119445.

Chakravarti, R., & VanRullen, R. (2012). Conscious updating is a rhythmic process. *Proceedings of the National Academy of Sciences of America, 109,* 10599-10604.

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.

Changizi, M. A., & Widders, D. M. (2002). Latency correction explains the classic geometrical illusions. *Perception, 31,* 1241-1262.

Changizi, M. A., Hsieh, A., Nijhawan, R., Kanai, R., & Shimojo, S. (2008). Perceiving the present and a systematization of illusions. *Cognitive Science, 32,* 459-503.

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Attention, Perception, & Psychophysics, 75,* 790-811.

Choi, H., & Scholl, B. J. (2006). Perceiving causality after the fact: postdiction in the temporal dynamics of causal perception. *Perception, 35,* 385-399.

Clay, E. R. (1882). *The Alternative: A Study in Psychology* (2nd Ed.). London: Macmillan.

Conrey, B., & Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *Journal of the Acoustical Society of America, 119,* 4065-4073.

Craig, J. C., & Busey, T. A. (2003). The effect of motion on tactile and visual temporal order judgments. *Perception and Psychophysics, 65,* 81-94.

Danquah, A. N., Farrell, M. J., & O'Boyle, D. J. (2008). Biases in the subjective timing of perceptual events: Libet et al. (1983) revisited. *Consciousness and Cognition, 17,* 616-627.

D'Ausilio, A., Jarmolowska, J., Busan, P., Bufalari, I., & Craighero, L. (2011). Tongue corticospinal modulation during attended verbal stimuli: priming and coarticulation effects. *Neuropsychologia, 49,* 3670-3676.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70,* 200-227.

Dehaene, S., Charles, L., King, J.-R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology, 25,* 76-84.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5,* e260.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117-1121.

Dennett, D. C. (1991). *Consciousness Explained*. New York: Little Brown.

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+.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research, 31,* 1619-1626.

Didierjean, A., % Marmèche, E. (2005). Anticipatory representation of visual basketball scenes by novice and expert players. *Visual Cognition, 12,* 265-283.

Di Lollo, V. (2012). The feature-binding problem is an ill-posed problem. *Trends in Cognitive Sciences, 16,* 317-321.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129,* 481-507.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception, 9,* 719-721.

Donaldson, I. M. L. (1979). Responses in cat suprasylvian cortex (Clare Bishop Area) to stretch of extraocular muscles. *Journal of Physiology, 296,* 60P-61P.

Donaldson, I. M. L. (2000). The functions of the proprioceptors of the eye muscles. *Philosophical Transactions of the Royal Society B, 355*, 1685-1754.

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in pariental cortex by intended eye movements. *Science, 255,* 90-92.

Eagleman, D. M. (2010). How does the timing of neural signals map on to the timing of perception? In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (pp. 216-231). Cambridge University Press.

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

Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: a unified explanation for the flash-lag, flash-drag, flash-jump, and Fröhlich illusions. *Journal of Vision*, *7*(4), No. 3.

Efron, R. (1963). The effect of handedness on the perception of temporal order. *Brain, 86,* 261-284.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345-352.

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2008). The spatiotemporal profile of processing leading up to visual perception. *Journal of Vision, 8,* (1), No. 12.

Feinkohl, A., Locke, S. M., Leung, J., & Carlile, S. (2014). The effect of velocity on auditory representational momentum. *Journal of the Acoustical Society of America, 136,* EL20-EL25.

Finke, R. A., Freyd, J. J., & Shyi, G. C.-W. (1986). Implied velocity and acceleration induce transformations of visual memory. *Journal of Experimental Psychology: General, 115,* 175-188.

Fourneret, P., & Jeannerod, M. (1998). Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia, 36,* 1133-1140.

Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., Leff, A., & Driver, J. (2013). Sight and sound out of synch: fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex, 49,* 2875-2887.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception and Psychophysics, 33,* 575-581.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94,* 427-438.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 126-132.

Freyd, J. J., Pantzer, T. M., & Cheng, J. L. (1988). Representing statics as forces in equilibrium. *Journal of Experimental Psychology: General, 117,* 395-407.

Futterweit, L. R., & Beilin, H. (1994). Recognition memory for movement in photographs: a developmental study. *Journal of Experimental Child Psychology, 57,* 163-179.

Galletti, C., & Fattori, P. (2003). Neuronal mechanisms for detection of motion in the field of view. *Neuropsychologia, 41,* 1717-1727.

Garrod, S., & Pickering, M. J. (2015). The use of content and timing to predict turn transitions. *Frontiers in Psychology, 6,* No. 751.

Geffen, G., Rosa, V., & Luciano, M. (2000a). Effects of preferred hand and sex on the perception of tactile simultaneity. *Journal of Clinical and Experimental Neuropsychology, 22,* 219-231.

Geldard, F. A., & Sherrick, C. E. (1972). The cutaneous "rabbit": a perceptual illusion. *Science, 178,* 178-179.

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision, 7,* (8), No. 8.

Getzmann, S., & Lewald, J. (2009). Constancy of target velocity as a critical factor in the emergence of auditory and visual representational momentum. *Experimental Brain Research, 193,* 437-443.

Getzmann, S., Lewald, J., & Guski, R. (2004). Representational momentum in spatial hearing. *Perception, 33,* 591-599.

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: behavioural and neurophysiological evidence. *Visual Cognition, 12,* 878-892.

Gollisch, T., & Meister, M. (2010). Eyes smarter than scientists believed: neural computations in circuits of the retina. *Neuron, 65,* 150-164.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15,* 20-25.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349,* 154-156.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature, 199,* 678-680.

Groen, I. I. A., Ghebreab, S., Prins, H., Lamme, V. A. F., & Scholte, H. S. (2013). From image statistics to scene gist: evoked neural activity reveals transition from low-level natural image structure to scene category. *Journal of Neuroscience, 33,* 18814-18824.

Grush, R. (2004).The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences, 27,* 377+.

Halliday. A. M., & Mingay, R. (1964). On the resolution of small time intervals and the effect of conduction delays on the judgement of simultaneity. *Quatrterly Journal of Experimental Psychology, 16,* 35-46.

Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Experimental Brain Research, 217,* 125-136.

He, D., Mo, C., & Fang, F. (2017). Predictive feature remapping before saccadic eye movements. *Journal of Vision, 17,* (5), No. 14.

Heinrich, A., & Schneider, B. (2006). Age-related changes in within- and between-channel gap detection using sinusoidal stimuli. *Journal of the Acoustical Society of America, 119,* 2316-2326.

Herzog, M. H., Kammer, T., & Scharnowski, F. (2016). Time slices: what is the duration of a percept? *Plos Biology, 14,* e1002433.

Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791-804.

Hogendoorn, H., Verstraten, F. A. J., & Cavanagh, P. (2015). Strikingly rapid neural basis of motion-induced position shifts revealed by high temporal-resolution EEG pattern classification. *Vision Research, 113,* 1-10.

Hohwy, J. (2013). *The Predictive Mind.* Oxford: Oxford University Press.

Holcombe, A. O. (2009). Seeing slow and seeing fast: two limits on perception. *Trends in Cognitive Sciences, 13,* 216-221.

Howard, C. J., Masom, D., & Holcombe, A. O. (2011). Position representations lag behind targets in multiple object tracking. *Vision Research, 51,* 1907-1919.

Hubbard, T. L. (1990). Cognitive representation of linear motion: possible direction and gravity effects in judged displacement. *Memory and Cognition, 18,* 299-309.

Hubbard, T. L. (1993). The effect of context on visual representational momentum. *Memory and Cognition, 21,* 103-114.

Hubbard, T. L. (1995). Auditory representational momentum: surface form, direction, and velocity effects. *American Journal of Psychology, 108,* 255-274.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: a review of the findings. *Psychonomic Bulletin and Review, 12,* 822-851.

Hubbard, T. L. (2013). Do the flash-lag effect and representational momentum involve similar extrapolations? *Frontiers in Psychology, 4,* No. 290.

Hubbard, T. L. (2014). The flash-lag effect and related mislocalizations: findings, properties, and theories. *Psychological Bulletin, 140,* 308-338.

Hubbard, T. L. (2015). The varieties of momentum-like experience. *Psychological Bulletin, 141,* 1081-1119.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception and Psychophysics, 44,* 211-221.

Hubbard, T. L., & Ruppel, S. E. (1999). Representational momentum and landmark attraction effects. *Canadian Journal of Experimental Psychology, 53,* 242-256.

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

Ichikawa, M., & Masakura, Y. (2010). Reduction of the flash-lag effect in terms of active observation. *Attention, Perception, and Psychophysics*, *72,* 1032-1044.

Ipser, A., Agolli, V., Bajraktari, A., Al-Alawi, F., Djaafara, N., & Freeman, E. D. (2017). Sight and sound persistently out of synch: stable individual differences in audiovisual synchronisation revealed by implicit measures of lip-voice integration. *Scientific Reports, 7,* No. 46413.

Itier, R. J., Alain, C., Kovacevic, N., & McIntosh, A. R. (2007). Explicit versus implicit gaze processing assessed by ERPs. *Brain Research, 1177,* 79-89.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2,* 194-203.

James, W. (1890). *The Principles of Psychology*. New York, NY: Holt.

Johnston, J., & Lagnado, L. (2015). General features of the retinal connectome determine the computation of motion anticipation. *ELife, 4,* No. 06250.

Joordens, S., Spalek, T. M., Razmy, S., & van Duijn, M. (2004). A clockwork orange: compensation opposing momentum in memory for location. *Memory and Cognition, 32,* 39-50.

Jordan, J. S., & Hunsinger, M. (2008). Learned patterns of action-effect anticipation contribute to the spatial displacvement of continuously moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 113-124.

Kahan, T. A., & Enns, J. T. (2014). Long-term memory representations influence perception before edges are assigned to objects. *Journal of Experimental Psychology: General, 143,* 566-574.

Kanai, R., Carlson, T. A., Verstraten, F. A. J., & Walsh, V. (2009). Perceived timing of new objects and feature changes. *Journal of Vision, 9,* (7), No. 5.

Kang, P., & Shevell, S. K. (2012). Feature binding of a continuously changing object. *Journal of the Optical Society of America A, 29,* A128-A132.

Kawabe, T., Yamada, Y., & Miura, K. (2007). Memory displacement of an object with motion lines. *Visual Cognition, 15,* 305-321.

Kerzel, D., & Gegenfurter, K. R. (2003). Neuronal processing delays are compensated in the sensorimotor branch of the visual system. *Current Biology, 13,* 1975-1978.

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and Cognition, 72,* 244-254.

Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The flash-lag effect as a motion-based predictive shift. *Plos Computational Biology, 13,* e1005068.

Khuu, S. K., Kidd, J. C., & Errington, J. A. (2010). The effect of motion adaptation on the position of elements in the visual saltation illusion. *Journal of Vision, 10 (14),* No. 19.

Kilgard, M. P., & Merzenich, M. M. (1995). Anticipated stimuli across skin. *Nature, 373,* 663.

Kim, S.-H., Feldman, J. & Singh, M. (2013). Perceived causality can alter the perceived trajectory of apparent motion. *Psychological Science, 24,* 575-582.

Kiss, M., & Eimer, M. (2008). ERPs reveal subliminal processing of fearful faces. *Psychophysiology, 45,* 318-326.

Kohler, P. J., Cavanagh, P., & Tse, P. U. (2015). Motion-induced position shifts are influenced by global motion, but dominated by component motion. *Vision Research, 110,* 93-99.

Koivisto, M., & Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neuroscience and Biobehavioral Reviews, 34,* 922-934.

Kolers, P., & von Grünau, M. (1976). Shape and color in apparent motion. *Vision Research, 16,* 329-335.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends in Neurosciences, 24,* 335-339.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31,* 32-59.

Lachter, J., & Durgin, F. H. (1999). Metacontrast masking functions: a question of speed? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 936-947.

Lachter, J., Durgin, F. H., and Washington, T. (2000). Disappearing percepts: evidence for retention failure in metacontrast masking. *Visual Cognition, 7,* 269-279.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience, 23,* 571-579.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience, 3,* 1340-1345.

Lakatos, P., Pincze, Z., Fu, K.-M. G., Javitt, D. C., Karmos, G., & Schroeder, C. E. (2005). Timing of pure tone and noise-evoked responses in macaque auditory cortex. *Neuroreport, 16,* 933-937.

Lewis, C. E., & McBeath, M. K. (2004). Bias to experience approaching motion in a three-dimensional virtual environment. *Perception, 33,* 259-276.

Li, Y., Zhong, W., Chen, Q., & Mo, L. (2013). Object detection is completed earlier than object categorization: evidence from LRP and N200. *Psychophysiology, 50,* 1255-1262.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain, 106,* 634-642.

Liegeois-Chauvel, C., Musolino, A., & Chauvel, P. (1991). Localization of the primary auditory area in man. *Brain, 114,* 139-153.

Lisberger, S. G. (2010). Visual guidance of smooth-pursuit eye movements: sensation, action, and what happens in between. *Neuron, 66,* 477-491.

Lloyd, D. (2012). Neural correlates of temporality: default mode variability and temporal awareness. *Consciousness and Cognition, 21,* 695-703.

Luo, Q., Holroyd, T., Jones, M., Hendler, T., & Blair, J. (2007). Neural dynamics for facial threat processing as revealed by gamma band synchronization using MEG. *Neuroimage, 34,* 839-847.

Maess, B., Mamashli, F., Obleser, J., Helle, L., & Friederici, A. D. (2016). Predictive signatures in the brain: semantic pre-activation during language comprehension. *Frontiers in Human Neuroscience, 10,* No. 591.

Marinovic, W., & Arnold, D. H. (2013). An illusory distortion of moving form driven by motion deblurring. *Vision Research, 88,* 47-54.

Marinovic, W., Plooy, A. M., & Arnold, D. H. (2012). The influence of visual motion on interceptive actions and perception. *Vision Research, 60,* 73-78.

Martínez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., Wong, E. C., Hinrichs, H., Heinze, H. J., & Hillyard, S. A. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience, 2,* 364-369.

Matin, L., Boff, K. R., & Pola, J. (1976). Vernier offset produced by rotary target motion. *Perception & Psychophysics, 20,* 138-142.

Matthews, P. B. C. (1982). Where does Sherrington's 'muscular sense' originate? Muscles, joints, corollary discharges? *Annual Review of Neuroscience, 5,* 189-218.

Maus, G. W., Khurana, B., & Nijhawan, R. (2010). History and theory of the flash-lag: past, present, and future. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in Perception and Action* (pp. 477-499). Cambridge: Cambridge University Press.

McCloskey, D. I. (1981). Corollary discharges: motor commands and perception. In V. B. Brooks (Ed.), *Handbook of Physiology. II. Nervous System* (pp. 1415-1447). Baltimore, MD: Williams and Wilkins.

McKee, S. P., & Welch, L. (1985). Sequential recruitment in the discrimination of velocity. *Journal of the Optical Society of America A, Optics and Image Science, 2,* 243-251.

McLeod, P., & Jenkins, S. (1991). Timing accuracy and decision time in high-speed ball games. *International Journal of Sport Psychology, 22,* 279-295.

Mégevand, P., Molholm, S., Nayak, A., & Foxe, J. J. (2013). Recalibration of the multisensory temporal window of integration results from changing task demands. *Plos One, 8,* e71608.

Merikle, P. M., Joordens, S., & Stolz, J. A. (1997). Measuring the relative magnitude of unconscious influences. *Consciousness and Cognition, 4,* 422-439.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks, 9,* 1265-1279.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences, 6,* 414-417.

Morel, S., Beaucousin, V., Perrin, M., & George, N. (2012). Very early modulation of brain responses to neutral faces by a single prior association with an emotional context: evidence from MEG. *Neuroimage, 61,* 1461-1470.

Motes, M. A., Hubbard, T. L., Courtney, J. R., & Rypma, B. (2008). A principal components analysis of dynamic spatial memory biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1076-1083.

Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London B: Biological Sciences, 264,* 393-399.

Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London B: Biological Sciences, 264,* 1407-1414.

Müsseler, J., Stork, S., & Kerzel, D. (2002). Comparing mislocalisations with moving stimuli: the Frölich effect, the flash-lag, and representational momentum. *Visual Cognition, 9, 120-138.*

Müsseler, J., & Tigglebeck, J. (2013). The perceived onset position of a moving target: effects of trial contexts are evoked by different attentional allocations. *Attention, Perception & Psychophysics, 75,* 349-357.

Nakamoto, H., Mori, S., Ikudome, S., Unenaka, S., & Imanaka, K. (2015). Effects of sports expertise on representational momentum during timing control. *Attention, Perception, & Psychophysics, 77,* 961-971.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature, 395,* 894-896.

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

Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature, 386,* 66-69.

Nijhawan, R. (2008). Visual prediction: psychophysics and neurophysiology of compensation for time delays. *Behavioral and Brain Sciences, 31,* 179+.

Nijhawan, R., & Kirschfeld, K. (2003). Analogoue mechanisms compensate for neural delays in the sensory and the motor pathways: evidence from motor flash-lag. *Current Biology, 13,* 749-753.

Nijhawan, R., & Wu, S. (2009). Compensating time delays with neural predictions: are predictions sensory or motor? *Philosophical Transactions of the Royal Society A, 369,* 1063-1078.

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,* 2109-2128.

Palmer, S. E., Marre, O., Berry, M. J., & Bialek, W. (2015). Predictive information in a sensory population. *Proceedings of the National Academy of Sciences of America, 112,* 6908-6913.

Parsons, B. D., Novich, S. D., & Eagleman, D. M. (2013). Motor-sensory recalibration modulates perceived simultaneity of cross-modal events at different distances. *Frontiers in Psychology, 4,* No. 46.

Perrinet, L. U., Adams, R. A., & Friston, K. J. (2014). Active inference, eye movements, and oculomotor delays. *Biological Cybernetics, 108,* 777-801.

Petrini, K., Russell, M., & Pollick, F. (2009). When knowing can replace seeing in audiovisual integration of actions. *Cognition, 110,* 432-439.

Pitcher, D., Goldhaber, T., Duchaine, B., Walsh, V., & Kanwisher, N. (2012). Two critical and functionally distinct stages of face and body perception. *Journal of Neuroscience, 7,* 15877-15885.

Pitts, M. A., Metzler, S., & Hillyard, S. A. (2014). Isolating neural correlates of conscious perception from neural correlates of reporting one's perception. *Frontiers in Psychology, 5,* No. 1078.

Prablanc, C., & Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *Journal of Neurophysiology, 67,* 455-469.

Prinzmetal, W., Shimamura, A., & Mikolinski, M. (2001). The Ponzo illusion and the perception of orientation. *Perception and Psychophysics, 63,* 99-114.

Puroshothaman, G., Patel, S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396,* 424.

Railo, H., Koivisto, M., & Revonsuo, A. (2011). Tracking the processes behind conscious perception: a review of event-related potential correlates of visual consciousness. *Consciousness and Cognition, 20,* 972-983.

Ramachandran, V. S., & Anstis, S. M. (1986). The perception of apparent motion. *Scientific American, 254,* 102-109.

Ramachandran, V. S., & Anstis, S. M. (1990). Illusory displacement of equiluminous kinetic edges. *Perception, 19,* 611-616.

Ramachandran, V. S., Madhusudhan Rao, V., & Vidyasagar, T. R. (1974). Sharpness constancy during movement perception. *Perception, 3,* 97-98.

Rao, H. M., Mayo, J. P., & Sommer, M. A. (2016). Circuits for presaccadic visual remapping. *Journal of Neurophysiology, 116,* 2624-2636.

Regan, D. (1992). Visual judgements and misjudgements in cricket and the art of flight. *Perception, 21,* 91-115.

Regan, D. (1997). Visual factors in hitting and catching. *Journal of Sports Science, 15,* 533-558.

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7,* 17-42.

Rodrigues, S. T., Vickers, J. N., & Williams, A. M. (2002). Head, eye and arm coordination in table tennis. *Journal of Sports Sciences, 20,* 187-200.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35,* 76-100.

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. *Vision Research*, *51,* 1297-1311.

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 7,* 629-630.

Rovelli, C. (2018). *The Order of Time.* London: Allen Lane.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50,* 696-735.

Scharnowski, F., Hermens, F., Kammer, T., Öğmen, H., & Herzog, M. H. (2007). Feature fusion reveals slow and fast visual memories. *Journal of Cognitive Neuroscience, 19,* 632-641.

Scharnowski, F., Rüter, J., Jolij, J., Hermens, F., Kammer, T., & Herzog, M. H. (2009). Long-lasting modulation of feature integration by transcranial magnetic stimulation. *Journal of Vision, 9,* (6), No. 1.

Schmiedchen, K., Freigang, C., Rübsamen, R., & Richter, N. (2013). A comparison of visual and auditory representational momentum in spatial tasks. *Attention, Perception, and Psychophysics, 75,* 1507-1519.

Schubotz, R. I., & von Cramon, D. Y. (2001). Functional organization of the lateral premotor cortex: fMRI reveals different regions activated by anticipation of object properties, location and speed. *Cognitive Brain Research, 11,* 97-112.

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *Neuroimage, 15,* 787-796.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research, 38,* 3569-3582.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience, 8,* 1391-1400.

Shafto, J. P., & Pitts, M. A. (2015). Neural signatures of conscious face perception in an inattentional blindness paradigm. *Journal of Neuroscience, 35,* 10940-10948.

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly pulsed ones. *Nature Neuroscience, 3,* 489-495.

Shi, Z., & Nijhawan, R. (2012). Motion extrapolation in the central fovea. *Plos One, 7,* e33651.

Shimojo, S. (2014). Postdiction: its implications on visual awareness, hindsight, and sense of agency. *Frontiers in Psychology, 5,* No. 196.

Simpson, W. A. (1994). Temporal summation of visual motion. *Vision Research, 34,* 2547-2559.

Snowden, R. J., & Braddick, O. J. (1991). The temporal integration and resolution of velocity signals. *Vision Research, 31,* 907-914.

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: when two locations are better than one. *Perception and Psychophysics, 65,* 318-328.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130,* 799-832.

Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43,* 482-489.

Stephens, G. J., Honey, C. J., & Hasson, U. (2013). A place for time: the spatiotemporal structure of neural dynamics during natural audition. *Journal of Neurophysiology, 110,* 2019-2026.

Sun, H.-M., Simon-Dacj, S. L., Gordon, R. D., & Teder, W. A. (2011). Contextual influences on rapid object categorization in natural scenes. *Brain Research, 1398,* 40-54.

Tapia, E., & Beck, D. M. (2014). Probing feedforward and feedback contributions to awareness with visual masking and transcranial magnetic stimulation. *Frontiers in Psychology, 5,* No. 1173.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520-522.

Tong, J., Patel, S. S., & Bedell, H. E. (2005). Asymmetry of perceived motion smear during head and eye movements: evidence for a dichotomous neural categorization of retinal image motion. *Vision Research, 45,* 1519-1524.

Tortosa, M. I., Lupiánez, J., & Ruz, M. (2013). Race, emotion, and trust: an ERP study. *Brain Research, 1494,* 44-55.

Tse, P. U., Whitney, D., Anstis, S., & Cavanagh, P. (2011). Voluntary attention modulates motion-induced mislocalization. *Journal of Vision, 11* (3), No. 12.

van de Grind (2002). Physical, neural, and mental timing. *Consciousness and Cognition, 11,* 241-262.

van der Lugt, A. H., Banfield, J. F., Osinsky, R, & Münte, T. F. (2012). Brain potentials show rapid activation of implicit attitudes towards young and old people. *Brain Research, 1429,* 98-105.

van Heijnsbergen, C. C. R. J., Meeren, H. K. M., Grèzes, J., & de Gelder, B. (2007). Rapid detection of fear in body expressions, an ERP study. *Brain Research, 1186,* 233-241.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia, 45,* 598-607.

Vaughn, D. A., & Eagleman, D. M. (2013). Spatial warping by oriented line detectors can counteract neural delays. *Frontiers in Psychology, 4,* No. 794.

Vollrath, M., Kazenwadel, J., & Krüger, H.-P. (1992). A universal constant in temporal segmentation of human speech. *Naturwissenschaften, 79,* 479-480.

Von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). *Naturwissenschaft, 37,* 464-476.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Attention, Perception, and Psychophysics, 72,* 871-884.

Westerink, J. H. D. M., & Teunissen, K. (1995). Perceived sharpness in complex moving images. *Displays, 16,* 89-97.

White, P. A. (2012). The experience of force: the role of haptic experience of forces in visual perception of object motion and interactions, mental simulation, and motion-related judgments. *Psychological Bulletin, 138,* 589-615.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends in Cognitive Sciences, 6,* 211-216.

Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: shifting the perceived position of remote stationary objects. *Nature Neuroscience, 3,* 954-959.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1,* 656-657.

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

Whitney, D., Wurnitsch, N., Hontiveros, B., & Louie, E. (2008). Perceptual mislocalization of bouncing balls by professional tennis referees. *Current Biology, 18,* R947-R949.

Wiegrebe, L., & Krumbholz, K. (1999). Temporal resolution and temporal masking properties of transient stimuli: data and an auditory model. *Journal of the Acoustical Society of America*, *105*, 2746-2756.

Wittmann, M. (2011). Moments in time. *Frontiers in Integrative Neuroscience, 5,* No. 66.

Wutz, A., & Melcher, D. (2014). The temporal window of individuation limits visual capacity. *Frontiers in Psychology, 5,* No. 952.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature, 414,* 302-305.

Yoshimoto, S., Uchida-Ota, M., & Takeuchi, T. (2014). The reference frame of visual motion priming depends on underlying motion mechanisms. *Journal of Vision*, *14* (1), No. 10.

Zago, M., & Lacquaniti, F. (2005). Cognitive, perceptual and action-oriented representations of falling objects. *Neuropsychologia, 43,* 178-188.

Zago, M., McIntyre, J., Senot, P., & Lacquaniti, F. (2009). Visuo-motor coordination and internal models for object interception. *Experimental Brain Research, 192,* 571-604.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research, 152,* 198-210.

Zeki, S. (2001). Localization and globalization in conscious vision. *Annual Review of Neuroscience, 24,* 57-86.

Zera, J., & Green, D. M. (1993). Detecting temporal onset and offset asynchrony in multicomponent complexes. *Journal of the Acoustical Society of America, 93,* 1038-1052.

Zimmermann, E., Morrone, M. C., & Burr, D. C. (2014). Buildup of spatial information over time and across eye-movements. *Behavioural Brain Research, 275,* 281-287.

Box 1

<u>Kinds of hypothesized temporal adjustments in perceptual processing mentioned in the review</u>

Extrapolation to the present: the hypothesis that processing latencies are compensated by using perceptual input information to generate a predictive model of what is out in the world now.

Extrapolation to the future: The hypothesis that perceptual input information is used to extrapolate a trajectory of a moving object to an anticipated location or time of interception.

Temporal sampling: the location of a moving object is sampled at intervals, not continuously, but the latency of the sampling process is neglected in the location estimate.

Differential latency: processing latencies are longer for static than for moving stimuli.

Temporal integration: any process in which a unified percept is generated by sampling over a period of time; specifically, temporal integration induces bias in computation of the location of a moving object.

Displacement: any consistent error in judged location at which a moving object disappeared.

Representational momentum: a specific form of displacement in which the judged location of disappearance is displaced forward in the direction of motion.

Representational gravity: downwards displacement possibly reflecting implicit knowledge of motion under gravity.

Deliberative compensation: an attempt, in post-perceptual processing, to overcome biases identified in judgments; potentially applicable to any of the research reviewed here.

Postdiction: adjustment of a percept in the light of subsequent perceptual input.

Motion extrapolation at the retina: speeding of retinal ganglion cell response associated with a stimulus moving across the retinal field.

Cross-modal synchronisation: integration processes that compensate for differential transmission and processing latencies in different modalities to generate subjectively synchronous percepts.

Figure captions

Figure 1: Hering illusion.

Figure 2: Ponzo illusion.

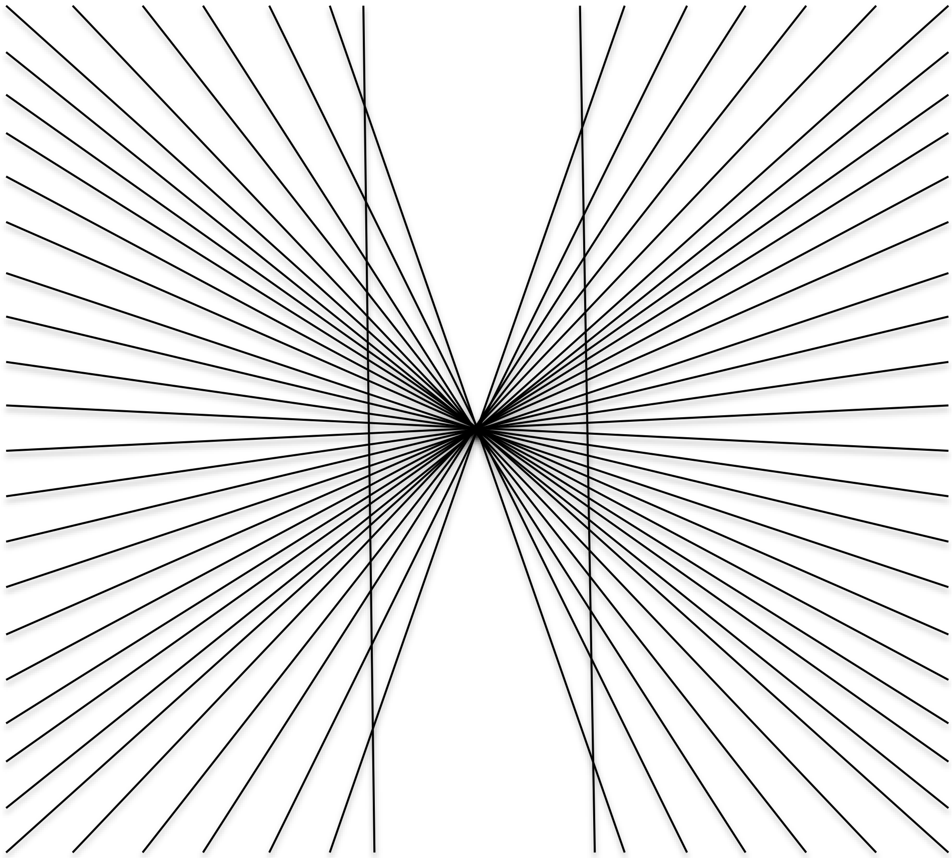

Figure 1: Hering illusion

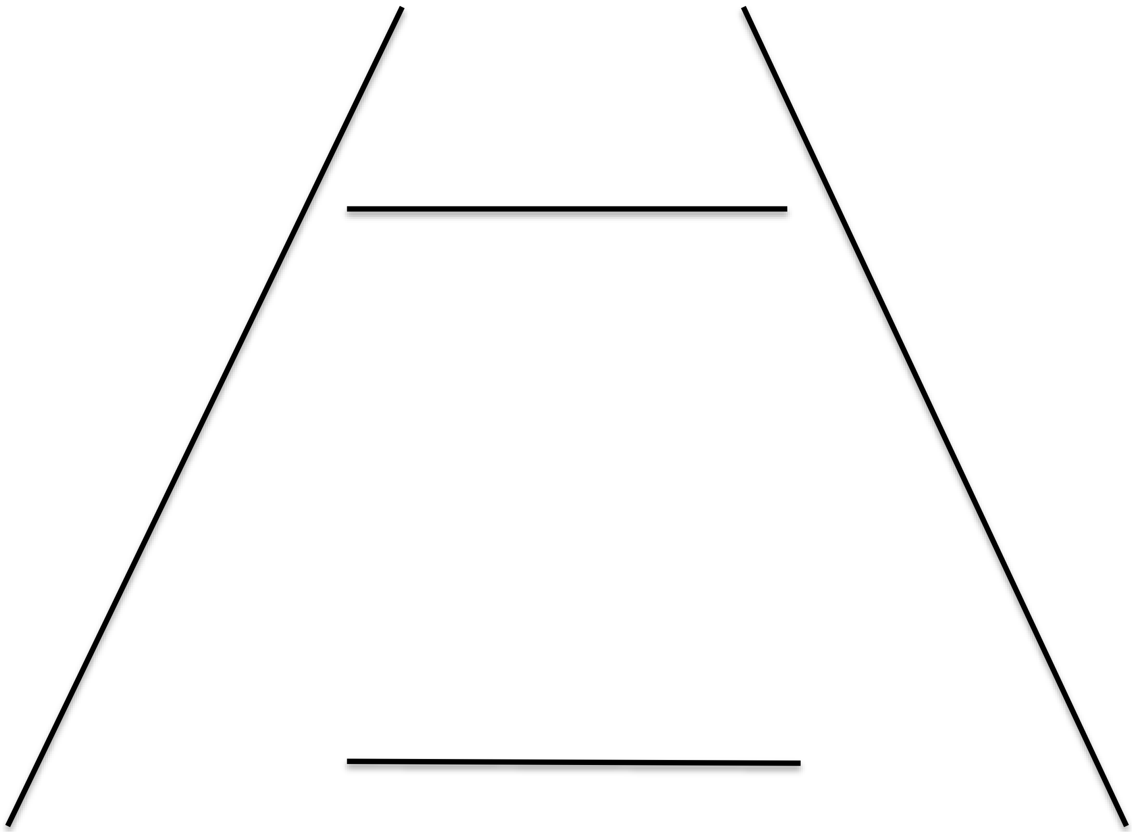

Figure 2: Ponzo illusion