

1 Title: Predictive coding of visual motion in both monocular and binocular human visual processing

2

3 Authors: Elle van Heusden<sup>1,2</sup>, Anthony M. Harris<sup>3,4</sup>, Marta Garrido<sup>4,5,6,7</sup> and Hinze Hogendoorn<sup>1,2</sup>

4 Affiliations:

5 1 Melbourne School of Psychological Sciences, The University of Melbourne, Australia

6 2 Helmholtz Institute, Dept of Experimental Psychology, Utrecht University , the Netherlands

7 3 Institute of Cognitive Neuroscience, University College London, United Kingdom,

8 4 Queensland Brain Institute, The University of Queensland, Australia

9 5 School of Mathematics and Physics, The University of Queensland, Australia

10 6 Centre for Advanced Imaging, The University of Queensland, Australia

11 7 Australian Research Council Centre of Excellence for Integrative Brain Function, The University of  
12 Queensland, Australia

13

14 Abstract:

15

16 Neural processing of sensory input in the brain takes time, and for that reason our awareness of visual events lags  
17 behind their actual occurrence. One way the brain might compensate to minimise the impact of the resulting delays  
18 is through extrapolation. Extrapolation mechanisms have been argued to underlie perceptual illusions in which  
19 moving and static stimuli are mislocalised relative to one another (such as the flash-lag and related effects).  
20 However, where in the visual hierarchy such extrapolation processes take place remains unknown. Here, we address  
21 this question by identifying monocular and binocular contributions to the flash-grab illusion. In this illusion, a brief  
22 target is flashed on a moving background that reverses direction. As a result, the perceived position of the target is  
23 shifted in the direction of the reversal. We show that the illusion is attenuated, but not eliminated, when the motion  
24 reversal and the target are presented dichoptically to separate eyes. This reveals extrapolation mechanisms at both  
25 monocular and binocular processing stages contribute to the illusion. We interpret the results in a hierarchical  
26 predictive coding framework, and argue that prediction errors in this framework manifest directly as perceptual  
27 illusions.

28

29

30 **Introduction**

31 Neural processing of sensory input in the brain takes time, and for that reason our awareness of visual events lags  
32 behind their actual occurrence. If the visual system did not somehow compensate for neural transmission delays,  
33 we would consistently mislocalise moving objects behind their actual position. Nevertheless, human observers are  
34 typically very accurate at localising moving objects, achieving near-zero lag when object trajectories are predictable  
35 (Brenner, Smeets, & de Lussanet, 1998). One explanation for how the brain might overcome its internal delay is  
36 through extrapolation: by exploiting knowledge about an object's past trajectory, the brain predicts its present  
37 position.

38 Although accurate interaction with moving objects could also be achieved by extrapolation in the motor system (e.g.  
39 Kerzel & Gegenfurtner, 2003), extrapolation mechanisms in the visual system have been hypothesized to underlie a  
40 class of visual illusions in which visual motion signals affect the perceived location of stationary objects. This includes  
41 the much-studied flash-lag effect, in which a moving object that is physically aligned with a stationary flash is  
42 perceived ahead of that flash (Nijhawan, 1994). In this interpretation, the brain extrapolates the position of the  
43 moving object along its expected trajectory to compensate for lag that would otherwise arise due to processing time.  
44 When the flash is presented aligned with the moving object, it is compared to the *extrapolated* position of the  
45 moving object, and hence appears to lag behind it.

46 In the years following Nijhawan's initial demonstrations of the flash-lag effect, numerous other motion-induced  
47 position shifts have been reported, including the flash-drag (Whitney & Cavanagh, 2000), flash-jump (Cai & Schlag,  
48 2001), and flash-grab (Cavanagh & Anstis, 2013) effects. Although the underlying mechanisms have been hotly  
49 debated (e.g. Eagleman & Sejnowski, 2000; B. Krekelberg, 2000; Patel, Ogmen, Bedell, & Sampath, 2000; Whitney &  
50 Murakami, 1998), convergent evidence points to an important role for predictive extrapolation mechanisms in  
51 causing these effects (Nijhawan, 2008). For instance, animal neurophysiology studies have demonstrated the  
52 existence of predictive extrapolation mechanisms in the retinae of salamanders, mice, and rabbits (Berry, Brivanlou,  
53 Jordan, & Meister, 1999; Schwartz et al., 2007), as well as in cat primary visual cortex (Jancke, Erlhagen, Schöner, &  
54 Dinse, 2004). In humans, it has been demonstrated that moving objects are extrapolated into regions of visual space  
55 where they could physically not be detected, such as the blind spot – ruling out explanations in terms of differential  
56 latencies (Maus & Nijhawan, 2008). Modeling studies have shown that a Bayesian model of perceived position that  
57 incorporates neural delays generates predictive position shifts such as seen in the flash-lag effect (Khoei, Masson, &  
58 Perrinet, 2017), and most recently an unsupervised predictive neural network exposed to natural video sequences  
59 (including motion) was found to have developed a pattern of response consistent with the flash-lag effect (Lotter,  
60 Kreiman, & Cox, 2018).

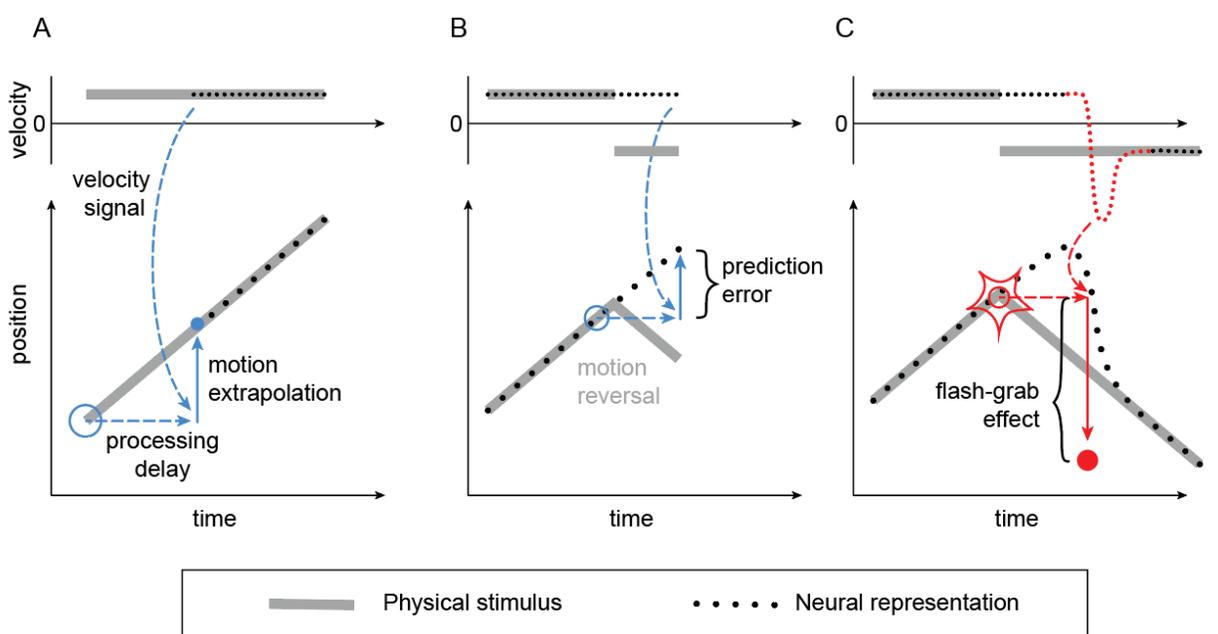
61 An important conceptual challenge to interpreting motion-induced position shifts as resulting from predictive  
62 mechanisms arises from the observation that the perceived position of a static event is biased primarily by motion

63 presented *after* that event, rather than before it. This led Eagleman & Sejnowski (2000) to coin the term *post-diction*,  
64 as a temporal counterpart to *prediction*. In this original post-diction account, events were essentially back-dated in  
65 perception, rewriting recent perceptual history. Several years later, the same authors presented a refined version of  
66 this model, in which local velocity signals integrated over a brief period *after* an event interact with local position  
67 signals to bias its perceived position (Eagleman & Sejnowski, 2007). Much has been made of what seems like reverse  
68 causality in the post-diction account, and the apparent contrast with predictive extrapolation mechanisms. However,  
69 the predictive model and the postdictive motion-biasing model are mechanistically the same, differing only in the  
70 time-window during which motion signals are integrated. Eagleman & Sejnowski (2007, p. 9) note that “Motion  
71 biasing will normally push objects closer to their true location in the world [...] by a clever method of updating signals  
72 that have become stale due to processing time”, which is precisely what motion extrapolation also does (Nijhawan,  
73 2008). Eagleman has more recently argued that prediction and postdiction cooperate to compensate for neural  
74 delays (Eagleman, 2008), simply because predictions by their nature sometimes do not come true, necessitating  
75 post-hoc revisions to the timeline of experience. Viewed more broadly, prediction and postdiction are simply two  
76 halves of the same mechanism, split along the line separating past from future. However, in the context of sensory  
77 processing, this line is artificial: due to neural delays, *all* cortical areas process information collected in the objective  
78 past. Although they seem polar opposites, predictive and postdictive accounts both push the representation of an  
79 object closer to its true location in the world at a given instant. Given the reality of neural processing delays, this  
80 means anticipating (i.e. predicting) the present.

81 Interestingly, this predictive rationale fits neatly with a more recent computational model aiming to unify position  
82 and motion perception (Kwon, Tadin, & Knill, 2015), which applied a Bayesian approach to another motion-position  
83 illusion (motion-induced position shifts; De Valois & De Valois, 1991; Ramachandran & Anstis, 1990). Kwon et al  
84 advocate a model in which motion and position judgments mutually interact to make optimal inferences about the  
85 generative causes underlying sensory signals. The model is implemented as a Kalman filter, and therefore the  
86 represented position at a given time explicitly depends only on velocity signals integrated before that time. However,  
87 the precise time-window over which signals are integrated relative to objective external time was not the focus of  
88 the model, nor does it invalidate the mechanistic similarities it shares with the account proposed earlier by Eagleman  
89 & Sejnowski (2007). Indeed, Eagleman & Sejnowski (2000) note that postdiction is commonplace in engineering,  
90 where it is simply known as smoothing. Most importantly, the model’s core feature – that it causes a position signal  
91 to be shifted in the direction of a motion signal – is the same: it is a predictive mechanism that causes anticipatory  
92 activation at the object’s future position.

93 A more recent illusion, the flash-grab effect, has provided the opportunity to study how predictive motion  
94 extrapolation mechanisms behave when motion vectors abruptly change, such that the anticipated future does not  
95 come true (Cavanagh & Anstis, 2013). In this illusion, a target is briefly flashed on a moving background as the motion  
96 unexpectedly reverses direction, which results in the perceived position of the flash being shifted in the direction of

97 the second motion sequence. Although neither Eagleman & Sejnowski (2007) nor Kwon et al (2015) made reference  
 98 to the flash-grab effect, the flash-grab effect can also be readily explained by the same mechanism. Figure 1 shows  
 99 schematically how this would work. A moving object is represented at a given level of processing with a certain delay.  
 100 It is possible to compensate for that delay by using information about the object's velocity to extrapolate the true  
 101 position of the object at that instant (e.g. (Krekelberg & Lappe, 2001; Nijhawan, 1994); see Figure 1A). When the  
 102 object reverses direction, it again takes time to detect the reversal, during which the object's position continues to  
 103 be extrapolated beyond the reversal point (Figure 1B). When sensory information about the object's actual  
 104 trajectory then becomes available, the represented position must rapidly shift from the predicted trajectory to the  
 105 new trajectory. This rapid shift in represented position equates to a brief spike in velocity (Figure 1C, upper plot).  
 106 Importantly, the key features in Figure 1 are not hypothetical: the overshoot in represented position and subsequent  
 107 acceleration to intercept the new trajectory exactly mirror population codes reported in the mouse and salamander  
 108 retina for such reversing stimuli (Schwartz et al., 2007). Although such mechanisms have not yet been directly  
 109 demonstrated in the brain itself, because prediction error signals arise already in the retina, they are passed on to  
 110 the rest of the visual processing hierarchy even if they would not be calculated there. The mechanism proposed by  
 111 Eagleman & Sejnowski (2007) then predicts that any stationary object flashed at the reversal point would interact  
 112 with this motion signal and be mislocalized. This is the effect we know as the flash-grab effect (Cavanagh & Anstis,  
 113 2013). Importantly, the magnitude of this mislocalization would be a direct reflection of how far into the future the  
 114 neural representation of the moving object has been extrapolated. As is evident in Figure 1A-C, the longer the  
 115 processing delay, the further into the (local) future a representation must be extrapolated in order to compensate,  
 116 and the longer it would take before a violation of that extrapolation is detected in that brain area. This would yield  
 117 a stronger velocity spike as the area "catches up", and a bigger flash-grab effect.



118

119 *Figure 1: Schematic illustration of extrapolation in the flash-grab effect. In each panel, the lower plot shows position*  
120 *as a function of time, and the upper plot shows velocity as a function of time. Solid grey traces indicate the properties*  
121 *of the physical stimulus as presented on the screen, and dotted black traces indicate (predictive) neural*  
122 *representations of the same stimulus, as demonstrated empirically in the retina by Schwartz et al (2007). (A). In order*  
123 *to accurately localize a moving object despite neural transmission delays, the visual system uses concurrent velocity*  
124 *signals to extrapolate the real-time position of the object (blue lines). (B). When an object unexpectedly reverses*  
125 *direction, at any given level of representation, some time elapses before the reversal is detected. During that time,*  
126 *the object will continue to be (erroneously) extrapolated into positions where it is never presented, creating a*  
127 *prediction error. (C). As the represented position shifts from the predicted trajectory to a new trajectory, the rapid*  
128 *shift in represented position creates a brief spike in the represented velocity (dotted red trace). If a (stationary) flash*  
129 *is presented at the same time as the reversal, then the position of the flash will interact with the (large) transient*  
130 *velocity signal and be mislocalised, resulting in the flash-grab effect (red lines).*

131 Consistent with this interpretation of the flash-grab effect as resulting from failed prediction, we recently  
132 demonstrated that the same neural mechanisms that cause the location of the target in the flash-grab effect to be  
133 misperceived, also influence saccades aimed at that target (Van Heusden, Rolfs, Cavanagh, & Hogendoorn, 2018).  
134 Most importantly, we showed that the degree of saccade error increased with increasing saccadic latency. This  
135 indicates that the visuomotor system was extrapolating the (stationary) target's position as if it was actually moving,  
136 confirming that mislocalisation in this illusion results from an extrapolation process. It can be readily appreciated  
137 from Figure 1 that if processing delays are larger (for whatever reason), then the original trajectory will be  
138 extrapolated further into the future, the object will move even further along its actual trajectory, and the total  
139 position error will be greater. The transient peak in velocity as the system adjusts will therefore also be greater, in  
140 turn leading to a larger flash-grab effect.

141 Where in the visual hierarchy the neural mechanisms responsible for extrapolation operate is still unknown. In  
142 animals, predictive neural mechanisms have been identified at multiple levels of the visual system, including the  
143 retina (Berry et al., 1999; Hosoya, Baccus, & Meister, 2005; Schwartz et al., 2007), lateral geniculate nucleus (Sillito,  
144 Jones, Gerstein, & West, 1994), primary visual cortex (Jancke et al., 2004), and V4 (Sundberg, Fallah, & Reynolds,  
145 2006). In humans, we recently demonstrated that the visual brain predicts the position of a moving object using an  
146 EEG classification paradigm (Hogendoorn & Burkitt, 2018a). This study revealed that for an object in apparent  
147 motion, the neural representation of the object's position is pre-activated when the object moves along a  
148 predictable trajectory. However, this was only true for neural representations evoked around 130 ms after stimulus  
149 presentation, whereas the latency of earlier neural position representations was not modulated by prediction. In  
150 contrast, a previous EEG study of the flash-grab effect revealed that the target's illusory position was represented in  
151 the EEG signal as early as 81 ms post-stimulus (Hogendoorn, Verstraten, & Cavanagh, 2015). These two studies

152 therefore seem to reveal extrapolation processes at different stages in the visual hierarchy. Hence, the limited  
153 evidence from humans is consistent with the evidence from animal neurophysiology.

154 With the cortical EEG reflecting the target's extrapolated position already at about 80 ms post-stimulus, the question  
155 is: where along the route from retina to cortex does this extrapolation take place? The vast majority of visual  
156 information reaches the cortex through the retino-geniculo-cortical pathway, passing from retina to the lateral  
157 geniculate nucleus of the thalamus before flowing on to the primary visual cortex (V1). Although there are alternative  
158 pathways to the cortex (a point to which we return in the discussion), given the severely limited timeframe, it is likely  
159 that the visual extrapolation mechanisms responsible for the flash-grab effect operate along the geniculate pathway.  
160 This would parallel the predictive mechanisms in the retina, LGN, and V1 revealed in animals, but it remains unknown  
161 whether (and if so, which of) these areas similarly carry out extrapolation in the human visual system.

162 In order to answer this question, here we make use of the fact that visual information from the two eyes does not  
163 converge until V1. Neurons that carry information either from the retina to LGN, or from LGN to V1, carry information  
164 from only one eye, with the first binocular neurons in the visual pathway located in V1 itself (Parker, 2007). We use  
165 the flash-grab effect, and employ dichoptic presentation to separate the different components of the flash-grab  
166 stimulus across the two eyes. In so doing, we prevent those components of the stimulus interacting at an early  
167 (monocular) stage of the visual hierarchy. We manipulate which components of the flash-grab stimulus sequence  
168 (motion prior to the flash, the flash itself, and the motion following the flash) are presented to which eye, and  
169 measure the strength of the resulting illusion. If the shift in the perceived position of a target presented in one eye  
170 is reduced when the flanking motion sequence is presented to the opposite eye, then this would be evidence that  
171 extrapolation mechanisms operate already at monocular stages of processing.

172 Here, we show that this is the case. The flash-grab effect is indeed attenuated when the flash is not presented in the  
173 same eye as both the preceding and the subsequent motion sequences. This makes a strong case for the existence  
174 of neural extrapolation mechanisms in early, monocular stages in the visual hierarchy. The fact that the illusion was  
175 not entirely eliminated in these conditions indicates that extrapolation also occurs in later binocular areas.  
176 Altogether, the results therefore point towards extrapolation computations being carried out at multiple stages of  
177 the early visual pathway.

## 178 **Methods**

179 *Observers:* Twenty observers performed the experiment. All observers had normal or corrected-to-normal vision  
180 and gave informed consent prior to participating. The experiment was conducted in accordance with the Declaration  
181 of Helsinki and was approved by the ethics committee of the Melbourne School of Psychological Sciences. Data from  
182 three observers were excluded from the analysis because target detection was lower than 50%. The remaining  
183 observers successfully detected an average of 88% of targets across conditions.

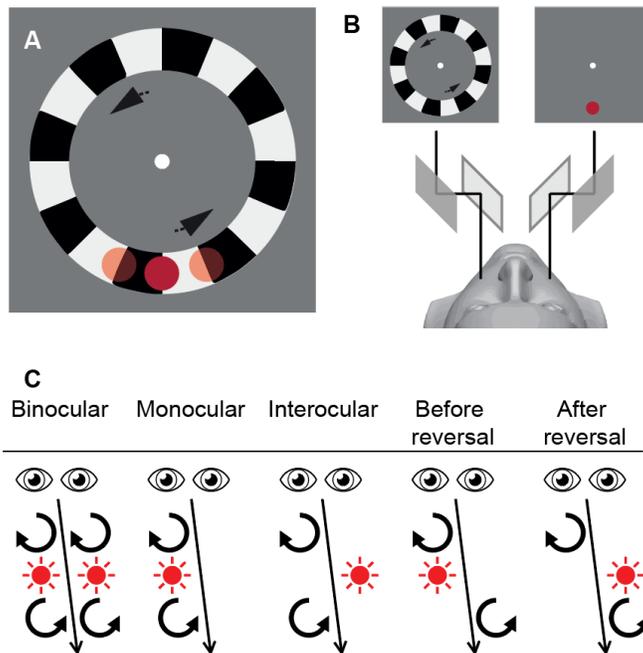
184 *Apparatus:* Stimuli were presented on an ASUS ROG Swift PG258Q monitor running at 100 Hz with a resolution of  
185 1920 x 1080 pixels, controlled by a Dell Precision computer. The experiment was presented using MATLAB (The  
186 MathWorks, Natick, MA) and Psychtoolbox 3.0.8 extensions (Brainard, 1997). A mirror-stereoscope set-up (including  
187 chin rest) was placed 50 cm away from the screen.

188 *Stimulus:* All stimuli were presented on a gray background. The stimuli consisted of two annuli (presented one to  
189 each eye), composed of 16 patches, which showed an alternating black and white pattern (see Figure 1A) and rotated  
190 at an angular velocity of 200 degrees per second. The annuli had inner and outer radii of 4.3 and 6.1 dva respectively.  
191 The two annuli were viewed through a mirror-stereoscope and fused into a single percept (Figure 1B). A black square  
192 was presented around both annuli throughout the experiment to assist in maintaining binocular fusion. The square  
193 was 9.3 dva wide, drawn with a linewidth of 0.8 dva. Fixation dots were presented in the center of both annuli  
194 (diameter : 0.6 dva). To give the observers some reference as to where they perceived the target, and to aid binocular  
195 fusion and avoid torsional eye movements, a white line was presented on the vertical meridian just below both  
196 annuli (width: 0.06 dva, height: 0.5 dva, 3.4 dva from fixation).

197 *Procedure:* On each trial, observers viewed a rotating annulus for either 1000, 1100, 1200, 1300, 1400 or 1500 ms  
198 (from now on referred to as the first motion sequence). During the very last frame of the motion sequence, a target  
199 (a red circle with a diameter of 0.6 dva) was presented at one of three possible target locations: 160, 180 or 200  
200 degrees polar angle offset from the top of the annulus, for a single frame (10 ms). Next, the direction of motion  
201 reversed and the annulus continued to rotate in the opposite direction for 400 ms, after which it gradually started  
202 to turn gray. The annulus was fully gray 100 ms later (these 500 ms are from now on referred to as the second motion  
203 sequence). This was done to ensure that participants were not distracted by the segments of the annulus when  
204 giving their response. At the end of each trial, observers used a mouse to report the position where they perceived  
205 the target. An image of the target was drawn at the cursor location for both eyes, and moved with the mouse cursor  
206 across the screen. When observers did not perceive the target, they were instructed to click at the location of the  
207 fixation dot.

208 *Experimental Design:* Although observers perceived the same series of events on every trial, we used a mirror  
209 stereoscope to manipulate the information presented to each eye across five different conditions (Figure 1C). 1) In  
210 the Binocular condition, all the information (first motion sequence, the target, and the second motion sequence)  
211 was presented to both eyes. 2) In the monocular condition, all information was presented to one eye only. 3) In the  
212 interocular condition, both the first and second motion sequence were presented to one eye, while the target was  
213 presented to the other eye. 4) In the Before Reversal condition, the first motion sequence and the target were  
214 presented to one eye, while the second motion sequence was presented to the other eye. 5) Lastly, in the After  
215 Reversal condition, the first motion sequence was presented to one eye, while the target and the second motion  
216 sequence were presented to other eye. The first motion sequence (and all consecutive events) occurred in the left  
217 and right eye with equal probability. The experiment consisted of 9 blocks, with 110 trials in each block. All conditions

218 were randomly interleaved within each block. On 10 % of the trials, no target was presented. These trials served as  
 219 catch-trials (1.8 % of which were wrongfully reported).



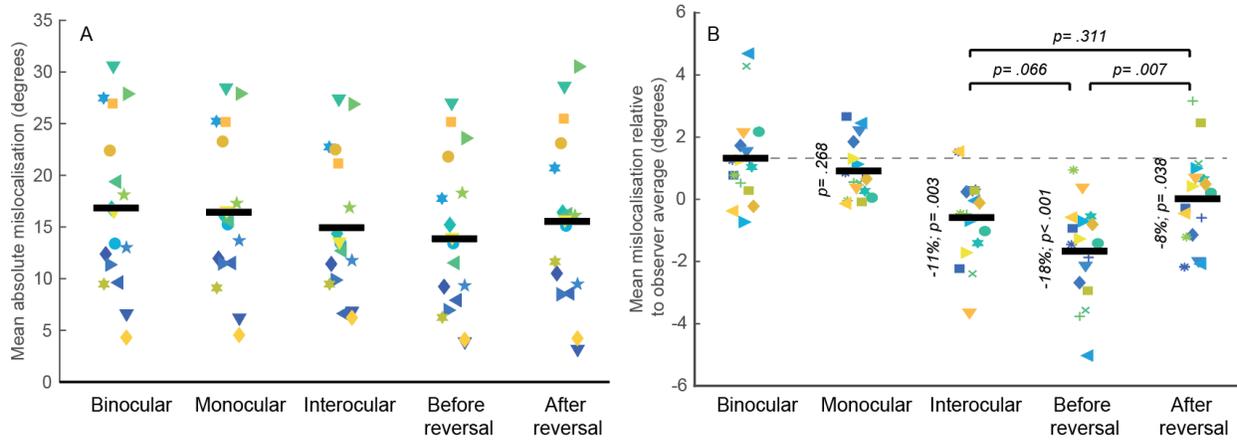
220  
 221 *Figure 2: Stimulus and Procedure. (A) Observers (N=17) viewed a flash-grab sequence consisting of a rotating*  
 222 *annulus that unexpectedly reversed its direction of motion. At the reversal, a red target disc was presented at one*  
 223 *of three possible target locations, and observers reported the perceived location of the target after the trial using a*  
 224 *mouse. (B) Using a mirror stereoscope, we manipulated the information presented to each eye. (C) Stimuli were*  
 225 *presented in five different conditions. Binocular condition: all information is presented to both eyes. Monocular*  
 226 *condition: all information is presented to one eye. Interocular condition: the moving annulus is presented to one*  
 227 *eye, while the target is presented to the other eye. Before reversal condition: the first motion sequence and the*  
 228 *target are presented to one eye, after which the annulus is presented to the other eye (rotating in the opposite*  
 229 *direction). After reversal condition: the first motion sequence is presented to one eye, after which both the annulus*  
 230 *(rotating in the opposite direction) and the target are presented in the other eye.*

231 **Results**

232 Observers viewed a flash-grab sequence (motion-flash-motion) in one of five conditions, in each case reporting the  
 233 perceived position of the flashed target (Figure 2). The strength of the illusion was calculated as the polar angle  
 234 between the reported position of the target and the target's real position, with errors in the direction of the second  
 235 motion sequence (i.e. post-reversal) taken as positive. Mean illusion strength in each condition is plotted in Figure  
 236 3, with and without baseline-correcting for variability in the mean strength of the illusion across observers. All  
 237 statistical analyses were carried out on the non-baselined data.

238

239



240

241 *Figure 3. Results. (A) Mean mislocalisation for individual observers in each of the five dichoptic presentation*  
 242 *conditions. Each combination of marker shape and color represents an individual observer. Solid black lines indicate*  
 243 *means across observers. (B) The same data after baseline-correction (subtracting the overall mean of each observer*  
 244 *from each of the conditions for that observer). This reveals that although observers vary widely in the magnitude of*  
 245 *the illusion, they all demonstrate a comparable pattern of illusion strength: in the three conditions in which the*  
 246 *complete stimulus sequence was not presented to the same eye (Interocular, Before Reversal, and After Reversal*  
 247 *conditions) the strength of the illusion was significantly attenuated. Statistical comparisons with paired-sample t-*  
 248 *tests are unaffected by the baseline correction and are illustrated only in Panel B for clarity. Vertical text indicates*  
 249 *comparisons against the Binocular condition.*

250 First, one-sample t-tests revealed that mislocalisation was evident in all conditions (all  $p < .001$ ). A repeated  
 251 measures analysis of variance subsequently revealed a highly significant effect of condition ( $F(4,64) = 10.3, p = 1.7 \times$   
 252  $10^{-6}$ , partial  $\eta^2 = .391$ ). To further interpret the results, we made planned comparisons between the conditions using  
 253 paired-samples t-tests. Each condition was compared to the baseline Binocular condition, and additional planned  
 254 pairwise comparisons were made between the three split conditions. The Monocular condition did not differ from  
 255 the Binocular condition ( $t(16) = -1.15, p = .26$ , Cohen's  $d = -.28$ ). Conversely, mislocalisation was significantly reduced  
 256 in the Interocular ( $t(16) = -3.43, p = .003$ , Cohen's  $d = -.83$ ), Before Reversal ( $t(16) = -4.67, p = .0003$ , Cohen's  $d = -$   
 257  $1.13$ ) and After Reversal ( $t(16) = -2.27, p = .037$ , Cohen's  $d = -.55$ ) conditions. Finally, mislocalisation did not differ  
 258 significantly between the Interocular and After Reversal conditions ( $t(16) = -1.0, p = .31$ , Cohen's  $d = -.25$ ), although  
 259 the Before Reversal condition was significantly reduced relative to the After Reversal condition ( $t(16) = -3.09, p =$   
 260  $.007$ , Cohen's  $d = -.75$ ), and there was a trend suggesting that the Before Reversal condition might also produce less  
 261 mislocalisation than the Interocular condition ( $t(16) = -1.98, p = .066$ , Cohen's  $d = -.75$ ).

262 The strength of the illusion was maximal when the entire stimulus sequence was presented to either one or both  
263 eyes. Importantly, the strength of the illusion was reduced when the motion and the flash were presented in  
264 separate eyes (*Interocular* condition) as well as when the first and second motion sequences were presented to  
265 separate eyes (*Before Reversal* and *After Reversal* conditions). Maximal reduction (the weakest illusion) was evident  
266 in the *Before Reversal* condition in which the flash was presented to the eye that received the first motion sequence.

267

## 268 **Discussion**

269 We have previously demonstrated that the flash-grab effect (Cavanagh & Anstis, 2013) involves neural extrapolation  
270 mechanisms that operate very early on in the visual pathway (Hogendoorn et al., 2015; Van Heusden et al., 2018).  
271 Here, we used dichoptic presentation to further narrow down when and where these mechanisms operate. Because  
272 visual input from the two eyes does not converge until primary visual cortex, separating the components of the flash-  
273 grab effect allowed us to discriminate whether motion extrapolation takes place at early, monocular stages (possibly  
274 retinal or subcortical), or at later binocular stages (V1 and/or beyond). The results reveal that dichoptic presentation  
275 attenuates, but does not eliminate the illusion, indicating that extrapolation mechanisms operate in both monocular  
276 and binocular visual processing.

277 This finding is important for two reasons. Firstly, it is consistent with animal work showing predictive mechanisms in  
278 monocular parts of the early visual pathway, including the retina (Berry et al., 1999; Hosoya et al., 2005; Schwartz et  
279 al., 2007) and LGN (Sillito et al., 1994). The present results reflect a similar involvement of predictive mechanisms in  
280 humans at both early, monocular stages (e.g. retina or LGN) and later, binocular stages (e.g. V1 and beyond; (Hubel  
281 & Wiesel, 1965, 1968)).

282 Secondly, both monocular and binocular neural populations contributed to the effect. This indicates that  
283 extrapolation occurs at multiple hierarchical (rather than single) processing stages. Extrapolation, or prediction, at  
284 multiple stages of sensory processing is the central property of hierarchical predictive coding, an influential  
285 theoretical and computational account of neural sensory processing (Huang & Rao, 2011; Rao & Ballard, 1999). In  
286 this model, successive layers of neurons “predict” their own input through feedback connections to earlier layers,  
287 feeding only the prediction error forward to higher layers (Rao & Ballard, 1999).

288 For example, in the visual system, a high-level neuron might represent a Gabor patch at a given position, spatial  
289 frequency, and orientation, and “predict” the local luminance of the lower-level neurons (with smaller, simpler  
290 receptive fields) that project to it. The lower-level neuron receiving the prediction then essentially compares the  
291 “prediction” to its input, and only feeds forward the deviations from that prediction – i.e. any properties of the  
292 stimulus not captured, or predicted, by the activity of the high-level neuron representing the Gabor. Conceptually,  
293 such a hierarchy would converge on patterns of connectivity and activation that minimize total prediction error in

294 the system. This would minimize metabolic requirements of sensory signaling, whilst optimizing information-  
295 theoretic properties of the network (a principle that has been dubbed the Free Energy Principle; Friston, 2005, 2010).

296 Importantly, the “predictions” in current models of predictive coding are predictive only in the hierarchical sense,  
297 but not in the temporal sense of predicting *future* activity (Bastos et al., 2012; Spratling, 2012, 2017; but see Friston  
298 (2005) for a discussion of predictive coding in time, and Garrido, Kilner, Stephan, & Friston (2009) for an empirical  
299 demonstration applied to expectation and mismatch negativity). However, neural transmission delays mean that for  
300 any time-variant input (such as visual motion), prediction errors at a given neural population are minimised by the  
301 higher area predicting that population’s *future*, rather than *current*, input. In a toy example, Area 1 sends visual  
302 information about the position of a moving object to Area 2, which in turn sends a “prediction” back to Area 1. That  
303 prediction is compared with input in Area 1 and any mismatch error is recursively minimised by adjusting the  
304 feedback signal to Area 1 to line up with its *input at the time the signal arrives there* (for details, see H. Hogendoorn  
305 & Burkitt, 2018). Minimising error therefore requires compensating for the delays incurred in both feed-forward and  
306 feed-back signaling. In the case of visual motion, compensating for these delays can be achieved by extrapolation:  
307 simply multiplying the instantaneous velocity of an object by the expected delay (feed-forward and feedback) yields  
308 a *spatiotemporal* prediction which is predictive in both the hierarchical and temporal sense. Indeed, several authors  
309 have proposed neural mechanisms for motion prediction, based on adaptation (Erlhagen, 2003) and Bayes-optimal  
310 motion-position estimation (Khoei et al., 2017; Kwon et al., 2015). However, these have not been related to the  
311 broader hierarchical predictive coding framework.

312 Because delays are incurred between each successive stage in the processing hierarchy, extrapolation must similarly  
313 occur at each stage if total prediction error is to be minimized. Extrapolation at each stage would require information  
314 about rate of change (i.e. velocity) at each stage. This is consistent with known properties of the early visual system:  
315 in lower vertebrates, velocity is extracted already in the retina (Amthor & Grzywacz, 1993), and although the  
316 proportion of direction-selective retinal ganglion cells in higher vertebrates is reduced (Bach & Hoffmann, 2000), in  
317 these animals direction-selective cells have been reported in the lateral geniculate nucleus (Niell, 2013). V1 itself of  
318 course also represents velocity (Hubel & Wiesel, 1962). Our finding that both monocular and binocular stages in the  
319 visual processing hierarchy carry out extrapolation is therefore anatomically plausible, and consistent with a version  
320 of hierarchical predictive coding that takes into account neural transmission delays.

321 Resulting as it does from an unexpected reversal of a moving background pattern, the flash-grab illusion is thought  
322 to occur due to a violation of expected motion (Cavanagh & Anstis, 2013; Hogendoorn et al., 2015). In the predictive  
323 coding framework, this amounts to a prediction error: a higher-level area extrapolates the position of the moving  
324 background, but by the time that predictive signal arrives at the lower-level area, the stimulus has reversed and the  
325 prediction (having been extrapolated in the initial direction, as indicated schematically in Figure 1) is very far from  
326 the new input. The resulting prediction error means that the represented position subsequently shifts rapidly over  
327 time, yielding a spike in velocity in the direction of the new motion (Figure 1C). As per the mechanism proposed by

328 Eagleman & Sejnowski (2007), and principally consistent with Kwon et al (2015), this velocity signal biases the  
329 perceived position of the (stationary) target that is briefly flashed superimposed on the background. In this sense,  
330 the flash-grab effect can be thought of as a direct reflection of prediction error. Two things about this interpretation  
331 remain to be elucidated. Firstly, even when no flash is presented, the reversal point of the sector edge on which the  
332 flash would otherwise be presented still undershoots the true physical point. Conversely, the neural representation  
333 of the edge, as measured in the retina (Schwartz et al., 2007) and proposed here to underlie the flash-grab effect,  
334 does not (Figure 1). The fate of these neural representations that do not reach awareness remains to be elucidated.  
335 In a similar vein, this sequence of neural representations generates a spike in the velocity signal (Figure 1C). We  
336 argue here that this causes a concurrent flash to be mislocalized, but perhaps this velocity spike also has other  
337 perceptual consequences. One interpretation could be that this velocity signal actually masks the final section of the  
338 position signal that represents the overshoot, comparable to the mechanisms proposed for saccadic suppression  
339 during eye movements (Ibbotson & Cloherty, 2009; Ibbotson, Crowder, Cloherty, Price, & Mustari, 2008), but this  
340 remains to be further explored.

341 The pattern of illusion strength in the three split conditions gives some further insight into the mechanisms that are  
342 likely to play a role. In the *Binocular* and *Monocular* conditions, in which the entire stimulus sequence is presented  
343 within a single eye, the violation of the background's motion direction can be detected at an early monocular stage.  
344 The prediction error therefore arises early, and is available to influence the neural representation of the target's  
345 position at both monocular (since the target is presented in the same eye) and binocular stages. This results in  
346 maximal prediction error (evident as maximal illusion strength, Figure 3). In the *Interocular* condition, the violation  
347 can be detected monocularly, but with the target being presented to the other eye, the target's representation can  
348 only be influenced when the monocular channels converge at a later binocular stage, thereby reducing the  
349 magnitude of the illusion. In the two other split conditions (*Before Reversal* and *After Reversal*), because the two  
350 motion sequences are presented to different eyes, the violation is only detected at a later binocular stage. Because  
351 receptive fields are generally larger further down the hierarchy, such that the discrepant extrapolated and actual  
352 positions of the target are more likely to fall within the same cell's receptive field as one looks further down the  
353 hierarchy, a given violation might be expected to yield a smaller error further down the hierarchy. Consistent with  
354 this, in our results, the illusion is attenuated in the two split conditions (paired-samples t-test of the two split  
355 conditions averaged together vs binocular;  $t(16)=4.3$ ,  $p < .001$ , Cohen's  $d = .96$ ). Finally, the illusion is more strongly  
356 reduced in the *Before Reversal* than *After Reversal* condition ( $t(16)=3.1$ ,  $p = .007$ , Cohen's  $d = .75$ ). This is a  
357 consequence of the so-called Frohlich effect (Kerzel, 2010), in which the onset position of a moving object is shifted  
358 in the direction of that object's subsequent trajectory. In the *After Reversal* condition, the second motion sequence  
359 does not violate a monocular motion prediction per se, but it still generates a monocular error signal due to the  
360 motion onset. In the *Before Reversal* condition, the same prediction error arises in the opposite eye to the target.  
361 Because this can only influence the target's position at the later binocular stage, this again leads to a smaller  
362 mislocalisation illusion.

363 The flash-grab effect has alternatively been explained in terms of trajectory shortening (Cavanagh & Anstis, 2013).  
364 It has been reported that the perceived trajectory of an object that reverses its direction is shortened (Sinico,  
365 Parovel, Casco, & Anstis, 2009), and that the perceived shift in the endpoint of the trajectory is linked to the position  
366 shift induced by the flash-grab effect (Cavanagh & Anstis, 2013). The trajectory-shortening explanation is formulated  
367 at a more abstract, computational level of description and therefore cannot offer any insight about how the effects  
368 should vary under monocular, binocular or dichoptic presentation. Cavanagh & Anstis (2013) argued that the flash-  
369 grab effect critically depends on attention, which might be taken as corresponding to a late, presumably binocular  
370 neural locus. As the available information at these stages would not be affected by dichoptic presentation, this  
371 interpretation in itself therefore does not provide a parsimonious explanation why the illusion would be attenuated  
372 in these conditions.

373 The proposition that motion and position signals interact already in monocular channels is consistent with a recent  
374 report studying motion-induced position shifts (the illusory displacement of the envelope of a Gabor patch when its  
375 carrier wave is moving; De Valois & De Valois, 1991). Hisakata, Hayashi, & Murakami (2016) observed that this  
376 illusory displacement is observed even when the carrier and the envelope are presented at widely divergent  
377 disparities, suggesting a disparity-insensitive monocular mechanism. They further showed that (as previously  
378 reported by Anstis, 1989) when illusory displacements are induced in the two eyes, the resulting illusory disparity  
379 yields an illusory depth percept, further supporting the involvement of motion-position interactions at monocular  
380 processing stages.

381 One limitation of the current study is that the three dichoptic conditions require binocular fusion, whereas the  
382 monocular and binocular presentations did not. Nevertheless, we do not believe this is a significant confound for a  
383 number of reasons. Firstly, a binocularly presented fixation point and large, high-contrast squares around the  
384 stimulus were presented to assist with binocular fusion, and during debrief none of the observers reported any  
385 difficulty with fusion. Furthermore, difficulty with binocular fusion mostly occurs when conflicting high-contrast  
386 stimuli are presented to each eye, which was never the case in our stimulus sequences. Instead, individual stimulus  
387 components were presented to one eye in the absence of any contrast energy in the other eye, a situation which  
388 does not negatively affect fusion (Alais & Blake, 2004). Finally, it is not clear how problems with binocular fusion  
389 would explain the observed differences between the different dichoptic conditions. Nevertheless, we cannot entirely  
390 rule out the possibility that differences in binocular fusion may have played a role in our experimental conditions.

391 In sum, we have used a dichoptic version of the flash-grab effect to study the monocular and binocular contributions  
392 to motion extrapolation in human visual motion perception. The results reveal that extrapolation mechanisms  
393 operate at both monocular and binocular processing stages – a finding that is consistent with an extension of the  
394 hierarchical predictive coding framework that accounts for neural transmission delays. The results further suggest  
395 that prediction errors in this framework can manifest directly as perceptual illusions.

396 **Acknowledgments**

397 EvH and HH were supported by the Australian Government through the Australian Research Council's Discovery  
398 Projects funding scheme (project DP180102268). MG acknowledges a University of Queensland Fellowship  
399 (2016000071).

400 **References**

- 401 Alais, D., & Blake, R. (2004). *Binocular Rivalry*. Cambridge, MA: MIT Press.
- 402 Amthor, F. R., & Grzywacz, N. M. (1993). Directional selectivity in vertebrate retinal ganglion cells. *Reviews of*  
403 *Oculomotor Research*, 5, 79–100. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8420563>
- 404 Anstis, S. M. (1989). Kinetic edges become displaced, segregated, and invisible. In D. Lam (Ed.), *Neural mechanisms*  
405 *of visual perception* (Vol. 2, pp. 247–260). Gulf Publishing Houston, TX.
- 406 Bach, M., & Hoffmann, M. B. (2000). Visual motion detection in man is governed by non-retinal mechanisms. *Vision*  
407 *Research*, 40(18), 2379–2385. [https://doi.org/10.1016/S0042-6989\(00\)00106-1](https://doi.org/10.1016/S0042-6989(00)00106-1)
- 408 Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012, November 21). Canonical  
409 Microcircuits for Predictive Coding. *Neuron*. NIH Public Access.  
410 <https://doi.org/10.1016/j.neuron.2012.10.038>
- 411 Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina.  
412 *Nature*, 398(6725), 334–338. <https://doi.org/10.1038/18678>
- 413 Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. Retrieved from  
414 <http://www.ncbi.nlm.nih.gov/pubmed/9176952>
- 415 Brenner, E., Smeets, J. B., & de Lussanet, M. H. (1998). Hitting moving targets. Continuous control of the  
416 acceleration of the hand on the basis of the target's velocity. *Experimental Brain Research*, 122(4), 467–474.  
417 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9827866>
- 418 Cai, R. H., & Schlag, J. (2001). Asynchronous feature binding and the flash-lag illusion. *Investigative Ophthalmology*  
419 *and Visual Science*, (42), S711.
- 420 Cavanagh, P., & Anstis, S. M. (2013). The flash grab effect. *Vision Research*, 91, 8–20.  
421 <https://doi.org/10.1016/j.visres.2013.07.007>
- 422 De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research*, 31(9),  
423 1619–1626. [https://doi.org/10.1016/0042-6989\(91\)90138-U](https://doi.org/10.1016/0042-6989(91)90138-U)
- 424 Eagleman, D. M. (2008, April 14). Prediction and postdiction: Two frameworks with the goal of delay  
425 compensation. *Behavioral and Brain Sciences*. Cambridge University Press.  
426 <https://doi.org/10.1017/S0140525X08003889>
- 427 Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science (New*  
428 *York, N.Y.)*, 287(5460), 2036–2038. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10720334>
- 429 Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: A unified explanation for the  
430 flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of Vision*, 7(4), 3–3.  
431 <https://doi.org/10.1167/7.4.3>
- 432 Erlhagen, W. (2003). Internal models for visual perception. *Biological Cybernetics*. [https://doi.org/10.1007/s00422-](https://doi.org/10.1007/s00422-002-0387-1)  
433 [002-0387-1](https://doi.org/10.1007/s00422-002-0387-1)

- 434 Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological*  
435 *Sciences*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- 436 Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.  
437 <https://doi.org/10.1038/nrn2787>
- 438 Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying  
439 mechanisms. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical*  
440 *Neurophysiology*, 120(3), 453–463. <https://doi.org/10.1016/j.clinph.2008.11.029>
- 441 Hisakata, R., Hayashi, D., & Murakami, I. (2016). Motion-induced position shift in stereoscopic and dichoptic  
442 viewing. *Journal of Vision*, 16(13), 3. <https://doi.org/10.1167/16.13.3>
- 443 Hogendoorn, H., & Burkitt, A. N. (2018a). Predictive coding of visual object position ahead of moving objects  
444 revealed by time-resolved EEG decoding. *NeuroImage*, 171.  
445 <https://doi.org/10.1016/j.neuroimage.2017.12.063>
- 446 Hogendoorn, H., & Burkitt, A. N. (2018b). Predictive coding with neural transmission delays: a real-time temporal  
447 alignment hypothesis. *BioRxiv*. <https://doi.org/10.1101/453183>
- 448 Hogendoorn, H., Verstraten, F. A. J., & Cavanagh, P. (2015). Strikingly rapid neural basis of motion-induced position  
449 shifts revealed by high temporal-resolution EEG pattern classification. *Vision Research*, 113(PA).  
450 <https://doi.org/10.1016/j.visres.2015.05.005>
- 451 Hosoya, T., Baccus, S. A., & Meister, M. (2005). Dynamic predictive coding by the retina. *Nature*, 436(7047), 71–77.  
452 <https://doi.org/10.1038/nature03689>
- 453 Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(5), 580–  
454 593. <https://doi.org/10.1002/wcs.142>
- 455 Hubel, D. H., & Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's  
456 visual cortex. *The Journal of Physiology*, 160(1), 106–154. Retrieved from  
457 <http://www.ncbi.nlm.nih.gov/pubmed/14449617>
- 458 Hubel, D. H., & Wiesel, T. (1965). BINOCULAR INTERACTION IN STRIATE CORTEX OF KITTENS REARED WITH  
459 ARTIFICIAL SQUINT. *Journal of Neurophysiology*, 28(6), 1041–1059.  
460 <https://doi.org/10.1152/jn.1965.28.6.1041>
- 461 Hubel, D. H., & Wiesel, T. (1968). Receptive Fields and Functional Architecture of monkey striate cortex. *Journal of*  
462 *Physiology*, 195(1), 215–243. [https://doi.org/papers://47831562-1F78-4B52-B52E-  
463 78BF7F97A700/Paper/p352](https://doi.org/papers://47831562-1F78-4B52-B52E-78BF7F97A700/Paper/p352)
- 464 Ibbotson, M. R., & Cloherty, S. L. (2009). Visual Perception: Saccadic Omission - Suppression or Temporal Masking?  
465 *Current Biology*, 19(12), R493–R496. <https://doi.org/10.1016/j.cub.2009.05.010>
- 466 Ibbotson, M. R., Crowder, N. A., Cloherty, S. L., Price, N. S. C., & Mustari, M. J. (2008). Saccadic Modulation of  
467 Neural Responses: Possible Roles in Saccadic Suppression, Enhancement, and Time Compression. *Journal of*  
468 *Neuroscience*, 28(43), 10952–10960. <https://doi.org/10.1523/JNEUROSCI.3950-08.2008>
- 469 Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for  
470 flashes in population responses of cat primary visual cortex. *Journal of Physiology*, 556(3), 971–982.  
471 <https://doi.org/10.1113/jphysiol.2003.058941>
- 472 Kerzel, D. (2010). The Fröhlich effect: past and present. In R. Nijhawan & B. Khurana (Eds.), *Space and Time in*  
473 *Perception and Action* (pp. 321–337). Cambridge: Cambridge University Press.  
474 <https://doi.org/10.1017/CBO9780511750540.019>
- 475 Kerzel, D., & Gegenfurtner, K. R. (2003). Neuronal Processing Delays Are Compensated in the Sensorimotor Branch

476 of the Visual System. *Current Biology*, 13(22), 1975–1978. <https://doi.org/10.1016/j.cub.2003.10.054>

477 Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The Flash-Lag Effect as a Motion-Based Predictive Shift. *PLOS*  
478 *Computational Biology*, 13(1), e1005068. <https://doi.org/10.1371/journal.pcbi.1005068>

479 Krekelberg, B. (2000). The Position of Moving Objects. *Science*, 289(5482), 1107a–1107.  
480 <https://doi.org/10.1126/science.289.5482.1107a>

481 Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends in Neurosciences*,  
482 24(6), 335–339. [https://doi.org/10.1016/S0166-2236\(00\)01795-1](https://doi.org/10.1016/S0166-2236(00)01795-1)

483 Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings*  
484 *of the National Academy of Sciences of the United States of America*, 112(26), 8142–8147.  
485 <https://doi.org/10.1073/pnas.1500361112>

486 Lotter, W., Kreiman, G., & Cox, D. (2018). A neural network trained to predict future video frames mimics critical  
487 properties of biological neuronal responses and perception, 1–18. Retrieved from  
488 <http://arxiv.org/abs/1805.10734>

489 Maus, G. W., & Nijhawan, R. (2008). Motion Extrapolation Into the Blind Spot. *Psychological Science*, 19(11), 1087–  
490 1091. <https://doi.org/10.1111/j.1467-9280.2008.02205.x>

491 Niell, C. M. (2013). Vision: More Than Expected in the Early Visual System. *Current Biology*, 23(16), R681–R684.  
492 <https://doi.org/10.1016/J.CUB.2013.07.049>

493 Nijhawan, R. (1994, July 28). Motion extrapolation in catching. *Nature*. Nature Publishing Group.  
494 <https://doi.org/10.1038/370256b0>

495 Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays.  
496 *Behavioral and Brain Sciences*, 31(2), 179–239. <https://doi.org/10.1017/S0140525X08003804>

497 Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, 8(5), 379–  
498 391. <https://doi.org/10.1038/nrn2131>

499 Patel, S. S., Ogmen, H., Bedell, H. E., & Sampath, V. (2000). Flash-lag effect: differential latency, not postdiction.  
500 *Science (New York, N.Y.)*, 290(5494), 1051. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11184992>

501 Ramachandran, V. S., & Anstis, S. M. (1990). Illusory Displacement of Equiluminous Kinetic Edges. *Perception*,  
502 19(5), 611–616. <https://doi.org/10.1068/p190611>

503 Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some  
504 extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87. <https://doi.org/10.1038/4580>

505 Schwartz, G., Taylor, S., Fisher, C., Harris, R., Berry, M. J., & Il. (2007). Synchronized firing among retinal ganglion  
506 cells signals motion reversal. *Neuron*, 55(6), 958–969. <https://doi.org/10.1016/j.neuron.2007.07.042>

507 Sillito, A. M., Jones, H. E., Gerstein, G. L., & West, D. C. (1994). Feature-linked synchronization of thalamic relay cell  
508 firing induced by feedback from the visual cortex. *Nature*, 369(6480), 479–482.  
509 <https://doi.org/10.1038/369479a0>

510 Sinico, M., Parovel, G., Casco, C., & Anstis, S. (2009). Perceived shrinkage of motion paths. *Journal of Experimental*  
511 *Psychology: Human Perception and Performance*, 35(4), 948–957. <https://doi.org/10.1037/a0014257>

512 Spratling, M. W. (2012). Predictive coding accounts for V1 response properties recorded using reverse correlation.  
513 *Biological Cybernetics*, 106(1), 37–49. <https://doi.org/10.1007/s00422-012-0477-7>

514 Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain and Cognition*, 112, 92–97.  
515 <https://doi.org/10.1016/j.bandc.2015.11.003>

- 516 Sundberg, K. A., Fallah, M., & Reynolds, J. H. (2006). A Motion-Dependent Distortion of Retinotopy in Area V4.  
517 *Neuron*, 49(3), 447–457. <https://doi.org/10.1016/j.neuron.2005.12.023>
- 518 Van Heusden, E., Rolfs, M., Cavanagh, P., & Hogendoorn, H. (2018). Motion extrapolation for eye movements  
519 predicts perceived motion-induced position shifts. *Journal of Neuroscience*.  
520 <https://doi.org/10.1523/JNEUROSCI.0736-18.2018>
- 521 Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: shifting the perceived position of remote  
522 stationary objects. *Nature Neuroscience*, 3(9), 954–959. <https://doi.org/10.1038/78878>
- 523 Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience*, 1(8), 656–  
524 657. <https://doi.org/10.1038/3659>
- 525