Prediction in goal-directed action

Katja Fiehler
Department of Psychology, Justus Liebig University, Giessen, Germany
Center for Mind, Brain, and Behavior (CMBB), Universities of Marburg and Giessen, Germany

Eli Brenner
Department of Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands

Miriam Spering
Department of Ophthalmology & Visual Sciences, University of British Columbia, Vancouver, Canada

Prediction allows humans and other animals to prepare for future interactions with their environment. This is important in our dynamically changing world that requires fast and accurate reactions to external events. Knowing when and where an event is likely to occur allows us to plan eye, hand, and body movements that are suitable for the circumstances. Predicting the sensory consequences of such movements helps to differentiate between self-produced and externally generated movements. In this review, we provide a selective overview of experimental studies on predictive mechanisms in human vision for action. We present classic paradigms and novel approaches investigating mechanisms that underlie the prediction of events guiding eye and hand movements.

Introduction

The ability to predict the sensory outcome of an action or event is considered critical for survival across many species. Even amphibians rely on trajectory prediction to capture prey and use sophisticated sensorimotor mechanisms to achieve this (Borghuis & Leonardo, 2015). In humans, predictions based on cognitive expectations or beliefs are of fundamental importance for many areas of visual perception, cognition, and the control of motor behavior. Predictions are needed to overcome sensory processing delays and to enable well-timed and accurate reactions to objects, events, or other individuals’ actions.

Predictive mechanisms work at a range of time scales, relying on a wide variety of sources of information, and guide many different kinds of judgments and actions. Predictions may be based on one’s knowledge of the physical environment, such as relying on one’s experience with gravity when anticipating the movement of a flying ball (McIntyre, Zago, Berthoz, & Lacquaniti, 2001; Zago, McIntyre, Senot, & Lacquaniti, 2008, 2009). They may also rely on memory of the particular visual scene, such as the likelihood of an object of interest being at a particular location (Draschkow, Wolfe, & Vô, 2014; Henderson, 2017; Vô & Wolfe, 2015), or on the statistics of events or items typically encountered in such scenes (Diaz, Cooper, Rothkopf, & Hayhoe, 2013; Hayhoe, McKinney, Chajka, & Pelz, 2012). Predictions can be related to temporal factors (e.g., when will an object be within reach; Rodriguez-Herreros, de Grave, Lopez-Moliner, Brenner, & Smeets, 2013) as well as spatial representations (e.g., completion of occluded parts of an object; Ekroll, Sayim, & Wagemans, 2017; van Lier, 1999).

Predictions are critical for planning eye, hand, and body movements and have been examined extensively across species and paradigms, both in highly constrained laboratory conditions and in real-world settings. Besides movement planning, predictions of the sensory consequences of actions allow us to dissociate self-produced from externally generated movements, influencing our sense of agency (Blakemore, Wolpert, & Frith, 1998). Acquiring expertise in predicting events and actions is important for proficiency in sports (Abernethy, Zawi, & Jackson, 2008; Cañal-Bruland, van der Kamp, & van Kesteren, 2010; Mann, Abernethy, & Farrow, 2010; Savelsbergh, Williams, Kamp, & Ward, 2002; Tomeo, Cesari, Aglioti, & Urgesi, 2013), illustrating the ability to improve predictions with training.

In this review, we provide a selective overview of paradigms and experimental studies on predictive mechanisms in human vision for action. We also

discuss some similarities with other senses. Studies conducted in different settings are included, ranging from controlled laboratory experiments to descriptive studies conducted in real-world situations. We do not particularly focus on prediction in real-world or natural task settings as other reviews on the topic have already discussed such studies (Hayhoe, 2017).

We begin by discussing a special form of short-term prediction: the efference copy. This mechanism appears to be important for predicting external object motion as well as the sensory consequences of eye and hand movements. Brain signals for motor control obviously provide quite reliable information about the upcoming movements, especially when dealing with movements of the eyes, because the eyes do not have to cope with issues such as changing inertial loads. The two following sections discuss the role of prediction in planning eye and hand movements, followed by a section on prediction in modulating sensory information. We conclude by presenting a selection of possible future research directions.

### The efference copy as a form of prediction

Imagine watching a train pass by. If you do not follow the train with your eyes, its image will sweep across the retina while the image of the context surrounding the train will not. If you do follow the train with your eyes, its image will be (more or less) stable on the retina, while the context’s image will produce a motion sweep. How does the brain figure out when to attribute retinal motion to object motion and when to attribute it to movements of the eyes? Von Holst and Mittelstaedt (1950) proposed that when a motor command is sent to the muscles that move the eyes, a copy of the efferent signal is simultaneously sent to visual areas of the brain. This information is presumably used to remove the predicted retinal slip signal arising from the eye movement from the actual retinal slip signal. Experimental support for the existence of this “efference copy” mechanism comes from studies with induced eye paralysis in healthy individuals using mechanical (Mach, 1959/1897; Stark & Bridgeman, 1983) or pharmacological interventions (Brindley, Goodwin, Kulikowski, & Leighton, 1976). When the eye muscles are paralyzed, individuals perceive motion of the external visual world when they attempt to move their eyes (despite the lack of proprioceptive input). An attempted saccade results in perceived displacement of the visual world in the direction of the intended saccade (Stevens et al., 1976). An efference copy (or corollary discharge) mechanism may therefore contribute to stable visual perception across saccades (Cavanaugh, Berman, Joiner, & Wurtz, 2016), or more generally to space constancy (Sperry, 1950; von Helmholtz, 1962; reviewed in Bridgeman, 2016), or more generally to space constancy (Sperry, 1950; von Helmholtz, 1962; reviewed in Bridgeman, 2010). The alternative would be to rely on information that is not influenced by eye movements (e.g., relative positions or motion; Brenner & van den Berg, 1994).

An important feature of using an efference copy mechanism to predict the sensory consequences of eye movements is that such a mechanism can provide information about the eye movement before it takes place (due to delays between motor commands leaving the brain and the eye muscles responding to such commands). This means that it could be used to anticipate the consequences of abrupt movements (such as saccades) before they start. Using an actual copy of the signals to the eye muscles to obtain information about the eyes’ future orientation can only predict eye orientation after exactly the time that it takes for the motor commands to move the eyes. However, the idea of using prior information about one’s intended actions to interpret future input might apply much more generally (Pickering & Clark, 2014). One might use signals that regulate the velocity of smooth pursuit eye movements to judge where gaze will be directed at different times. One might also use signals related to hand and arm movements to anticipate sensory action effects that can be used to correct the current movement as well as to update future predictions. When planning a complicated or continuous action it might be beneficial to use the signals that drive a component of the action to help plan the next component, rather than planning the whole action in advance or always waiting for the first component to end so that one can rely on feedback. In the remainder of this review we will concentrate on predictions that go beyond predicting the visual input that one will obtain after the duration of the oculomotor delay. We will start with predictions about external objects, and then return to predicting the consequences of our own actions for interpreting sensory input.

### Prediction in eye movements

Predictions are of fundamental importance for the control of a range of goal-directed oculomotor behaviors. The ability to make adequate predictive eye and head movements during natural tasks like ball sports has been associated with superior sports performance (Land & McLeod, 2000; Mann, Sratford, & Abernethy, 2013). Many aspects of smooth pursuit eye movements—the eyes’ key response to moving visual objects—rely on predictive signals. Smooth pursuit is a continuous tracking behavior, making it suitable for investigating the temporal
properties of predictive responses. This section provides an overview of the literature on prediction in eye movements with a focus on pursuit, discussing both the smooth component and corrective saccades. A note on terminology: The eye movement literature occasionally distinguishes between anticipatory pursuit, made in expectation of an upcoming target’s trajectory prior to initial target onset, and predictive pursuit, made in expectation of a reappearing target that observers had been exposed to previously. Both types of response are based on prediction of a future event and will here be discussed jointly as anticipatory pursuit.

Anticipatory pursuit is driven by cognitive expectations about a target’s motion

Human and nonhuman primates can reliably initiate anticipatory pursuit as early as 200 milliseconds before the initial onset of object motion (Badler & Heinen, 2006; Barnes, 2008; Barnes & Schmid, 2002; de Hemptinne, Lefèvre, & Missal, 2006; Dodge, Travis, & Fox, 1930; Freyberg & Ilg, 2008; Kowler, 1989, 2011; Kowler, Aitkin, Ross, Santos, & Zhao, 2014; Westheimer, 1954). Anticipatory pursuit usually occurs when the onset and motion direction of a moving target are predictable, either due to the stimulus configuration (e.g., when the peripheral position of a fixation spot on the screen is followed by centripetal target motion in a predefined direction) or because it is indicated by a cue (e.g., an arrow pointing in the direction of the upcoming target). A series of seminal studies by Kowler (1989) introduced a paradigm designed to investigate the origin of anticipatory pursuit. Pursuit prior to target motion onset could be guided by oculomotor learning (or habit formation) following the repetition of several trials with identical motion direction. Alternatively, it could be triggered by cognitive expectation, generated by a symbolic cue. In Kowler’s (1989) original paradigm, observers viewed a ball-like object falling down a tube that branched into two arms (Figure 1). In some trials, observers did not know which way the ball would travel ahead of time. In these trials, anticipatory pursuit was generated only if the ball travelled down the same branch for several trials in a row (Figure 1a). In other trials, observers were cued as to which direction the ball would travel, either by auditory instruction or by presenting a visual barrier (Figure 1b). A symbolic cue consistently produced stronger anticipatory pursuit than pure stimulus repetition, indicating the important role of the certainty of the expectation in guiding predictive eye movements.

These early studies employed cues that indicated upcoming motion direction with high certainty. However, cues in our natural environment are usually not so conclusive. To mimic naturalistic requirements and reductions in certainty, studies have employed cues that carry probabilistic information (e.g., probability of motion direction to the left or right or of the ball traveling down the left or right arm, see Figure 1c). Anticipatory pursuit is finely tuned to probabilistic information. The velocity of the anticipatory pursuit response increases in proportion to target motion probability (Santos & Kowler, 2017). The exact relation between anticipatory pursuit strength and probability level depends on the type of cue. Anticipatory pursuit in response to symbolic probabilistic cues (Figure 1c) underweighted low probabilities and overweighted high probabilities. The reverse finding was observed when anticipatory pursuit was based on stimulus repetition, i.e., on memory (as shown in Figure 1a; Santos & Kowler, 2017).

When timing, direction, or velocity is unpredictable, anticipatory pursuit is not necessarily abolished. However, an increase in temporal uncertainty leads to an increase in the width of the pursuit latency
distribution (de Hemptinne, Nozaradan, Duvivier, Lefèvre, & Missal, 2007) and to temporal averaging (Heinen, Badler, & Ting, 2005). When the predictive signal is noisy, an estimate of target motion probability is likely derived from short-term memory or past experience (Barnes & Collins, 2008). It is known that the pursuit system rapidly and continuously updates such memory-based priors—even after one-trial exposure (Maryott, Noyce, & Sekuler, 2011)—and integrates them with sensorimotor evidence (Bogadhi, Montagnini, & Masson, 2013; Darlington, Tokiyama, & Lisberger, 2017; Deravet, Blohm, Orban de Xivry, & Lefèvre, 2018; Orban de Xivry, Coppe, Blohm, & Lefèvre, 2013). Moreover, anticipatory pursuit scales with reinforcement contingencies. Whereas an expected monetary reward per se only had a small effect on anticipatory pursuit, reward contingencies interacted with a probability effect; i.e., reward modulated anticipatory pursuit strongly in the presence of a strong probability-driven direction bias (Damasse, Perrinet, Madelain, & Montagnini, 2018). These findings indicate that anticipatory pursuit can be considered as an operant behavior.

To summarize, anticipatory pursuit can be driven by cognitive expectation of target motion (Kowler, 1989; Kowler, Aitkin, Ross, Santos, & Zhao, 2014). It is sensitively tuned to cue probability and can effectively be used as a model of sensorimotor integration.

**Predictive pursuit is finely tuned to target properties**

The ability to predict the future position and trajectory of moving objects reflects an understanding of the physical properties of the object and the surrounding world (Freyd, 1987; Diaz et al., 2013; Nusseck, Lagarde, Bardy, Fleming, & Bültthoff, 2007; Zago, Bosco, Maffei, Iosa, Ivanenko, & Lacquaniti, 2004). Predictions can be based on visual cues (Gray & Regan, 1998; Regan & Beverley, 1979) that are weighted and combined effectively to optimally respond to the properties of the object and the demands of the task (Rushton & Wann, 1999). When visual judgments are poor, as is for instance the case for judgments of acceleration (Brouwer, Brenner, & Smeets, 2002; Werkhoven, Snippe, & Toet, 1992), relying on predictions can lead to errors (Brenner et al., 2016; Tresilian, 1995). Sometimes, such judgment errors can be avoided by combining online visual information with prior knowledge of the target’s dynamics (Körding & Wolpert, 2004). For example, visual signals about the position and velocity of a target can be combined with an internal estimate of gravity (McIntyre et al., 2001). Alternatively, predictions might simply be adjusted to feedback about errors (Brenner et al., 2016).

How predictive signals guide smooth pursuit eye movements has been investigated by applying variations of a classic paradigm in which the target is blanked; i.e., it transiently disappears from view (Barnes, 2008; Becker & Fuchs, 1985; Mitrani & Dimitrov, 1978; Orban de Xivry, Bennett, Lefèvre, & Barnes, 2006; Orban de Xivry, Missal, & Lefèvre, 2008; Figure 2a). Smooth pursuit eye movements can be maintained for several seconds during this blanking period, albeit at a decreased velocity level (Figure 2b). According to a simple reflexive model of pursuit, such velocity maintenance in the absence of a visual signal may be driven by basic predictive signals based on eye-velocity memory. An efference copy signal could serve as an input signal to perpetuate ongoing smooth pursuit when no visual input is available. In the absence of visual feedback from image velocity and acceleration, the efference copy is not updated, resulting in a gradual decrease of eye velocity over time. Interestingly, if the duration of the blanking period is predictable (i.e., the target is repeatedly blanked for the same amount of time from trial to trial), observers’ pursuit commonly accelerates towards the end of the blanking period in anticipation of the reappearing target (Bennett & Barnes 2003, 2004; Churchland, Chou, & Lisberger, 2003; Orban de Xivry et al., 2006; Figure 2b). This predictive acceleration can build up within three trials (Bennett, Orban de Xivry, Lefèvre, & Barnes, 2010). Importantly, it scales with visual properties of the reappearing target such as velocity (Bennett & Barnes, 2004; Orban de Xivry et al., 2006).
and acceleration (Bennett & Barnes, 2003; Bennett, Orban de Xivry, Barnes, & Lefèvre, 2007; Bennett et al., 2010). These findings show that smooth pursuit is at least partly under cognitive control, especially if there is a strong expectation about future events such as the timing of target reappearance (Bennett & Barnes, 2003; Churchland et al., 2003).

Signals that are neither related to vision nor to eye movements can help the eyes to continue tracking targets that are temporarily occluded, and it is interesting to discuss how these signals might be integrated with visual and eye movement information to drive pursuit. For example, smooth pursuit is generally enhanced if gaze is directed at a target that is actively moved by the observer’s hand (Gauthier & Hofferer, 1976; Gauthier, Vercher, Mussa Ivaldi, & Marchetti, 1988). This finding indicates that the oculomotor system might have access to a limb efference copy signal (Ariff, Donchin, Nanayakkara, & Shadmehr, 2002) and might use an internal representation of the dynamics of the hand to control the eye (Landelle, Montagnini, Madelain, & Danion, 2016). Consistent with these assumptions, smooth pursuit can be enhanced by haptic feedback during target occlusion (Danion, Mathew, & Flanagan, 2017). Similarly, pursuit maintenance in the absence of visual signals can be enhanced by a reinforcing auditory stimulus (Madelain & Krauzlis, 2003). In this study, observers viewed a visual target that was temporarily occluded. They heard an auditory signal that provided feedback about the real-time tracking performance: A continuous tone was played as long as pursuit was free of saccades; when a saccade was detected, the tone was interrupted for 100 ms. If the tone was interrupted for less than 20% of the trial, the trial was considered successful. The success rate contributed to the monetary reward paid out at the end of the session. Such real-time sonification of smooth pursuit may successfully improve performance in several tasks (Boyer, Portron, Bevilacqua, & Lorenceau, 2017), but requires extensive practice (several hours).

Synergies between pursuit and saccades in predicting object motion

Many of the studies summarized above focus on the smooth pursuit component of eye movements. However, saccades are also important for understanding prediction in goal-directed action. Catch-up saccades occur frequently during pursuit. They are thought to reduce retinal position errors that would otherwise accumulate over time whenever pursuit is not perfect (de Brouwer, Yuksel, Blohm, Missal, & Lefèvre, 2002; Orban de Xivry & Lefèvre, 2007). However, such saccades are not simply triggered whenever there is a retinal error, as is evident from stimulus configurations that are specially designed to prevent catch-up saccades. When a pursuit target steps in one direction (e.g., to the left of fixation) and then moves in the opposite direction (to the right), observers appear to anticipate that the target motion will reduce the position error, and refrain from making an initial saccade in the direction of the step to align gaze with the target. Instead, observers initiate a smooth eye movement in the direction of the target’s motion, especially if the target passes the initial fixation point after the time it usually takes to initiate a pursuit response (Rashbass, 1961). Moreover, although the primary role of catch-up saccades appears to be to realign gaze with the target of pursuit, such saccades do influence judgments of the target’s speed and future position (Goettker, Brenner, Gegenfurtner, & de la Malla, 2019).

Besides making catch-up saccades, people also make anticipatory saccades during ongoing pursuit, such as when the batter in cricket makes a predictive saccade to a future bounce location (Land & McLeod, 2000) or when a participant in a simple interception task makes a predictive saccade to where the target is to be hit (de la Malla, Smeets, & Brenner, 2017). The timing of these predictive saccades is tightly linked to the execution of interceptive hand movements and is correlated with the outcome of the action (Fooken & Spering, 2019). The characteristics of catch-up saccades are a good probe of where the observer thinks the target will be in the near future, and which action goal is selected (e.g., to try or not to try to hit the ball).

The endpoints of saccades towards where occluded or blanked targets are anticipated to reappear also provide a good probe of predictive mechanisms. When monkeys were trained to make saccades to where a moving target that disappeared would reappear if it were deflected by a visible structure, they corrected initial targeting errors even when the target was not visible (Ferrera & Barborica, 2010). Thus, the correction had to be driven by an internal signal. Neurophysiological evidence pointed to the frontal eye fields as the origin of the error signal that led to this behavior, suggesting that this area plays a role in detecting and correcting movement errors based on internal feedback signals (Ferrera & Barborica, 2010; Teichert, Yu, & Ferrera, 2014).

Predictive components in catch-up saccade programming and in error correction are also reflected in saccade curvature. When asked to track a target that jumped and moved unpredictably, human observers often made saccades that were highly curved (Schreib- er, Missal, & Lefèvre, 2006). When the target’s position and motion changed twice in rapid succession, the saccade that was executed to deal with the change was initially oriented towards the anticipated target position. It then gradually curved towards the final
position. Evaluating such curvature can provide a continuous measure of where the brain predicts the target to be located in the future (Orban de Xivry & Lefèvre, 2007; Schreiber, Missal, & Lefèvre, 2006).

### How predictive signals drive eye movements when interacting with dynamic objects

The ability to anticipate future events plays a critical role in many daily contexts. It is particularly evident and important in certain sports. In cricket, for example, skilled batsmen fixate the ball as it leaves the bowler’s hand and then quickly make a saccade to near where the ball will bounce, waiting for its image to come back onto the fovea (Land & MacLeod, 2000). The brain must have combined the visual information that was available soon after the bowler released the ball with a learned model of the ball’s motion to predict the location of the bounce and how the ball will move after the bounce. The learned model considers issues such as gravity (Zago et al., 2009) and the elasticity of the ball (Diaz et al., 2013). Saccades sometimes land above a future bounce point (Diaz et al., 2013; Hayhoe et al., 2012), possibly to facilitate tracking the ball after the bounce by directing gaze to where the ball will pass shortly after it bounces. Expertise in sports may partly consist of being better at anticipating. Whereas novice baseball batters are usually unable to track the fastest balls with their gaze, elite batters rely on a combination of eye and head rotations to track a fastball (Higuchi, Nagami, Nakata, & Kanosue, 2018). Elite batters can even track the ball up to the moment of bat-ball contact in cricket (Mann et al., 2013).

Interestingly, when participants were asked to observe two players throwing a ball back and forth instead of actively throwing and catching it themselves, predictive saccades were less accurate, and fewer periods of smooth pursuit were observed between bounce and catch (Hayhoe, Minnie, Sullivan, & Gorgos, 2005). Similarly, laboratory studies have shown that actively engaging in tasks such as manual tracking, catching, or pointing can enhance simultaneous pursuit eye movements (Chen, Valsecchi, & Gegenfurtner, 2016; Danion & Flanagan, 2018; Fooken, Lalonde, Mann, & Spering, 2018). Conversely, pursuing a ball seems to provide information about the ball’s motion, which is important for predicting its trajectory (Bennett, Baures, Hecht, & Benguigui, 2010; Spering, Schütz, Braun, & Gegenfurtner, 2011) and for controlling the timing and accuracy of interceptive movements directed at the ball (Brenner & Smeets, 2011; de la Malla et al., 2016; de la Malla et al., 2017; Fooken, Yeo, Pai, & Spering, 2016; Leclercq, Blohm, & Lefèvre, 2013). Thus, when dealing with moving targets, the control of eye and hand movements is intricately related to the ability to anticipate how the target will move.

To highlight the role of prediction in the control of eye and hand movements, Fooken and Spering (2019) developed a go/no-go paradigm that requires observers to predict the trajectory of a briefly presented ball. Observers viewed the initial launch of a ball on the computer screen, and instinctively tracked it with their eyes as it moved towards a designed hit zone. Importantly, the ball was only shown for up to 300 ms. If observers thought the ball would traverse the hit zone, they had to intercept it with their finger once it reached the hit zone (go trial; Figure 3a). If observers thought the ball would miss the hit zone, they were not...
to make a hand movement (no-go trial, Figure 3b). Go/no-go decisions were reflected in distinct eye movement responses: Early predictive saccades were more common when a go response was required (bottom panels of Figure 3). The timing of the predictive saccade could be used to predict the decision with >75% accuracy, suggesting that eye movements can provide a sensitive and continuous readout of human decision-making.

**Prediction in guiding the hand during manual interception**

In the previous section we discussed the fact that people predict where important visual information will become available so that they can direct their eyes accordingly. We saw that people are willing to direct their eyes to where they expect new information to become available or to where they expect it to be beneficial to direct their gaze in the near future. They make saccades to where they expect an occluded target to appear or to where they expect a target to abruptly change its direction of motion, and they move their eyes smoothly in anticipation of the target’s motion in order to maintain their gaze on the moving target. Gaze can presumably constantly be directed to where the most relevant information is available because eye movements are fast and not very costly. Indeed, people would rather make additional eye movements than for instance try to memorize simple object properties (Ballard, Hayhoe, & Pelz, 1995). However, making reliable predictions may be more important for guiding other parts of the body than the eyes, precisely because such movement are less fast and more costly. In this section we will discuss prediction in the context of manual interception. To intercept a moving target is to prevent it from continuing on its path, for instance by hitting or catching it. How might predictions help one to do so?

The term “prediction” can be interpreted in many ways in the context of the visual control of action. One distinction that we already mentioned is between predictions about external events, such as how someone approaching you is likely to pursue their path, and predictions about your own actions, which is most evident in relation to determining what information you will need to execute them successfully. That people can predict what information they will need to guide their own actions is evident from studying gaze. There is abundant evidence that whenever people are performing a task rather than just looking at a scene, their eyes are directed towards where information that will be useful in the near future is likely to be found. This has been demonstrated for various everyday tasks such as making tea (Land, Mennie, & Rusted, 1999), making a sandwich (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003), walking (Matthis, Yates, & Hayhoe, 2018) and driving (Land & Lee, 1994; Wilkie & Wann, 2003). It has also been demonstrated for more specialized activities such as reading the score when playing music (Furneaux & Land, 1999). In all these cases, gaze precedes and guides movements of the arm or leg. Similar eye movements guide tasks that do not involve such movements. For instance, when reading (Rayner, 1998), searching for an object (Eckstein, 2011), or identifying faces (Peterson & Eckstein, 2013), observers move their eyes to where they anticipate to find the most relevant information at each moment. In general, people anticipate when they will need certain information (Ballard, Hayhoe, Li, & Whitehead, 1992) and where it is likely to be (Võ & Wolfe, 2015). One could use this to infer when certain information is relevant in tasks for which it is less evident, although there is an obvious danger of falling into circular reasoning.

Predicting the movements or actions of other objects or agents is more difficult. A particularly difficult situation is when the other agent is trying to deceive you, which is sometimes the case in sports (Güldenpenning, Kunde, & Weigelt, 2017; Helm, Weech, Munzert, & Troje, 2016) and magic (Beth & Eklund, 2014). In such cases, the opponent or magician intentionally tries to induce an incorrect prediction. In the laboratory, the experimenter can also bring about incorrect predictions, for instance by manipulating a virtual ball’s acceleration in a way that is inconsistent with gravity (Senot et al., 2012) or even in a way that is unusual but that one might encounter in real life if there is a strong side-wind or if the ball is spinning (Dessing & Craig, 2010). Our experience in the world tells us that objects do not accelerate too abruptly unless they hit other objects, so one might in such cases anticipate that the ball will continue to move at the velocity at which it is moving (Brenner & Smeets, 2018). Doing so could provide reliable predictions if the target is moving at a constant velocity, but as mentioned in the introduction people can intercept targets that are accelerated by gravity (Zago et al., 2008, 2009) and even ones that undergo quite arbitrary accelerations (Brenner et al., 2016).

People anticipate that a flying ball will accelerate downwards due to gravity (McIntyre et al., 2001; Senot, Zago, Lacquinti, & McIntyre, 2005) and that it will bounce when it reaches the ground (Diaz et al., 2013; Land & McLeod, 2000). They appear to be reluctant to ignore gravitational acceleration, even after prolonged experience with its absence (McIntyre et al., 2001). However, they do learn to anticipate a repeated arbitrary acceleration after being exposed to it for only several trials (Brenner et al., 2016). They also appear to learn how far ahead of the target to aim from the feedback on recent trials (Brenner, Cañal-Bruland, &
So, to what extent do people predict how the target will move or how they should move to intercept the target? Or more generally, how are interceptive movements controlled?

**Control of interceptive movements**

There are many ways in which interceptive movements could be controlled. The simplest one is obviously to just continuously move towards the target, without predicting anything about how the target will move. If one constantly aimed at the instantaneously judged target position (dark blue trajectories in Figure 4) one would follow a curved path towards the target. More importantly, due to sensorimotor delays, one would always be aiming towards a position that the target had already passed, and would therefore miss the target. This could be avoided by always aiming slightly ahead of the target. If one were to aim as far ahead of the target as the target travels during the sensorimotor delay, one would hit the target. In this case one would be predicting how the target will move, but for a fixed time. One would not have to predict when one will reach the target. Again, the path would curve in a systematic manner (green trajectories in Figure 4).

In order to avoid moving along a systematically curved path, one has to predict where one will hit the target. If one correctly predicts when and where one will hit the target, one can move straight towards the predicted position and hit the target there (red trajectories in Figure 4). This would appear to be the ideal strategy. However, it is only ideal if one can predict perfectly. There are two reasons why perfect prediction is unlikely. The first is that our sensory judgments have limited precision and are sometimes even biased. The second is that targets may change their motion unexpectedly. One way to deal with the resulting prediction errors is by constantly updating one’s predictions (light blue trajectories in Figure 4).

This allows one to start moving in about the correct direction. As time passes, errors that arise from misjudging the target’s motion decrease as anticipated displacement is replaced by perceived displacement. Thus, even if the sensory input on which the prediction is based does not become more reliable during the movement, the prediction of the endpoint improves because one predicts over an ever-decreasing time interval (until the sensorimotor delay prevents further improvement; Brenner & Smeets, 2015).

This way of controlling movements by constantly updating one’s prediction can deal with changes in the target’s motion as well as with perceptual errors without relying on additional assumptions. It provides performance that is reasonably robust with respect to issues such as ignoring acceleration (Brenner & Smeets, 2018). Assuming that the posterior parietal cortex is critical in updating movements (Pisella et al., 2000), the proposed way of controlling movements can be reconciled with the finding that performance only suffers from disrupting processing of the part of the posterior parietal cortex that corresponds with the instantaneous (rather than a future) target position (Reid & Dessing, 2018) by assuming that the posterior parietal cortex is spatially organized in accordance with the sensory input rather than the predicted movement endpoint (Buneo & Andersen, 2006).

The fact that using simplified models, that for instance do not consider acceleration, can give rise to small but systematic prediction errors need not be a big problem because people readily learn to compensate for recent errors. For instance, in the case of ignoring acceleration, only a few trials are required to eliminate the bias (Brenner et al., 2016). Presumably, if the target is repeatedly accelerating, people simply learn to hit slightly further ahead of where they predict that the target will be on the basis of its judged position and speed.

The main advantage of relying on constantly updated predictions is that these predictions do not have to be very good. Predictions are unlikely to be very good for the reasons that have already been mentioned. A more extreme alternative is to circumvent making predictions altogether by relying on constantly updated associations between specific target properties and fitting motor responses (Dessing, Peper, Lieke, Bullock, & Beek, 2005; Lee, Georgopoulos, Clark, Craig, & Port, 2001; Warren, 2006; Zhao & Warren, 2017). One example of a control mechanism that might lead to successful interception without requiring any prediction is to constantly aim a certain angle ahead of the target (green trajectories in Figure 5).

Although this proposal sounds very similar to aiming a fixed distance ahead of the target (green trajectory in Figure 4), there are a number of fundamental differences. One attractive feature of this proposal is that it ultimately leads one to the target,
irrespective of the chosen angle, because how far ahead of the target one is aiming decreases as one approaches the target. It does so as a result of the distance becoming smaller. For interception with the head rather than the hand, an additional attraction is that it does not require judgments of the distance to the target or anything like that. All it requires is the ability to determine in which direction the target is moving and to regulate one’s movement so that the direction in which one is moving is a fixed angular distance from the target in the direction in which the target is moving.

An alternative but comparable control mechanism that could lead to successful interception without requiring any prediction is to pick an interception point and move straight towards it while adjusting movement speed to keep the angle between the target and where one is heading constant (magenta trajectories in Figure 5). This mechanism only requires one to be able to predict that the target will cross one’s path in a certain direction at some moment in the future. An advantage of such mechanisms, as with any continuously updated mechanism, is that they do not fail nor have to be replaced altogether if the target or person does not move exactly as anticipated. There are, however, a number of serious drawbacks to the more specific control mechanisms illustrated in Figure 5: They do not consider sensorimotor delays, they are quite specific to the task for which they were designed, and they are not easy to expand to consider additional constraints (such as obstacles or limitations to the actor’s own abilities; Brenner & Smeets, 2018). Moreover, on their own they cannot account for people’s ability to intercept occluded targets (Zhao & Warren, 2015).

Predicting sensory consequences of movements

The preceding sections focused on predicting how targets will move in order to guide an eye or hand movement. When discussing eye movements, we mentioned how studying where people look at various moments can tell us when certain input is needed. We also discussed how expertise in acquiring such information might underlie exceptional performance in sports. However, people’s ability to select input on the basis of anticipated task requirements is not restricted to what they can achieve through eye movements. In this section we will focus on other ways in which predictions about the merits of sensory input can influence the sensory input itself. These predictions function across longer time periods as described in the previous two sections, certainly beyond sensorimotor delays. It is important to realize that the mechanisms underlying predictions are complex and can derail. For example, impaired ability to differentiate external from self-produced sensory events has been linked to neuropsychiatric symptoms in schizophrenia (Blakemore, Smith, Steel, Johnstone, & Frith, 2000; Lemaître, Luyat, & Lafargue, 2016; Lindner, Thier, Kircher, Haarmeier, & Leube, 2005; Spering, Dias, Sanchez, Schütz, & Javitt, 2013). This important topic has been covered by recent reviews (Bansal, Ford, & Spering, 2018; Thakkar, Diwadkar, & Rolfs, 2017) and will not be discussed here. Instead, we will review the mechanisms underlying sensory regulation based on normally functioning predictions of sensory action consequences. The ability to predict the sensory consequences of self-generated movements have often been related to efference copy mechanisms (Bays, Flanagan, & Wolpert, 2006; Blakemore et al., 1998; Cullen, Brooks, & Sadeghi, 2009; Crapse & Sommer, 2008), but the principle may be more general, even including predictions based on planned actions or on actions of others.

Predictive models exploit efference copy signals related to a forthcoming movement to generate predictions about the sensory action effects (Wolpert & Flanagan, 2001; Wolpert, Ghahramani, & Jordan, 1995). The predicted sensory consequences are then compared with the actual sensory feedback of the movement: the reafferent signals. The differences are used to correct the current movement as well as future predictions (Pickering & Clark, 2014). They can also be used to help select incoming signals. Whenever the predicted and the actual feedback signals match, the reafferent signal is attenuated so that it is perceived as less intense, presumably to reflect the fact that self-produced action effects convey little novel information. The strength of attenuation is known to be inversely proportional to the error between the predicted sensory feedback and the actual sensory feedback produced by the movement (Claxton, 1975; Weiskrantz, Elliot, & Darlington, 1971). Moreover, sensory attenuation increases with age, indicating a greater reliance on predictive than on sensory signals with age (Klever,
When removing the predictable component from the incoming sensory signals, the relative salience of unpredictable and potentially harmful external events is increased. Enhancing such unanticipated effects might be important for survival. Predicting the sensory consequences of actions also helps to attribute agency and ownership to movements, actions, and their sensory consequences: to distinguish movements that one generated oneself from those generated by another agent (Bansal et al., 2018; Farrer, Bouchereau, Jeannerod, & Franck, 2008; Farrer, Valentin, & Hupé, 2013; Kilteni & Ehrsson, 2017; Lindner et al., 2005). The attenuation of anticipated sensory signals facilitates this distinction because whenever the predicted and the actual sensory signals mismatch, the afferent signals are not attenuated.

**Predictive signals change perception**

One of the prime examples of predictive sensory attenuation is the phenomenon that we cannot tickle ourselves. Humans experience a touch on their own body as less ticklish if they produce the tactile stimulation themselves (Blakemore, Frith, & Wolpert, 1999). If a brief temporal delay or a slight change in the spatial pattern of the touch is introduced between the self-produced action and the action effect, humans experience a stronger tickle sensation. Sensory attenuation is temporally tuned and centered on the expected time of the action effect; i.e., it is highest at the time of the expected action effect and reduced with an increasing temporal delay between the expected and the actual time of occurrence. In a seminal study by Bays, Wolpert, & Flanagan (2005), participants tapped a force sensor, mounted above their left index finger, with their right index finger. When a motor generated a tap on the left finger synchronously with the tap performed by the right finger, the perception of force in the left finger was strongly attenuated (central dip in Figure 6a). This attenuated sensitivity gradually decreased as the motor tap on the left finger was either delayed (right section of curve in Figure 6a) or advanced (left section of curve in Figure 6a) relative to the performed tap of the right finger. Importantly, reduced sensitivity can also be observed when the action is aborted or when the action is performed but fails to have the anticipated effect (Bays et al., 2005; Bays et al., 2006). This highlights a strong predictive component linked to an expected action effect.

Attenuation of such predicted sensory consequences of a movement has frequently been found in audition and somatosensation, e.g., we hear our own voice less loudly (Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000) or perceive a self-produced touch as less forceful (Shergill, Bays, Frith, & Wolpert, 2003). Corresponding observations in vision are rather scarce. Cardoso-Leite, Mamassian, Schütz-Bosbach, & Waszak (2010) reported attenuation in the visual system for learned associations between self-generated actions (left or right key press) and visual action effects (left or right tilted Gabor patches). Visual sensitivity was substantially reduced for the Gabor patches when these stimuli were triggered by the action they had previously been associated with. Moreover, the visual event-related N1 component is attenuated when the visual feedback of the participant’s reaching arm is provided in real-time compared with when it is delayed, creating a mismatch between the predicted and actual visual consequences of the movement (Benazet, Thénault, Whittingstall, & Bernier, 2016). These findings demonstrate that self-
generated actions change the perception of predicted action effects.

Sensory attenuation has also been observed for externally applied tactile stimuli. For example, an externally applied touch stimulus is perceived as weaker and as later in time when the limb is moving than when the limb is at rest. This has been shown for a wide variety of visuo-motor tasks, such as reaching (Fraser & Fiehler, 2018; Gertz, Voudouris, & Fiehler, 2017), grasping (Colino & Binsted, 2016; Colino, Buckingham, Cheng, van Donkelaar, & Binsted, 2014; Voudouris, Broda, & Fiehler, 2019), juggling (Juravle & Spence, 2011), and during gait (Duysens et al., 1995; Morita, Petersen, & Nielsen, 1998). Sensitivity to the external stimulus typically starts decreasing around 100 ms before movement onset and reaches its minimum during the movement (Buckingham, Carey, Colino, Degrosbois, & Binsted, 2010; Figure 6b). Because tactile attenuation of external stimuli is also present during passive movements, it has been questioned whether it arises from a predictive mechanism (Chapman, Bushnell, Miron, Duncan, & Lund, 1987; Williams & Chapman, 2002). An alternative is that the processing of movement-related reafferences during movement execution leads to backward masking of the external stimulus and thereby to reduced sensitivity to the stimulus. Contrary to this alternative, reduced tactile sensitivity on a moving limb has been observed when a movement is planned (Buckingham et al., 2010; Jackson, Parkinson, Pears, & Nam, 2011; Voss, Ingram, Wolpert, & Haggard, 2008), planned but not executed (Haggard & Whitford, 2004; Voss, Ingram, Haggard, & Wolpert, 2006) or just imagined (Kilteni, Andersson, Houborg, & Ehrsson, 2018), arguing for a central predictive component.

Efference copy mechanisms can generate predictions about action consequences both within a single sensory system and across different senses. Using a temporal prediction task, van Kemenade, Arikan, Kircher, and Straube (2016) asked participants to report a delay between a self-initiated button press and a sensory action effect that was either a dot on the screen, or a tone, or both. Delay detection was better for the combined visual-auditory than for the unisensory visual or auditory action effect. These findings support the notion of enhanced prediction based on signals of multiple sensory action effects. Further evidence for a multisensory predictive mechanism comes from brain imaging studies showing that predicted compared to unpredicted visual-auditory action consequences lead to less activation in the respective sensory cortices (Straube et al., 2017). The angular gyrus seems to function as a supramodal comparator area evaluating the discrepancies between the predicted and the actual feedback signals from different sensory modalities (van Kemenade et al., 2017; for evidence for unisensory visual feedback comparison, see Farrer et al., 2008; Leube, Knoblich, Erb, & Kircher, 2003).

**Task relevance influences predictive sensory attenuation**

Given the importance of predictive sensory attenuation in highlighting novel and unpredictable events, deteriorated sensory feedback about the outcome of one’s own actions may impair sensorimotor learning and control. This dilemma can be solved by dynamically adjusting the strength of sensory attenuation with respect to the relevance of the predicted sensory signals for a given task. For example, when performing a precision grasp, tactile sensitivity is less attenuated at the index finger, which is involved in the grasp, compared to the pinky or the forearm of the grasping limb (Colino et al., 2014; Colino & Binsted, 2016; Juravle, Colino, Meleqi, Binsted, & Farné, 2018). This effect is temporally tuned to the time shortly before the start of the grasp, i.e., the time period critical for movement planning. Reduced tactile attenuation can also be found in haptic exploration when tactile information is required to discriminate object surfaces made of different materials (Juravle, McGlone, & Spence, 2013). These results show that sensory signals relevant for the task are less attenuated when they are needed for successful task performance. Presumably this is also true for signals that are used to adjust or compare predictions.

Such task-related decrease of sensory attenuation of predicted movement consequences may be associated with an increase in attenuation of task-irrelevant information. Gertz and colleagues (2017) observed stronger tactile attenuation on the moving limb when participants reached to their other hand than when they reached to an external target without visual feedback. They argued that the increase of tactile attenuation on the moving limb may free capacities for enhanced processing of somatosensory information of the target hand which served as movement goal. In line with this hypothesis, enhanced tactile sensitivity could be observed at the target hand during reaching (Voudouris & Fiehler, 2017a, 2017b). This potential compensatory mechanism seems to be restricted within one modality (at least for somatosensation), as additional processing of task-relevant visual information failed to increase tactile attenuation on the moving limb (Gertz, Fiehler, & Voudouris, 2018). Overall, these findings show that predictive sensory attenuation is not an all-or-nothing mechanism but can be flexibly adjusted to the task requirements, a phenomenon that can also be found for predictive mechanisms in eye movements and interceptive movements.
One evident distinction is between presenting images on a monitor to a static observer, as opposed to having an observer interact with real objects in the real world without requiring that the observer remains static. There is recent evidence showing fundamental behavioral (Gomez, Skiba, & Snow, 2018; Squires, MacDonald, Culham, & Snow, 2016) and cortical differences (Freud et al., 2018; Gomez et al., 2018; Snow et al., 2011) in how humans process real objects compared to pictures of the same objects. Real-world environments are not only more complex in terms of visual structure, but also provide additional information (e.g., from motion parallax; de la Malla et al., 2016; Graham & Rogers, 1979). Thus, more natural circumstances make it necessary to consider head and body movements and vergence eye movements as well as horizontal and vertical eye movements. Through evolution and a lifetime of experience, humans are presumably adapted to their natural environment, so we need to study behavior under these conditions to elicit the processes of prediction that we want to understand.

Many years have passed since the early seminal studies describing eye movements during everyday tasks (Hayhoe & Ballard, 2005; Land, 2006; Land & Hayhoe, 2001). The technology for mobile eye tracking and motion capture has become much more accessible, but the number of studies that make use of such technology to systematically and carefully investigate predictive strategies during tasks such as walking (Matthis et al., 2018), driving (Tuhkanen et al., 2019), or ball sports (Diaz, Binaee, & Phillips, 2016; Higuchi et al., 2018) is still quite limited. Combining such technology with virtual or augmented reality can provide experimenters with precise control of sensory variables of interest and allow them to manipulate target features such as ball elasticity or gravity (Diaz et al., 2013; Jörges & López-Moliner, 2019; Russo et al., 2017), while providing an immersive experience mimicking real-world settings. Virtual reality is especially valuable for creating conflict situations in space and time, for example between the proprioceptive and the (virtual) visual feedback of one’s own body movements (Ma & Hommel, 2015), to test the flexibility and the limits of sensory prediction of action consequences. Control of task requirements can be achieved with specific equipment, such as using ball-launching devices to control ball trajectories in order to assess how such trajectories influence observers’ eye and interceptive movements (Cesqui, Mezzetti, Lacquaniti, & d’Avella, 2015). Moreover, the ability to automatically process detailed movement trajectories makes it possible to interpret performance in less precisely controlled circumstances (López-Moliner, Brenner, Louw, & Smeets, 2010).

The complexity of prediction under natural circumstances requires new computational techniques that are suited for quantifying and analyzing human behavior in more complex tasks involving multiple goals (López-Moliner & Brenner, 2016), multisensory action effects (van Kemenade et al., 2016), or sequential dependencies (Matthis et al., 2018). There has been a recent breakthrough in modeling sequential behavior when mastering the board game Go (Silver et al., 2016). However, modeling dynamic human behavior in everyday tasks raises an even bigger challenge as we need to consider sensorimotor delays as well as a variety of expected and unexpected uncertainties, factors that challenge successful predictions. There are recent advances in studying predictive mechanisms in sequential human behavior in the real world and in virtual environments (for a review, see Hayhoe, 2018). Recurrent neural networks (RNNs) provide a tool to model predictive behavior by incorporating trial-by-trial behavioral variations. Results of these models even argue against the hypothesis that accurate prediction requires internal models of physical dynamics (Zago et al., 2009), but instead suggest a discrete mapping between information integrated over time and a temporally distant motor output (Binaee, Starwynska, Pelz, Kanan, & Diaz, 2018). This example shows that studying natural behavior can add significantly to our current understanding of predictive mechanisms in goal-directed actions.

Predictions over longer time scales (seconds to minutes) are especially likely to benefit from more complex and natural circumstances and analyses. However, unnatural, controlled circumstances can sometimes reveal mechanisms that might not be detected in natural situations. One disadvantage of very
complicated experiments, circumstances, or models is that one may miss hidden assumptions and therefore draw incorrect conclusions from the data.

Some of the research presented in this review has demonstrated the importance of eye movements for vision and action, and their potential role as models of sensorimotor prediction. For eye movements to fulfill this role, a few things have to be better described and understood. For example, most studies on the role of eye movements in interceptive tasks focus on hand movements and only provide rudimentary descriptions of what the eye does (e.g., where a saccade lands relative to the finger). A more detailed analysis of the eyes’ temporal movement properties would enhance our understanding of the link between both systems. Such behavioral studies could be paralleled by neurophysiological primate studies that employ new techniques to record the activity of large populations of neurons across multiple brain areas (Pesaran et al., 2018). The basic cortical and subcortical substrates of eye and hand movement control are well known (Battaglia-Mayer & Caminiti, 2018). Future studies could now aim at identifying detailed functionality in terms of sensory signal processing and motor output generation across multiple brain areas. Such research could have tremendous impact on our understanding of disease states, for example, deficits of predicting the sensory consequences of one’s own action in schizophrenia, which likely involve many different regions in the brain (Bansal et al., 2018; Thakkar et al., 2017).

Another evident future direction is to further develop realistic models of how arm and eye movements might be controlled, based on various assumptions, and to use this knowledge to identify primary underlying mechanisms. This is very much a question about prediction, because one of the main debates in motor control is about the extent to which movements are planned based on predictions rather than emerging from simple sensorimotor strategies. Many of the important topics are related to this question, such as to what extent ongoing movements are controlled, to what extent movements are optimized, to what extent knowledge about the world is considered, how expertise can best be achieved, and so on. We are not yet certain about which information is used and which is not. Moreover, although we better understand where people look when performing a wide variety of everyday tasks (de Haas, Iakovidis, Schwarzkopf, & Gegenfurtner, 2019; Matthis et al., 2018), how people learn to look at specific places and act at exactly the right moment is still largely unknown.

Keywords: prediction, smooth pursuit eye movements, manual interception, sensory attenuation, efference copy

Acknowledgments

This work was supported by the German Research Foundation-Collaborative Research Centre SFB/TRR 135, Project A4 and the DFG FI 1567/6-1 (TAO, The active observer) to KF, by NSECR Discovery Grant and Accelerator Supplement to MS, and by project number 464.18.111 (TAO, The active observer) of the Netherlands Organization for Scientific Research (NWO) to EB.

Commercial relationships: none.

Corresponding author: Katja Fiehler.

Email: katja.fiehler@psychol.uni-giessen.de.

Address: Department of Psychology, Justus Liebig University, Giessen, Germany; Center for Mind, Brain, and Behavior (CMBB), Universities of Marburg and Giessen, Germany.

References


Barnes, G. R., & Collins, C. J. S. (2008). Evidence for a link between the extra-retinal component of random-onset pursuit and the anticipatory pursuit of
predictable object motion. *Journal of Neurophysiology*, 100(2), 1135–1146.


updating one’s predictions underlies successful interception. *Journal of Neurophysiology*, 120(6), 3257–3274.


de Haas, B., Iakovidis, A. L., Schwarzkopf, D. S., &


Kuhn, G., & Land, M. F. (2006). There’s more to magic than meets the eye. *Current Biology*, 16(22), R950–R951.


7, 1–16, https://doi.org/10.1167/11.1.7. [PubMed] [Article]


