CHAPTER

Accumulating visual information for action

4

Eli Brenner[1], **Jeroen B.J. Smeets**
Vrije Universiteit, Amsterdam, The Netherlands
[1]*Corresponding author: Tel.: +31-20-5982571, e-mail address: eli.brenner@vu.nl*

Abstract

In this chapter, we discuss the way in which visual information is gathered and try to relate this to the task at hand. It is well established that people direct their gaze toward the places at which they expect to be able to gather the most useful information. Studies of gaze during goal-directed actions show that people also make sure to gather information precisely at the moment that they need it. We argue that the eye movements that people make during interception tasks and the precision that people achieve in such tasks suggest that people constantly update their estimates of the details that are needed to successfully hit the target. The updating predominantly consists of replacing the previously acquired information, rather than of adding to such information.

Keywords

Visual representation, Updating, Interception, Gaze, Human

1 INTRODUCTION

Many of our everyday actions rely on visual information. Our access to such information depends on how we ourselves move, including where we direct our eyes. Consequently, accumulating visual information is an active process. The information that we acquire at any instant is limited, so we should ensure that such information is always the information that is most important at that instant. However, there might be other important information elsewhere, or other important information may have been available at a different time, so it may be worthwhile to also retain some information for later use. We will use the term *visual representation* to refer to the information that has been acquired, and possibly retained, rather than to the information that is potentially available. Before discussing how such visual representations might be updated, it is useful to consider the possible nature of the representations. How might such information be represented?

Progress in Brain Research, Volume 236, ISSN 0079-6123, https://doi.org/10.1016/bs.pbr.2017.07.007

2 THE NATURE OF VISUAL REPRESENTATIONS

We know that many brain areas are activated by visual stimulation, each in its own specific way with its own specific latency (Schmolesky et al., 1998). Most of what we understand about such activation is related to how it arises rather than to how it is used. We understand that the mapping between the image on the retina and its representation in various areas of the brain is not uniform due to the inhomogeneity of the receptors in the retina and of the further processing. We understand that the representation of color arises from the presence of different cone types in the retina and from the further processing having to separate surface properties from properties of the illumination. We understand that the representation of motion relies on correlations between changes at different places at different times. However, we understand much less about how all this allows us to recognize objects and interact with them. One reason for this may be that it is difficult not to think about all these representations as providing information about our surrounding, rather than thinking of them as tools for interacting with our surrounding. Approaching visual representations in terms of how they could be used to guide our investigations of—and actions toward—the objects that surround us might help us understand more about our ability to recognize objects and interact with them (O'Regan and Noë, 2001).

Much of the information that falls on the retina, and therefore presumably influences the activity in at least some of the visual areas of the brain to some extent, does not appear to influence people's judgments or actions. We therefore consider it to be important to understand how the visual information that does have an influence on judgments and actions is selected and used. The prototypical example of the fact that not everything that is represented on the retina and therefore presumably also to some extent in several visual areas of the brain is actually "used" is the failure of many people to detect a gorilla walking through a scene when counting the number of times that people within the scene throw a ball to each other (Simons and Chabris, 1999). If they had not been counting the throws but just looking at the scene, they would probably have all noticed the gorilla, because it was quite conspicuous. That people only observe, or at least only remember having observed, what is relevant to them at any particular moment can be illustrated by many more examples. For instance, when searching for a certain object, looking at an object that is not what one was looking for discourages people from looking at that object's position again (Peterson et al., 2001). Presumably, people remember where they have already looked, because doing so can evidently be important. However, having looked at the object that was not what one was looking for does not help one find that object if it later becomes the target of the search (Võ and Wolfe, 2012). This is probably because at the time there was no way to tell that remembering what one had seen might become important. The use and retention of visual information is apparently quite selective.

Another limitation of the retention of visual information is that it cannot be very detailed. When asked to detect a change in a static scene, so that one is aware that all items are potentially relevant but the item that changes is not known, large changes can go by undetected if the change itself occurs too slowly to be noticed or is masked by a blink, saccade, or some other transient (O'Regan et al., 1999; Rensink, 2002; Simons and Rensink, 2005). Masking is necessary if the changes are abrupt because particularly salient things, such as sudden changes, are noticed even if they are irrelevant to the task at hand (van der Stigchel et al., 2009). Importantly, even looking at the right place is not always enough. People can miss dramatic identity changes across cuts in films (Levin and Simons, 1997). They can even miss identity changes when someone else appears after the person they were talking to temporarily disappears from sight (Simons and Levin, 1998). In the latter cases, the identity of the person in question is probably not relevant enough at that moment, so constancy is assumed rather than perceived. Thus, our visual representations retain very little detail, even for a short time and even of what we were looking at.

That visual information is only processed at the moment that it is needed is also nicely illustrated by a study in which virtual bricks of two sizes had to be grasped and placed elsewhere (Triesch et al., 2003). In some trials a brick changed size once it had been grasped, near the time at which the participant moved his or her eyes to the position at which it was to be placed. The eye movement masked detection of the change in size. When the task was to first move the large bricks to one position and then the small ones to a different position, so that the size of the brick was relevant both when grasping and when placing the bricks, participants often placed bricks that had changed size at the wrong position. They failed to notice that the size no longer matched the position at which they were placing the brick, although they were looking at the brick when placing it at the location that matched the original size. Thus, not only are visual representations limited to details about the surroundings that are directly relevant to the task, but such details are only considered at the moment at which they are essential. Presumably, once participants had decided where to put the brick, there was no longer any need to consider its size. People do not simply ignore all information that is not directly relevant for guiding the current movement once the movement is underway, because if options for more effective movements arise they respond quickly and adequately (Brenner and Smeets, 2015a). Thus, it is the anticipated relevance for attaining their goals that determines whether information is used.

The previous examples show that there is more to selecting information than only directing gaze. Nevertheless, where one looks does obviously strongly determine what one sees, so the information that one wishes to acquire should direct gaze to appropriate locations. One reason for maintaining a persistent representation of one's surroundings is for knowing where to look for certain information when such information becomes relevant. When relevance has an external cause, such as a car suddenly heading toward you or someone calling your name, the instantaneous sensory information can direct your response toward the position of interest. However, if you decide to get the milk from the fridge (Land et al., 1999), you need to know that the milk is in the fridge and that the fridge is behind you.

Examining people's gaze behavior is one way to try to deduce what information they were trying to acquire. When there is no task, gaze is initially directed to the center of the scene (Tatler, 2007) or to salient objects in the scene (Peters et al., 2005). If one knows what one is looking for but not where it is, gaze is directed to objects that resemble what one is looking for in some way (Becker and Lewis, 2015). If one knows where an object of interest is located, one obviously looks there (Aivar et al., 2005). The ability to direct one's gaze appropriately is probably largely based on peripheral vision, supplemented by specific experience with various aspects of the task in similar environments. In what form such supplementary information is available, and what kinds of representations this involves, is unclear. Different aspects of such information are undoubtedly represented in different ways in separate parts of the brain. In some way, the task at hand must determine which aspects of such supplementary information are relevant, and access such information together with the latest visual input accordingly. Moreover, it must do so at the correct time.

Experiments have shown that people would rather make additional eye movements than rely on an enduring visual representation (Ballard et al., 1992, 1995). However, people must have some enduring representation of their surroundings because they do rely on enduring visual representations to some extent (Aivar et al., 2005; Brouwer and Knill, 2007, 2009). If people had not had some kind of enduring visual representation, they would not know where to look for relevant items that are not visible at that instant. People direct their gaze in anticipation of where relevant information for the next action is to be found (Hayhoe et al., 2002, 2003; Land et al., 1999; Mennie et al., 2007), even if there is no useful information at that position at the moment that gaze reaches that position (for instance when making saccades to a future bouncing point; Diaz et al., 2013a,b). This indicates that they have a sufficiently good representation of the static and dynamic aspects of the scene to direct their eyes appropriately.

That there are diverse visual representations within the brain is hardly controversial. There is abundant evidence from neurology and neuroscience for the existence of many anatomically segregated visual areas in the brain. It makes sense to process different attributes in different brain areas because different visual attributes must be coded in fundamentally different ways. That different attributes are processed in different ways is evident from the fact that one sometimes observes inconsistencies between perceived attributes (Brenner and van Damme, 1999; de Grave et al., 2004; Gillam and Chambers, 1985; Smeets and Brenner, 2008; Smeets et al., 2002). For example, a moving background can influence a target's apparent motion without influencing its apparent position accordingly (Smeets and Brenner, 1995). That different attributes are processed in different ways is also evident from the information that is considered. When parts of an object are presented in rapid succession while the eye's orientation is changing, the object's overall apparent shape is determined by the parts' relative positions in the retinal image, completely ignoring the changes in eye orientation. However, the eye orientation at the time of the presentation *is* considered when judging the position of the object as a whole (Brenner and Cornelissen, 2000). If we accept the fact that there are diverse, fundamentally different visual representations, there is no reason to expect all the representations to use the same reference frame. Similarly, there is no reason to expect all the representations to be updated in the same manner.

3 UPDATING REPRESENTATIONS

There are several ways to update a representation. The way in which a representation is updated obviously depends on the reason for updating the representation: doing so when looking around in order to maintain a gaze-centered representation despite moving one's eyes is different from doing so while fixating to accumulate information in order to improve the reliability of a representation. The way in which information is updated probably also depends on the kind of representation: when considering an object's motion, there is no point in accumulating local information in a putative allocentric frame of reference, because the object is constantly changing position in such a representation, whereas it might be meaningful to accumulate local information in a retinal frame of reference if one is pursuing the object with one's eyes (as in Brenner and Cornelissen, 2000), and it would be meaningful to do so in a putative object-centered representation.

The simplest and probably most ubiquitous method of updating a representation is by "overwriting" it. Transient representations must obviously continuously be updated, because otherwise they disappear. The measured physiological responses in many areas of the brain, including the lateral geniculate nucleus and the initial visual cortical areas, appear to respond to visual stimulation in the way that one would expect from such representations. These representations persist to some extent due to the sluggishness of retinal processing, and are not completely oblivious to earlier stimulation due to adaptive processes, but they are inherently transient. Of course, only input that is present at that instant is represented in such a representation, so task-dependent gaze patterns influence the content. More enduring representations might also be updated by overwriting them, but in that case the representation will not change when there is no new input. Presumably, a representation of your teacup's position on the table, or of its color, does not need to be constantly updated in order to be useful in the future. When it is updated, it is reasonable to overwrite the location or color, rather than to remember all previous values. However, there might also be conditions in which it is beneficial to retain previous values. What can we expect of enduring representations and how are they likely to be updated?

As already mentioned, one might need an enduring representation for knowing where to look for certain information when it becomes relevant, such as the example given above of knowing that the milk is in the fridge and that the fridge is behind you. Such knowledge could reside in a representation of where objects are relative to you. If so, the representation must be updated whenever you move (Clemens et al., 2012; Medendorp et al., 1999; Schutz et al., 2013; van Pelt and Medendorp, 2007). Such updating would have to include everything that might become relevant. However, such representations cannot remain reliable for very long, because updating is needed with every movement. As updating cannot be perfect, errors will quickly accumulate and thereby render the representation useless after several movements (Smeets et al., 2006). One way to increase precision might be to consider relative positions as well as positions relative to oneself (Fiehler et al., 2014; Klinghammer et al., 2016). One might even altogether store information such as that the milk is in the fridge and the fridge is opposite the sink in terms of relative positions.

An advantage of relying on relative positions is that the representation only needs to be updated when something relevant changes in the environment, or when one detects new relevant things in one's environment. Having established that there is a cake on the table, there is no need update anything about the cake until something changes, such as seeing someone move it or eat it. Details, such as that it is a chocolate cake, are unlikely to ever need to be updated, because they are unlikely to change. However, such a representation does not tell you where things are relative to yourself. In order to use the information to guide your gaze (or other actions), you would either have to keep track of your own position within the remembered scene or have to deduce your position from what you see at each instant, or consider both. What you see at each instant is a combination of what you are looking at and what you see in less detail in peripheral vision. Combining this with keeping track of your actions on the basis of your intentions and the sensory feedback can probably normally provide a reasonable estimate of your position within the scene. Combining what you see with knowing where you are, or at least knowing what sort of environment you are in, might usually be enough to make use of such a representation.

It is evident that not everything that one sees is stored in an enduring representation. The details that are stored in some such representation, and therefore that one remembers about the circumstances at a certain moment, are probably determined by whether one considered remembering those details to be important at the time. Consequently, most details are not remembered. This is consistent with the fact that once one knows what is changing, masking the change (Simons and Rensink, 2005) no longer prevents one from seeing it. It also explains why having looked at an object does not help one later find it unless its identity was already known to be important when one first looked (Võ and Wolfe, 2012). The critical issue is not whether the details were considered to be relevant at the time, but whether remembering them was considered to be relevant at that time. Obviously, when looking for a certain object, other objects' identities are determined to some extent, in order to judge whether they match the object that one is looking for. However, if they do not match the target object, there is no need to remember what they were, as long as one remembers not to look at them again. Similarly, not noticing that a bar has suddenly become smaller just before one places it at a location that is reserved for large bars (Triesch et al., 2003) can be explained by people no longer considering remembering the size to be important once they know where to put the selected bar. Thus, the existence of some form of enduring visual representation is hardly controversial, but the amount of detail that is stored in such a representation is probably quite limited and how such an enduring visual representation is achieved is still largely unknown.

As already mentioned, most of the measured neuronal responses to visual stimulation rapidly terminate when the target disappears or when eye movements shift the target's image to another part of the retina. Thus they appear to correspond with transient gaze-centered representations. There are also some neurons that respond to visual stimuli at certain positions relative to the head, body, or even hand, irrespective of the orientation of the eyes (Fattori et al., 2004, 2009; Graziano, 2001; Lacquaniti and Caminiti, 1998; Thompson and Henriques, 2011). Some such neurons even appear to contribute to an enduring representation, in the sense that the response does not stop when the target disappears, but this is probably because their activity is related to the upcoming movement rather than to the stimulus (Fattori et al., 2004; Ryou and Wilson, 2004). The eyes' orientations and the orientations of other parts of the body must be considered when using retinal stimulation to guide actions such as arm movements, but there is little evidence for enduring head-, body-, or hand-centered visual representations. Experiments on how shifts in gaze influence remembered locations are more consistent with a gaze-centered representation (Henriques et al., 1998). Moreover, the advantage of a head-, body-, or hand-centered representation is not evident because such representations would need to be updated whenever one moves one's head, body, or hand.

It is often assumed that people have some kind of gaze-centered enduring visual representation that is updated when they move their eyes. This assumption has led to many questions as to how information is combined across eye movements and how we perceive a stable world despite the images in our eyes constantly changing (Collins et al., 2009; Deubel et al., 2002; He et al., 2017; Herwig, 2015; Irwin et al., 1990; Melcher, 2007; Niemeier et al., 2003; Oostwoud Wijdenes et al., 2015). Much of the literature on the updating of visual representations is concerned with these questions. However, these are only problems for enduring gaze-centered representations. If people do indeed combine large amounts of gaze-centered information across time, rather than primarily relying on what they see at each moment, they must constantly update the enduring gaze-centered representation as they move their eyes. If they also move their head, they must also update the representation in accordance with their head movements, and so on.

Probably the most compelling support for the existence of a gaze-centered representation that is updated when the eyes move is that various visual neurons' receptive fields shift temporarily just before saccades (Duhamel et al., 1992; Kusunoki and Goldberg, 2003; Walker et al., 1995). In accordance with this, flashes presented in total darkness near the time of saccades are judged to be at positions that are consistent with combining the retinal stimulation with a later orientation of the eyes (Honda, 1993; Mateeff, 1978; Matin et al., 1970; Schlag and Schlag-Rey, 2002), and the retinal position that is considered to be relevant for a perceptual task (as revealed by enhanced performance at that position) shifts in anticipation of the eye movement (Mathôt and Theeuwes, 2010; Rolfs et al., 2011). However, it is not clear how such a transient shift can help construct an enduring representation.

People associate flashes with a later orientation of the eyes not only during saccades but also during smooth pursuit (Brenner et al., 2001; Rotman et al., 2004). They also associate flashes with a later position of a moving target when the eyes are not moving at all (Nijhawan, 1994). The similarity between these three circumstances suggests that the perceptual errors are related to flashes rather than to eye movements. Perhaps flashes are judged to have occurred later than they actually occurred. This might be so because the sluggishness of the retina makes it impossible to distinguish between the extremely short bright flashes that are used in this kind of experiment and less intense ones that last longer (Brenner and Smeets, 2010). Another possible explanation is that because the orientation of the eye or the position of the moving target is constantly changing, it has to be sampled. In a typical experiment, this sampling will occur in response to the flash, so that only positions some time after the flash are considered (Brenner and Smeets, 2000).

The shift in the region of enhanced processing just before saccades (Mathôt and Theeuwes, 2010; Rolfs et al., 2011) could be considered to be the result of selecting a different region as being relevant for the task at that moment (which is what is often referred to as a shift of attention). The eyes then follow this decision, so the enhanced processing precedes the eye movement. This interpretation gives the selection of information a prominent role in guiding eye movements. Following this reasoning, the task will dictate the choice of attributes and positions that are considered to be important, and therefore what information is considered (and therefore noticed), what additional information one tries to acquire by moving one's eyes and body, and how to move to fulfill the task itself. A persistent visual representation in terms of what is present (there is a chocolate cake) and of relative positions (the cake is on the table) is useful because it is not affected by changes in gaze or self-motion and therefore does not need to be updated whenever we move. As for the perceived stability of the world when we make eye movements, we are probably not confused by images flying across the retina during saccades because the motion is too fast and the new image quickly overwrites the old one (Matin et al., 1972) rather than because we compensate for the change. Our movements, including those of our eyes, are considered whenever we judge objects' positions, so positions after a movement are consistent with those before we moved. In fact, consistency between what we saw in peripheral vision before saccades and what we see in foveal vision after saccades is probably even used to compensate for errors in judging the magnitudes of our movements (Deubel et al., 1998, 2010). The world is assumed to be stable, while the retinal image changes dramatically.

Rather than updating a representation to keep the representation up to date, one might also update a visual representation to improve one's judgments of selected measures. In that case, the representation will not be updated when one shifts gaze, but will be updated whenever one does not. In sports, better performance is associated with keeping looking at specific points; and people are even trained to do so (Gonzalez et al., 2017; Lebeau et al., 2016). We know that people look at a position for a longer time if they suspect that the visual processing will take more time (Becker, 2011; Hooge and Erkelens, 1998). However, it is not clear why looking longer helps. Some attributes take more time to judge than others, and combining several attributes may improve precision. An example of an attribute that might take relatively long to judge is retinal motion, because in order to judge retinal motion, the brain has to detect changes at two different times and places (van de Grind et al., 1986; van Doorn and Koenderink, 1982). Another possibility is that information about a single attribute is somehow accumulated. This would require an enduring gaze-centered visual representation but not necessarily one that is updated when the eyes move. A final possibility is that the spatiotemporal sensitivity of the retina makes what we see depend on where we have just looked, so the precise eye movement pattern during fixations, including slow drifts and microsaccades, influences how the retinal stimulation is processed (Rucci and Victor, 2015; Rucci et al., 2007). The retina is particularly sensitive to change. Large changes in fixation give rise to large changes at all spatial frequencies. Small changes in fixation only give rise to large changes at high spatial frequencies. Thus, the signals from the retina depend not only on where one is looking but also on how long one has been looking near there. Looking longer at the same place with some drift and microsaccades increases the relative contribution of high spatial frequencies in the retinal signal and therefore increases the sensitivity to fine detail at that position (Boi et al., 2017). Thus, even without any representation that is integrating information across time, it might be advantageous to look longer if one needs to detect finer details. This may be responsible for how long one looks being task dependent.

Thus, altogether, it would appear that we mainly rely on the information that is instantaneously available (Horowitz and Wolfe, 1998). Some information that is considered to be likely to be useful in the future is retained, but in quite minimal detail, presumably partly in a form that is independent of gaze. This is necessary to not keep looking in the same place when you misplaced your mug, or briefly remembering where an object appears to be relative to oneself when relating its apparent position before a saccade to what one sees after the saccade. The positions and attributes that are deemed relevant at any instant appear to largely determine how one directs one's gaze and therefore what information is available to guide our judgments and actions. One consequence of primarily relying on instantaneous information is that precisely *when* one looks *where* can be critical for the control of our actions.

4 CONTINUOUS CONTROL OF ACTION

An action that is very suitable for studying the timing of the use of information is interception. To intercept a target, people must combine the perceived target position and velocity to anticipate where the target will be some time in the future. Any error in judging the position or velocity, or in judging one's own capabilities, will have consequences for the progression or even the outcome of the action. If one is constantly judging the position and velocity and using these judgments to update the estimate of where to hit the target, most of the random components in the above-mentioned errors will average out as the ever-changing estimate is smoothed by the mechanical properties of the arm, resulting in an outcome that depends on a weighted average of previous estimates. Later estimates will generally have more impact, except when they are so late that the arm can no longer respond. Thus, in contrast to what one would expect if people would improve their judgment of the velocity by accumulating visual information about the velocity across time, one would not expect seeing a target that is moving at a constant velocity for a longer time *before* one approaches it to improve performance. Indeed, when asked to tap on such a target when it reached a specified position, whether it was visible for 1200 ms or only for 600 ms before reaching that position did not influence how precisely participants hit at the right time (Brenner and Smeets, 2015b). One would expect seeing the target longer, *while* one is approaching it to matter, and indeed how long before the hit such a target disappears from view does clearly matter (Brenner and Smeets, 2011), as does whether a falling ball is visible for 530 ms or only 370 ms before being hit (Brenner et al., 2014). Relying on continuously updated judgments of the position and velocity makes it essential to keep updating the information until doing so becomes useless due to sensorimotor delays (Brenner and Smeets, 2015b). This is probably why people naturally follow the target with their gaze (Brenner and Smeets, 2007, 2009a; Soechting et al., 2009) and seldom blink while trying to intercept it (but rather delay the blink until just after the tap; unpublished observation).

In our experience with interception tasks, the only way to reliably stop people from looking (exclusively) at the target is to make the required precision at another position even more important for the task. If another object than the target is relevant to the task, but its position does not need to be determined with more precision, people occasionally take a quick look (Brenner and Smeets, 2009a), but such glimpses are brief and rare. If people have to intercept a large target at a very precise position, they direct their gaze toward the position at which they have to tap. Since doing so reduces the precision of judgments of the target's position (due to its retinal eccentricity) and possibly of its velocity, the precision in timing the interception becomes lower when one does so (Brenner and Smeets, 2011, 2015c).

From the preceding findings, one might expect that if people must look elsewhere at some time while trying to catch a ball, they would always do so as early as possible. In a study that was designed to examine this, five of six participants tested did indeed predominantly shift their gaze away from the ball as early as possible, but the sixth participant did so as late as possible. Nevertheless, this participant did not drop exceptionally many balls (López-Moliner and Brenner, 2016). Why not? In particular, why did participants not always look away early if it is critical to look at the ball later in time? The reason is probably that if one's prediction is very incorrect and one only realizes that it is incorrect just before the last moment at which corrections can still be made (due to sensorimotor delays) there will no longer be time to make the required adjustments. Normally, constantly adjusting one's movement to the instantaneous estimate helps achieve a high precision, because continuous small adjustments integrate across time to provide a precise action. It is therefore useful to update information for some time before the last moment at which adjustments can be made. If one does so, systematically misjudging properties such as the target's velocity will only influence performance to the extent that it influences the action during that last time period (Brenner and Smeets, 2015b; de la Malla et al., 2017). However, if the initial error is large, there will be too large an error to correct, so one will not achieve a high precision despite looking at the ball until it is caught.

The studies mentioned earlier demonstrate that people continuously adjust their movements when targets move in a predictable manner: at a constant velocity or accelerated by gravity. When the target is moving erratically or actively trying to avoid being hit, it is obviously even more beneficial to constantly update one's representation of the target's position and movement. The many studies using perturbed targets show that people automatically rapidly adjust to many changes (Aivar et al., 2008; Brenner and Smeets, 2009b; Brenner et al., 1998; Goodale et al., 1986; Prablanc and Martin, 1992; Veerman et al., 2008), so all the underlying judgments must rapidly be updated. When trying to intercept virtual targets that move with various interleaved constant accelerations, people make errors that correspond with the targets' acceleration during the time during which no more adjustments can be made, exactly as one would expect if they continuously adjust their movement on the basis of the target's position and velocity, ignoring the acceleration (Brenner and Smeets, 2015b). People do not make such errors when catching or hitting real balls, because the systematic error disappears almost altogether when a given acceleration is repeated a few times (Brenner et al., 2016). These findings reveal that the interceptive action relies on a constantly updated visual representation of the target's position and velocity to predict where the target will be some time in the future, but also relies on experience during previous trials to use these measures to make the predictions.

When one knows that a ball's motion will change at a certain place, for instance because there is a surface on its path, people quickly divert their gaze to that position (Land and McLeod, 2000), or to where the ball will be slightly after reaching that position (Diaz et al., 2013b), presumably so that the eyes will be on the ball during the critical moments of the action (Mann et al., 2013). Keeping one's eyes on the ball after the bounce is advantageous in terms of updating one' visual representation at the critical moment. It is particularly advantageous if the ball's path can be diverted slightly when it bounces off the surface due to the ball spinning or the surface being uneven (McLeod and Jenkins, 1991). Interception is by no means the only task in which people consistently direct their gaze where critical information for updating their actions is likely to be. Gaze is also consistently directed toward an object that one intends to grasp (Voudouris et al., 2016) or to where one will soon place one's foot when walking over terrain that constrains foot placement (Matthis and Fajen, 2014; Matthis et al., 2015).

Continuously acting on instantaneous updates of the critical visual information is also not specific to interception, but is similarly evident in other movements such as moving one's hand to a static object or reaching to grasp an object. Rapid adjustments to such movements are found when the objects' positions change (Brenner and Smeets, 2015b; Paulignan et al., 1991b), but also when their sizes or orientations change (Brenner and Smeets, 2009b; Castiello et al., 1993; Paulignan et al., 1991a; van Mierlo et al., 2009; Voudouris et al., 2013). Some of these changes even occur when people do not notice that anything has changed (Goodale et al., 1986; Prablanc and Martin, 1992). The latter finding is considered to be surprising, but it is completely consistent with relying on a visual representation in which information is constantly being replaced to adjust one's action. In that case, there is no need notice the change to respond to it (Smeets et al., 2016). Noticing a change or detecting motion might give rise to an additional response, because one might not only make adjustments to reach the new position but actually initially aim even further in the direction of the displacement because seeing the target move might initially make one expect it to continue to move.

Adjustments can also be made when things in the surrounding change. This might sometimes be because the surrounding is considered to be static, so that motion of the surrounding is interpreted as being due to self-motion (Gomi, 2008; Saijo et al., 2005), but it also makes sense to respond to actual displacements of items in the surrounding, especially if they are close to the target (Brenner and Smeets, 2015c), because such items might become, or at least be considered to become, obstacles (Aivar et al., 2008). Rapid adjustments are also found when we acquire new information about our own actions (Brenner and Smeets, 2003; Reichenbach et al., 2014; Saunders and Knill, 2003, 2005).

When a change in the environment calls for a more elaborate response than a simple adjustment to the updated information, such as when an object rotates so that it would be better to grasp it at different positions, the hand initially simply responds to the updated grasping points, but some time later a more suitable response is initiated (Voudouris et al., 2013). This is probably just because the changes in position are updated faster than the changes in orientation, so each digit's endpoint is initially adjusted on the basis of the changes in the positions of the points by which one intended to grasp the object (Smeets and Brenner, 1999), after which the updated orientation leads to a new selection of points for grasping the object. Such a sequence of events is consistent with the idea that different sets of measures are used separately for different aspects of the task and take different amounts of time to update and process.

5 MOVEMENT STRATEGIES

We have been arguing that sensory information is continuously updated and that the updated information is used as soon as it becomes available. Obviously, there are considerable delays involved. Overcoming such delays requires some kind of prediction, both in relation to the motion of external objects and in relation to one's own movements (Scott, 2012; Wolpert and Flanagan, 2001). The prediction must therefore also constantly be updated. The various inputs for the predictions might be updated with different delays, and the resolution of each input may change over time or as a function of gaze. There may be discontinuities in the input when performance on the task dictates that one shifts one's gaze, perhaps because looking elsewhere is advantageous for an upcoming component of the task. There are obviously discontinuities in the updating when the task itself changes. As already mentioned, the information acquired during a fixation (or period of smooth pursuit) takes some time to stabilize, both because measures such as the speed of retinal motion take more time to process than others and because the relative sensitivity to different spatial frequencies changes during the first tens of milliseconds after saccades (Rucci and Victor, 2015). Whenever the input changes, there is the possibility that adjustments to the action will be considered to be necessary. Thus, it is not too surprising that despite continuously controlling movements on the basis of continuously updated visual information, the movements themselves are not always completely smooth (Khan et al., 2006; Lee et al., 1997).

Having to make predictions means that one must incorporate knowledge about the delays. Moreover, despite continuous control based on instantaneous information, there are many aspects of movements that might benefit from considering previous experience. One can do so quite directly, for instance adjusting the timing of one's movements in response to feedback indicating that the circumstances are slightly different than expected. One can also do so on the basis of prolonged experience. For example, although there is much to be said for moving as smoothly as possible (minimizing jerk; Flash and Hogan, 1985) in terms of energy consumption and risk of injury, it may sometimes be advantageous to move differently in order to make better use of one's ability to acquire and use information to guide the task at hand. Finding the best way to move is probably quite a complicated process that requires a lot of practice to optimize.

Movements are not always precisely the same under identical circumstances. This is partly because the movements are constantly guided by the latest noisy sensory estimates, as mentioned earlier, and partly due to noise in the neuromuscular pathway. Additional variability probably arises from considering recent experience and possibly from slightly varying one's movements in order to search for the best strategy. There is ample evidence that details of the way we move are influenced by recent experience, both in terms of the stimuli that were encountered and in terms of feedback about one's performance (Brenner et al., 2013, 2016; de Lussanet et al., 2001; Elliott et al., 2010, 2017). Finding evidence for attempts to find the optimal strategy is much more difficult, but the fact that many actions improve with practice indicates that something must be changing. We know very little about how such changes arise and how they are linked to the appropriate circumstances.

An example of a strategy that does not directly arise from continuous control but that is advantageous if one relies on continuous control is to move the hand asymmetrically, starting quickly in order to rapidly bring it close to the target so that it can then slow down during the final stages of the movement. This is advantageous because seeing the hand move slowly close to the target improves the precision of the sensory signals that guide the hand to the target, such as estimates of the hand's distance from the target and of its position and velocity at the last moment at which adjustments can be made. Moreover, moving slowly near the end of the movement provides more time to make adjustments, while the hand is close to the target. Any further distinction between an early and a late movement phase (Elliott et al., 2010, 2017; Woodworth, 1899) may simply be caused by the changing information about the hand that is acquired during the movement, together with predictions based on experience as to how such information will become available. People must somehow learn that a certain compromise between starting very fast and ending very slowly to maximize precision and performing smooth movements to optimize other factors is advantageous if one wants to achieve a high final spatial precision.

One difficulty in learning how best to move is that different strategies apply under different circumstances. The timing of goal-directed movements is affected by the way in which one moves (Schmidt, 1969; Schmidt et al., 1979). When moving toward a static target, it is advantageous to end one's movement slowly in order to be able to adjust the final path, but when moving toward a target that is moving fast, timing the end of the movement precisely becomes critical. Timing the moment of contact precisely can be achieved by moving fast near the moment of contact (Brenner et al., 2012). There are also conditions in which it is advantageous to move along a curved path (Brenner and Smeets, 2015b). It is therefore not surprising that wanting to end a movement with a high velocity in a direction that is not the main direction of motion, while a long movement time is imposed, can lead to deviations from smooth paths (Lee et al., 1997). This does not refute the idea of continuous control as long as one acknowledges that smoothness of the movement, and therefore probably energy consumption, is not the only consideration. If seeing one's approaching hand is important, but fixating the target is even more important, one should not be surprised to see adjustments made when the hand comes into view or as it comes closer to the target or slows down so that its position can be judged more precisely. With continuous updating, adjustments will usually tend to be gradual, but if one wants to hit a target at an indicated position and realizes that one is arriving too late, one will have to increase the movement speed, so there will be an additional acceleration.

We have mainly discussed continuous control in cases in which the target moved completely predictably, so that gathering more information might increase one's precision but is not essential for performing the task. There are many natural conditions in which continuous control is critical, and would even be so if one would be able to judge positions, velocities, and accelerations very precisely at a glimpse. One might survive being unable to accurately intercept an object that is rolling away across an uneven surface or actively trying to escape. However, our distant ancestors would probably not have survived if they could not adjust their grip when the wind pushed a branch away. We ourselves would probably get into many embarrassing situations if we could not adjust our movements when others hand us objects. Thus, in the real world, it is probably important to continuously update one's judgments and quickly adjust one's actions accordingly, rather than integrating information over long periods of time before planning the perfect action.

Our actions are sometimes guided by what we anticipate rather than directly by what we see. Considering uncertainties such as those mentioned in the previous paragraph, it is useful to consider one's experience when converting the instantaneous visual representations into the anticipations that guide our actions. It might for instance be useful to consider that falling objects generally accelerate due to gravity, even if one does not actually notice the acceleration. More generally, it might be useful to consider the outcome of previous actions in which the circumstances were similar. We know that people adjust their actions to quite arbitrary circumstances, such as unusual target accelerations (Brenner et al., 2016) or artificial delays in the consequences of their own actions (de la Malla et al., 2014). Adjustments to such circumstances need not be adequate for the next action. When they are not, they are likely to introduce errors that need to be compensated for during the movement, as evidence that they are unsuitable emerges. One might even start moving before knowing the precise requirements: the eyes have been seen to move toward the global position at which targets appear before any target has actually appeared (Brenner and Smeets, 2009a) and to start pursuing a moving target before the target has started to move (Kowler et al., 2014). The hand has been seen to start to move before one knows precisely where it will be moving to (Brenner and Smeets, 2015a; Hening et al., 1988; van Sonderen and Denier van der Gon, 1991). Thus, in order to draw conclusions from a pattern of movement, one must be certain that one understands the history and the precise constraints. We argue that the large variability in how movements are made, across individuals but especially across similar actions by the same individual, can be understood in terms of people relying on instantaneous visual information together with various more abstract aspects of their recent experience, such as that they should aim less far ahead of a moving target (Brenner et al., 2013) or further to the left than on the previous trial (van Beers, 2009).

6 CONCLUSION

We propose that the visual representations in the brain are continuously updated as new information enters our eyes. The way we direct our gaze largely determines the content of these representations. The task determines how we direct our gaze in order to ensure that the representations contain relevant information. The task probably also determines what information is actually used and what is retained in a more enduring representation. The information that is retained is probably very limited, consisting of some basic information about objects and their relative positions, information about what information is relevant for the task and how to respond to it, and any strategies that might improve performance when one encounters the same or a similar task in the future. Such information is combined with continuous control on the basis of the information that is available at each moment to guide our actions.

REFERENCES

Aivar, M.P., Hayhoe, M.M., Chizk, C.L., Mruczek, R.E., 2005. Spatial memory and saccadic targeting in a natural task. J. Vis. 5 (3), 177–193.

Aivar, M.P., Brenner, E., Smeets, J.B.J., 2008. Avoiding moving obstacles. Exp. Brain Res. 190, 251–264.

Ballard, D.H., Hayhoe, M.M., Li, F., Whitehead, S.D., 1992. Hand-eye coordination during sequential tasks. Philos. Trans. R. Soc. Lond. B Biol. Sci. 337 (1281), 331–338.

Ballard, D.H., Hayhoe, M.M., Pelz, J.B., 1995. Memory representations in natural tasks. J. Cogn. Neurosci. 7 (1), 66–80.

Becker, S.I., 2011. Determinants of dwell time in visual search: similarity or perceptual difficulty? PLoS One 6 (3), e17740.

Becker, S.I., Lewis, A.J., 2015. Oculomotor capture by irrelevant onsets with and without color contrast. Ann. N. Y. Acad. Sci. 1339, 60–71.

Boi, M., Poletti, M., Victor, J.D., Rucci, M., 2017. Consequences of the oculomotor cycle for the dynamics of perception. Curr. Biol. 27 (9), 1268–1277.

Brenner, E., Cornelissen, F.W., 2000. Separate simultaneous processing of egocentric and relative positions. Vision Res. 40, 2557–2563.

Brenner, E., Smeets, J.B.J., 2000. Motion extrapolation is not responsible for the flash-lag effect. Vision Res. 40, 1645–1648.

Brenner, E., Smeets, J.B.J., 2003. Fast corrections of movements with a computer mouse. Spat. Vis. 16, 365–376.

Brenner, E., Smeets, J.B.J., 2007. Flexibility in intercepting moving objects. J. Vis. 7, 14.

Brenner, E., Smeets, J.B.J., 2009a. Sources of variability in interceptive movements. Exp. Brain Res. 195, 117–133.

Brenner, E., Smeets, J.B.J., 2009b. Modifying one's hand's trajectory when a moving target's orientation changes. Exp. Brain Res. 196, 375–383.

Brenner, E., Smeets, J.B.J., 2010. How well can people judge when something happened? Vision Res. 50, 1101–1108.

Brenner, E., Smeets, J.B.J., 2011. Continuous visual control of interception. Hum. Mov. Sci. 30, 475–494.

Brenner, E., Smeets, J.B.J., 2015a. Quickly making the correct choice. Vision Res. 113, 198–210.

Brenner, E., Smeets, J.B.J., 2015b. How people achieve their amazing temporal precision in interception. J. Vis. 15 (3), 8.

Brenner, E., Smeets, J.B.J., 2015c. How moving backgrounds influence interception. PLoS One 10 (3), e0119903.

Brenner, E., van Damme, W.J., 1999. Perceived distance, shape and size. Vision Res. 39, 975–986.

Brenner, E., Smeets, J.B.J., de Lussanet, M.H.E., 1998. Hitting moving targets—continuous control of the acceleration of the hand on the basis of the target's velocity. Exp. Brain Res. 122 (4), 467–474.

Brenner, E., Smeets, J.B.J., van den Berg, A.V., 2001. Smooth eye movements and spatial localisation. Vision Res. 41, 2253–2259.

Brenner, E., van Dam, M., Berkhout, S., Smeets, J.B.J., 2012. Timing the moment of impact in fast human movements. Acta Psychol. (Amst) 141, 104–111.

Brenner, E., Canal-Bruland, R., van Beers, R.J., 2013. How the required precision influences the way we intercept a moving object. Exp. Brain Res. 230, 207–218.

Brenner, E., Driesen, B., Smeets, J.B.J., 2014. Precise timing when hitting falling balls. Front. Hum. Neurosci. 8, 342.

Brenner, E., Abalo, I., Estal, V., Schootemeijer, S., Mahieu, Y., Veerkamp, K., Zandbergen, M., van der Zee, T., Smeets, J.B.J., 2016. How can people be so good at intercepting accelerating objects if they are so poor at visually judging acceleration? Iperception 7, 2041669515624317.

Brouwer, A.M., Knill, D.C., 2007. The role of memory in visually guided reaching. J. Vis. 7 (5), 6.1–6.12.

Brouwer, A.M., Knill, D.C., 2009. Humans use visual and remembered information about object location to plan pointing movements. J. Vis. 9 (1), 24.1–24.19.

Castiello, U., Bennett, K.M., Stelmach, G.E., 1993. Reach to grasp: the natural response to perturbation of object size. Exp. Brain Res. 94 (1), 163–178.

Clemens, I.A.H., Selen, L.P.J., Koppen, M., Medendorp, W.P., 2012. Visual stability across combined eye and body motion. J. Vis. 12 (12), 8, 1–11.

Collins, T., Rolfs, M., Deubel, H., Cavanagh, P., 2009. Post-saccadic location judgments reveal remapping of saccade targets to non-foveal locations. J. Vis. 9 (5), 29.1–29.9.

de Grave, D.D., Brenner, E., Smeets, J.B.J., 2004. Illusions as a tool to study the coding of pointing movements. Exp. Brain Res. 155, 56–62.

de la Malla, C., López-Moliner, J., Brenner, E., 2014. Dealing with delays does not transfer across sensorimotor tasks. J. Vis. 14 (12), 8.

de la Malla, C., Smeets, J.B.J., Brenner, E., 2017. Errors in interception can be predicted from errors in perception. Cortex. pii: S0010-9452(17)30086-2. https://doi.org/10.1016/j.cortex.2017.03.006.

de Lussanet, M.H., Smeets, J.B.J., Brenner, E., 2001. The effect of expectations on hitting moving targets: influence of the preceding target's speed. Exp. Brain Res. 137, 246–248.

Deubel, H., Bridgeman, B., Schneider, W.X., 1998. Immediate post-saccadic information mediates space constancy. Vision Res. 38 (20), 3147–3159.

Deubel, H., Schneider, W.X., Bridgeman, B., 2002. Transsaccadic memory of position and form. Prog. Brain Res. 140, 165–180.

Deubel, H., Koch, C., Bridgeman, B., 2010. Landmarks facilitate visual space constancy across saccades and during fixation. Vision Res. 50 (2), 249–259.

Diaz, G., Cooper, J., Hayhoe, M., 2013a. Memory and prediction in natural gaze control. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368 (1628)20130064.

Diaz, G., Cooper, J., Rothkopf, C., Hayhoe, M., 2013b. Saccades to future ball location reveal memory-based prediction in a virtual-reality interception task. J. Vis. 13 (1), 1–14, pii: 20.

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

Elliott, D., Hansen, S., Grierson, L.E., Lyons, J., Bennett, S.J., Hayes, S.J., 2010. Goal-directed aiming: two components but multiple processes. Psychol. Bull. 136 (6), 1023–1044.

Elliott, D., Lyons, J., Hayes, S.J., Burkitt, J.J., Roberts, J.W., Grierson, L.E., Hansen, S., Bennett, S.J., 2017. The multiple process model of goal-directed reaching revisited. Neurosci. Biobehav. Rev. 72, 95–110.

Fattori, P., Breveglieri, R., Amoroso, K., Galletti, C., 2004. Evidence for both reaching and grasping activity in the medial parieto-occipital cortex of the macaque. Eur. J. Neurosci. 20 (9), 2457–2466.

Fattori, P., Breveglieri, R., Marzocchi, N., Filippini, D., Bosco, A., Galletti, C., 2009. Hand orientation during reach-to-grasp movements modulates neuronal activity in the medial posterior parietal area V6A. J. Neurosci. 29 (6), 1928–1936.

Fiehler, K., Wolf, C., Klinghammer, M., Blohm, G., 2014. Integration of egocentric and allocentric information during memory-guided reaching to images of a natural environment. Front. Hum. Neurosci. 8, 636.

Flash, T., Hogan, N., 1985. The coordination of arm movements: an experimentally confirmed mathematical model. J. Neurosci. 5 (7), 1688–1703.

Gillam, B., Chambers, D., 1985. Size and position are incongruous: measurements on the Müller-Lyer figure. Percept. Psychophys. 37, 549–556.

Gomi, H., 2008. Implicit online corrections of reaching movements. Curr. Opin. Neurobiol. 18, 558–564.

Gonzalez, C.C., Causer, J., Miall, R.C., Grey, M.J., Humphreys, G., Williams, A.M., 2017. Identifying the causal mechanisms of the quiet eye. Eur. J. Sport Sci. 17 (1), 74–84.

Goodale, M.A., Pelisson, D., Prablanc, C., 1986. Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. Nature 320 (6064), 748–750.

Graziano, M.S., 2001. Is reaching eye-centered, body-centered, hand-centered, or a combination? Rev. Neurosci. 12 (2), 175–185.

Hayhoe, M., Aivar, P., Shrivastavah, A., Mruczek, R., 2002. Visual short-term memory and motor planning. Prog. Brain Res. 140, 349–363.

Hayhoe, M.M., Shrivastava, A., Mruczek, R., Pelz, J.B., 2003. Visual memory and motor planning in a natural task. J. Vis. 3 (1), 49–63.

He, D., Mo, C., Fang, F., 2017. Predictive feature remapping before saccadic eye movements. J. Vis. 17 (5), 14.

Hening, W., Favilla, M., Ghez, C., 1988. Trajectory control in targeted force impulses. V. Gradual specification of response amplitude. Exp. Brain Res. 71 (1), 116–128.

Henriques, D.Y.P., Klier, E.M., Smith, M.A., Lowy, D., Crawford, J.D., 1998. Gaze-centered remapping of remembered visual space in an open-loop pointing task. J. Neurosci. 18, 1583–1594.

Herwig, A., 2015. Transsaccadic integration and perceptual continuity. J. Vis. 15 (16), 7.

Honda, H., 1993. Saccade-contingent displacement of the apparent position of visual stimuli flashed on a dimly illuminated structured background. Vision Res. 33, 709–716.

Hooge, I.T., Erkelens, C.J., 1998. Adjustment of fixation duration in visual search. Vision Res. 38 (9), 1295–1302.

Horowitz, T.S., Wolfe, J.M., 1998. Visual search has no memory. Nature 394 (6693), 575–577.

Irwin, D.E., Zacks, J.L., Brown, J.S., 1990. Visual memory and the perception of a stable visual environment. Percept. Psychophys. 47 (1), 35–46.

Khan, M.A., Franks, I.M., Elliott, D., Lawrence, G.P., Chua, R., Bernier, P.M., Hansen, S., Weeks, D.J., 2006. Inferring online and offline processing of visual feedback in target-directed movements from kinematic data. Neurosci. Biobehav. Rev. 30 (8), 1106–1121.

Klinghammer, M., Schutz, I., Blohm, G., Fiehler, K., 2016. Allocentric information is used for memory-guided reaching in depth: a virtual reality study. Vision Res. 129, 13–24.

Kowler, E., Aitkin, C.D., Ross, N.M., Santos, E.M., Zhao, M., 2014. Davida Teller Award Lecture 2013: the importance of prediction and anticipation in the control of smooth pursuit eye movements. J. Vis. 14 (5), 10.

Kusunoki, M., Goldberg, M.E., 2003. The time course of perisaccadic receptive field shifts in the lateral intraparietal area of the monkey. J. Neurophysiol. 89, 1519–1527.

Lacquaniti, F., Caminiti, R., 1998. Visuo-motor transformations for arm reaching. Eur. J. Neurosci. 10 (1), 195–203.

Land, M.F., McLeod, P., 2000. From eye movements to actions: how batsmen hit the ball. Nat. Neurosci. 3, 1340–1345.

Land, M., Mennie, N., Rusted, J., 1999. The roles of vision and eye movements in the control of activities of daily living. Perception 28 (11), 1311–1328.

Lebeau, J.C., Liu, S., Sáenz-Moncaleano, C., Sanduvete-Chaves, S., Chacón-Moscoso, S., Becker, B.J., Tenenbaum, G., 2016. Quiet eye and performance in sport: a meta-analysis. J. Sport Exerc. Psychol. 38 (5), 441–457.

Lee, D., Port, N.L., Georgopoulos, A.P., 1997. Manual interception of moving targets. II. On-line control of overlapping submovements. Exp. Brain Res. 116 (3), 421–433.

Levin, D.T., Simons, D.J., 1997. Failure to detect changes to attended objects in motion pictures. Psychon. Bull. Rev. 4, 501–506.

López-Moliner, J., Brenner, E., 2016. Flexible timing of eye movements when catching a ball. J. Vis. 16 (5), 13.

Mann, D.L., Spratford, W., Abernethy, B., 2013. The head tracks and gaze predicts: how the world's best batters hit a ball. PLoS One 8 (3), e58289.

Mateeff, S., 1978. Saccadic eye movements and localization of visual stimuli. Percept. Psychophys. 24, 215–224.

Mathôt, S., Theeuwes, J., 2010. Evidence for the predictive remapping of visual attention. Exp. Brain Res. 200 (1), 117–122.

Matin, L., Matin, E., Pola, J., 1970. Visual perception of direction when voluntary saccades occur. II. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented before the saccade. Percept. Psychophys. 8, 9–14.

Matin, E., Clymer, A.B., Matin, L., 1972. Metacontrast and saccadic suppression. Science 178 (4057), 179–182.

Matthis, J.S., Fajen, B.R., 2014. Visual control of foot placement when walking over complex terrain. J. Exp. Psychol. Hum. Percept. Perform. 40 (1), 106–115.

Matthis, J.S., Barton, S.L., Fajen, B.R., 2015. The biomechanics of walking shape the use of visual information during locomotion over complex terrain. J. Vis 15 (3), 10, 1–13, pii: 10.

McLeod, P., Jenkins, S., 1991. Timing accuracy and decision time in high-speed ball games. Int. J. Sport Psychol. 22, 279–295.

Medendorp, W.P., Van Asselt, S., Gielen, C.C.A.M., 1999. Pointing to remembered visual targets after active one-step self-displacements within reaching space. Exp. Brain Res. 125, 50–60.

Melcher, D., 2007. Predictive remapping of visual features precedes saccadic eye movements. Nat. Neurosci. 10 (7), 903–907.

Mennie, N., Hayhoe, M., Sullivan, B., 2007. Look-ahead fixations: anticipatory eye movements in natural tasks. Exp. Brain Res. 179 (3), 427–442.

Niemeier, M., Crawford, J.D., Tweed, D.B., 2003. Optimal transsaccadic integration explains distorted spatial perception. Nature 422 (6927), 76–80.

Nijhawan, R., 1994. Motion extrapolation in catching. Nature 370, 256–257.

Oostwoud Wijdenes, L., Marshall, L., Bays, P.M., 2015. Evidence for optimal integration of visual feature representations across saccades. J. Neurosci. 35 (28), 10146–10153.

O'Regan, J.K., Noë, A., 2001. A sensorimotor account of vision and visual consciousness. Behav. Brain Sci. 24 (5), 939–973.

O'Regan, J.K., Rensink, R.A., Clark, J.J., 1999. Change-blindness as a result of 'mudsplashes'. Nature 398 (6722), 34.

Paulignan, Y., Jeannerod, M., MacKenzie, C., Marteniuk, R., 1991a. Selective perturbation of visual input during prehension movements. 2. The effects of changing object size. Exp. Brain Res. 87 (2), 407–420.

Paulignan, Y., MacKenzie, C., Marteniuk, R., Jeannerod, M., 1991b. Selective perturbation of visual input during prehension movements. 1. The effects of changing object position. Exp. Brain Res. 83 (3), 502–512.

Peters, R.J., Iyer, A., Itti, L., Koch, C., 2005. Components of bottom-up gaze allocation in natural images. Vision Res. 45 (18), 2397–2416.

Peterson, M.S., Kramer, A.F., Wang, R.F., Irwin, D.E., McCarley, J.S., 2001. Visual search has memory. Psychol. Sci. 12, 287–292.

Prablanc, C., Martin, O., 1992. Automatic control during hand reaching at undetected two-dimensional target displacements. J. Neurophysiol. 67 (2), 455–469.

Reichenbach, A., Franklin, D.W., Zatka-Haas, P., Diedrichsen, J., 2014. A dedicated binding mechanism for the visual control of movement. Curr. Biol. 24 (7), 780–785.

Rensink, R.A., 2002. Change detection. Annu. Rev. Psychol. 53, 245–277.

Rolfs, M., Jonikaitis, D., Deubel, H., Cavanagh, P., 2011. Predictive remapping of attention across eye movements. Nat. Neurosci. 14 (2), 252–256.

Rotman, G., Brenner, E., Smeets, J.B.J., 2004. Mislocalization of targets flashed during smooth pursuit depends on the change in gaze direction after the flash. J. Vis. 4, 564–574.

Rucci, M., Victor, J.D., 2015. The unsteady eye: an information-processing stage, not a bug. Trends Neurosci. 38 (4), 195–206.

Rucci, M., Iovin, R., Poletti, M., Santini, F., 2007. Miniature eye movements enhance fine spatial detail. Nature 447 (7146), 851–854.

Ryou, J.W., Wilson, F.A., 2004. Making your next move: dorsolateral prefrontal cortex and planning a sequence of actions in freely moving monkeys. Cogn. Affect. Behav. Neurosci. 4 (4), 430–443.

Saijo, N., Murakami, I., Nishida, S., Gomi, H., 2005. Large-field visual motion directly induces an involuntary rapid manual following response. J. Neurosci. 25 (20), 4941–4951.

Saunders, J.A., Knill, D.C., 2003. Humans use continuous visual feedback from the hand to control fast reaching movements. Exp. Brain Res. 152, 341–352. PMID: 12904935.

Saunders, J.A., Knill, D.C., 2005. Humans use continuous visual feedback from the hand to control both the direction and distance of pointing movements. Exp. Brain Res. 162, 458–473. PMID: 15754182.

Schlag, J., Schlag-Rey, M., 2002. Through the eye, slowly: delays and localization errors in the visual system. Nat. Rev. Neurosci. 3, 191–200.

Schmidt, R.A., 1969. Movement time as a determiner of timing accuracy. J. Exp. Psychol. 79, 43–47.

Schmidt, R.A., Zelaznik, H., Hawkins, B., Frank, J.S., Quinn, J.T., 1979. Motor-output variability: a theory for the accuracy of rapid motor acts. Psychol. Rev. 47, 415–451.

Schmolesky, M.T., Wang, Y.C., Hanes, D.P., Thompson, K.G., Leutgeb, S., Schall, J.D., Leventhal, A.G., 1998. Signal timing across the macaque visual system. J. Neurophysiol. 79 (6), 3272–3278.

Schutz, I., Henriques, D.Y.P., Fiehler, K., 2013. Gaze-centered spatial updating in delayed reaching even in the presence of landmarks. Vision Res. 87, 46–52.

Scott, S.H., 2012. The computational and neural basis of voluntary motor control and planning. Trends Cogn. Sci. 16 (11), 541–549.

Simons, D.J., Chabris, C.F., 1999. Gorillas in our midst: sustained inattentional blindness for dynamic events. Perception 28 (9), 1059–1074.

Simons, D.J., Levin, D.T., 1998. Failure to detect changes to people during a real-world interaction. Psychon. Bull. Rev. 5, 644–649.

Simons, D.J., Rensink, R.A., 2005. Change blindness: past, present, and future. Trends Cogn. Sci. 9 (1), 16–20.

Smeets, J.B.J., Brenner, E., 1995. Perception and action are based on the same visual information: distinction between position and velocity. J. Exp. Psychol. Hum. Percept. Perform. 21, 19–31.

Smeets, J.B.J., Brenner, E., 1999. A new view on grasping. Motor Control 3, 237–271.

Smeets, J.B.J., Brenner, E., 2008. Why we don't mind to be inconsistent. In: Calvo, P., Gomila, T. (Eds.), Handbook of Cognitive Science—An Embodied Approach. Elsevier, San Diego, Oxford, Amsterdam, pp. 207–217.

Smeets, J.B.J., Brenner, E., de Grave, D.D., Cuijpers, R.H., 2002. Illusions in action: consequences of inconsistent processing of spatial attributes. Exp. Brain Res. 147, 135–144.

Smeets, J.B.J., van den Dobbelsteen, J.J., de Grave, D.D.J., van Beers, R.J., Brenner, E., 2006. Sensory integration does not lead to sensory calibration. Proc. Natl. Acad. Sci. U.S.A. 103 (49), 18781–18786.

Smeets, J.B.J., Oostwoud Wijdenes, L., Brenner, E., 2016. Movement adjustments have short latencies because there is no need to detect anything. Motor Control 20, 137–148.

Soechting, J.F., Juveli, J.Z., Rao, H.M., 2009. Models for the extrapolation of target motion for manual interception. J. Neurophysiol. 102 (3), 1491–1502.

Tatler, B.W., 2007. The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. J. Vis. 7 (14), 4.1–17.

Thompson, A.A., Henriques, D.Y., 2011. The coding and updating of visuospatial memory for goal-directed reaching and pointing. Vision Res. 51 (8), 819–826.

Triesch, J., Ballard, D.H., Hayhoe, M.M., Sullivan, B.T., 2003. What you see is what you need. J. Vis. 3 (1), 86–94.

van Beers, R.J., 2009. Motor learning is optimally tuned to the properties of motor noise. Neuron 63, 406–417.

van de Grind, W.A., Koenderink, J.J., van Doorn, A.J., 1986. The distribution of human motion detector properties in the monocular visual field. Vision Res. 26 (5), 797–810.

van der Stigchel, S., Belopolsky, A.V., Peters, J.C., Wijnen, J.G., Meeter, M., Theeuwes, J., 2009. The limits of top-down control of visual attention. Acta Psychol. (Amst) 132 (3), 201–212.

van Doorn, A.J., Koenderink, J.J., 1982. Temporal properties of the visual detectability of moving spatial white noise. Exp. Brain Res. 45 (1–2), 179–188.

van Mierlo, C.M., Louw, S., Smeets, J.B.J., Brenner, E., 2009. Slant cues are processed with different latencies for the online control of movement. J. Vis. 9 (3), 25.1–25.8.

van Pelt, S., Medendorp, W.P., 2007. Gaze-centered updating of remembered visual space during active whole-body translations. J. Neurophysiol. 97 (2), 1209–1220.

van Sonderen, J.F., Denier van der Gon, J.J., 1991. Reaction-time-dependent differences in the initial movement direction of fast goal-directed arm movements. Hum. Mov. Sci. 10 (6), 713–726.

Veerman, M.M., Brenner, E., Smeets, J.B.J., 2008. The latency for correcting a movement depends on the visual attribute that defines the target. Exp. Brain Res. 187 (2), 219–228.

Võ, M.L., Wolfe, J.M., 2012. When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. J. Exp. Psychol. Hum. Percept. Perform. 38 (1), 23–41.

Voudouris, D., Smeets, J.B.J., Brenner, E., 2013. Ultra-fast selection of grasping points. J. Neurophysiol. 110, 1484–1489.

Voudouris, D., Smeets, J.B.J., Brenner, E., 2016. Fixation biases towards the index finger in almost-natural grasping. PLoS One 11 (1), e0146864.

Walker, M.F., Fitzgibbon, E.J., Goldberg, M.E., 1995. Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. J. Neurophysiol. 73, 1988–2003.

Wolpert, D.M., Flanagan, J.R., 2001. Motor prediction. Curr. Biol. 11 (18), R729–R732.

Woodworth, R.S., 1899. The accuracy of voluntary movement. Psychol. Rev. 3, 1–119. Monograph Supplement.