



Multisensory integration in visuospatial recalibration

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Manuel Bayer
aus Düsseldorf

Düsseldorf, Oktober 2023

aus dem Institut für Experimentelle Psychologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Eckart Zimmermann
2. Prof. Dr. Christian Bellebaum

Tag der mündlichen Prüfung: 04.12.23

Contents

Summary	1
Introduction	2
The Perceptual system	2
Embodied depth perception	5
Perception of a stable world	10
Overview of studies	17
Study 1	17
Study 2	23
Study 3	29
General Discussion	34
Integration of bodily information in visual perception	35
The role of experience in visual perception	37
The influence of gaze shifts on spatial representations	41
Conclusion	45
References	47
Affidavit	61
Acknowledgements	62
Appendix	63

Summary

The information the visual system provides is essential for our interaction with the environment. To what extent the perceptual system uses and relies on body-related signals and sensorimotor experiences is a topic for discussion. While early theories regarded the visual system as a mere input system, recent findings suggest a more embodied perspective. The studies presented in this dissertation examined the influence of body-related and sensorimotor information and experience on visual perception. The first study investigated how the perceived self-location in space alters our perception of depth. It provides evidence for a use of body-related information to interpret ambiguous signals and to provide accurate distance estimations. The second study examined visual stability, our ability to perceive a stable world despite the constant movements we perform. The results suggest that constant recalibrations are performed to adjust our expectations of the visual consequences of performed movements based on previous sensorimotor experience. The third study aimed to explore space constancy, which refers to our ability to keep an accurate spatial representation of surrounding objects. The data suggest that we use all information sources available to use in order to create the most accurate and actionable spatial representation of objects. This dissertation delivers further support for an embodied account of visual perception by providing evidence for the inclusion of body-related information and sensorimotor experience in visual processes.

Introduction

The Perceptual system

Visual perception

Human visual perception is restricted by a visual field of 180 ° horizontally and 120 ° vertically and relies on light falling into the eye (Johnson et al., 2011). Before being projected onto the retina, the light passes through the cornea, lens and pupil, which adjusts to the amount of light passing through it by varying its diameter between 2 and 9 mm (Laeng & Endestad, 2012). Photoreceptors in the retina enable the conversion of light into an electrical signal that is transmitted by the nervus opticus over the chiasma opticum and the tractus opticus to the thalamus, in particular the nucleus geniculatum laterale, the pulvinar and the colliculus superior. The nucleus geniculatum laterale then transmits the signal to the primary visual cortex via the radiatio optica. The retina can perceive light with a wavelength of approximately 350 nm to 750 nm (Wald, 1945) and has varying density levels of photoreceptors across its surface. The highest density of photoreceptors can be found on the fovea, the point of our highest visual acuity. The punctum caecum on the other hand is the point that is unable to convert light into an electrical signal due to the lack of photoreceptors as the nervus opticus passes through the retina at this spot. The primary visual cortex is responsible for the initial processing of the visual input. It is structured in columns that cover selective parts of the visual field and are sensitive to different properties of the visual input, e.g., a specific orientation (Hubel & Wiesel, 1962). Higher order processing of the visual input takes place in other areas of the visual cortex, e.g., the secondary and tertiary visual cortex.

Proprioception

The visual system provides valuable information about our surroundings and can also be used to get an accurate representation of our body in space. When we use the visual system to perceive and process objects in our environment, we cannot use it simultaneously to obtain information about our position in space. We are therefore dependent on other systems, like the proprioceptive and vestibular system, to provide us with the necessary information. *Proprioception* is not a uniformly defined construct (Hillier et al., 2015). Here, we define proprioception as a bundling of several perceptual abilities, including the perception of joint position, force, velocity, acceleration, muscular effort and heaviness and kinesthesia, the sense of the movement of our limbs (Niessen et al., 2009; Ogard, 2011; Proske, 2006; Proske & Gandevia, 2012). Similar to the specialized receptors of the retina, proprioception utilizes mechanosensory neurons, also termed *proprioceptors*, which cover most of our body (Tuthill & Azim, 2018). Signals of these proprioceptors are transmitted to spinal ganglions, which forward the signal to nuclei in the medulla spinalis. There are two superordinate pathways to the cerebellum and to the somatosensory cortex, both starting at the medulla spinalis. On the cerebellar pathway, proprioceptive signals are processed in cerebellar nuclei like the pedunculus cerebellaris superior or inferior. On the pathway to the somatosensory cortex, they pass the ventral posterolateral nucleus in the thalamus before being further processed in the primary and secondary somatosensory cortex.

The vestibular system

The vestibular system contributes to maintaining balance of the body and to keeping an accurate spatial orientation, especially in the context of

movement (Khan & Chang, 2013). In order to accomplish these tasks, the vestibular system provides information about head motion and gravitational forces (Khan & Chang, 2013). This information is gathered by five structures of the inner ear, which include the lateral, superior, and posterior semicircular canals, the saccule, and the utricle (Khan & Chang, 2013). The semicircular canals are sensitive to angular acceleration, e.g., rotation of the head. Due to their position in relation to the spatial planes, they are sensitive to angular acceleration in the sagittal and axial plane, covering head pitch and yaw (Lee, 2011). Once they experience angular acceleration, the cupula, a gelatinous substance, is displaced. Hair cells, the receptors of the vestibular system, are located in the cupula and are depolarized during this displacement. This causes the transmission of an electrical signal to the vestibular ganglion (Khan & Chang, 2013). From the vestibular ganglion, the signal is transmitted via the vestibular nerve, which later merges with the auditory nerve to form the vestibulocochlear nerve. The vestibular signals end up in the brainstem, pons, and the cerebellum. The saccule and the utricle provide information about the orientation of the head, linear acceleration and gravitational forces (Barrett et al., 2012; Khan & Chang, 2013). Both contain a macula, a sensory neuroepithelium, which is able to perceive horizontal plane motion in the case of the utricle and vertical plane motion in the case of the saccule (Khan & Chang, 2013). The hair cells in the saccule and the utricle are more sensitive than the ones found in the semicircular canals and can therefore additionally also perceive gravitational forces, linear movement and tilting of the head (Khan & Chang, 2013).

Embodied depth perception

Depth perception

In order to grasp or manipulate objects in our environment, we rely on an accurate spatial representation of these objects. While the visual system provides us with the input we need to create such a spatial map of our surroundings, the input has to be interpreted in order to be actionable. This interpretation relies on monocular cues, e.g., accommodation and motion parallax, and binocular cues, e.g., convergence and stereopsis (Servos et al., 1992). *Accommodation* is the mechanism by which the eye is able to focus distant or close objects and is accomplished by changing the curvature of the lens via the contraction and relaxation of the ciliary eye muscles, respectively. *Motion parallax* on the other hand is a depth cue obtained from moving one's head while observing the perceived motion of objects at varying distances. Objects beyond the fixation distance are perceived to move with the self-produced motion while closer objects move against it, giving the observer information about the distance of objects relative to one another and the fixation distance (Ferris, 1972). Estimating the distance to an object based on accommodation of the eye or motion parallax is inaccurate and unreliable (Ferris, 1972; Fisher & Ciuffreda, 1988). *Convergence* on the other hand, can be used as a reliable and accurate source of depth information but is hard to explicitly access (Hill, 1972; Servos et al., 1992). Convergence refers to the angle that the center line of sight of both eyes have to one another. This angle becomes larger with increasing distance to the fixation point and is smaller the closer this point is. *Stereopsis* is the ability to interpret two images taken from different angles as one. While the binocular cues seem to be of greater importance (Nicolle et al., 1995; Servos et al., 1992), it is likely

that our perceptual system is taking all available information into account to represent the environment as accurately as possible with regard to depth.

Multisensory integration

While there are overlaps regarding what the senses can perceive, there is usually always one sense specialized to perceive the respective input. We can for example see where our limbs are in relation to us, but we can also perceive their relative location based on our proprioception. In most cases in which our visual system can be used to perceive the input of interest, it is the most accurate of our perceptual systems. Despite the possibility to simply perceive the respective input with the most specialized sense, we integrate the information of several senses in order to obtain the most accurate and in-depth representation of the input we can perceive. The underlying mechanism is called *multisensory integration* and its result is beyond a mere sum of the input of the individual senses (Stein & Stanford, 2008). Multisensory integration manipulates the weight assigned to the different modality specific inputs (Calvert et al., 2004). By weighting the inputs of the different senses, the consideration of the information provided by one sense can be enhanced. There is evidence for situation dependent adjustments of these processes, e.g., via attention (Macaluso et al., 2016), further suggesting that multisensory integration goes far beyond simply adding the individual inputs up. One thoroughly researched aspect of multisensory integration is the integration of conflicting information from different senses (Gelder & Bertelson, 2003). Conflicting in this context means that different senses receive input that is contradictory to one another, forcing a decision regarding which of the inputs to rely on. In such situations, the information of the most accurate sense is weighted more and biases the perceived input of contradictory senses in the

direction of the most accurate one (Ernst & Banks, 2002).

Full-body illusion

The full-body illusion is an illusion that exploits the mechanisms of multisensory integration (Blanke et al., 2015; Hohwy & Paton, 2010; Pfeiffer et al., 2014). It can be induced, e.g., by presenting a video showing the application of tactile stimulation on the participants' back while the participant experiences a congruent tactile stimulation (Lenggenhager et al., 2007). Other studies used avatars in virtual environments instead of videos of the participant (Lenggenhager et al., 2011). Since vision is the most accurate sense for the inputs during the induction of the full-body illusion, multisensory integration biases the sense of touch and proprioception towards vision (Ehrsson, 2007; Lenggenhager et al., 2007). This visual bias leads to an identification with and a perceived self-location shift towards the avatar (Aspell et al., 2009; Ionta et al., 2011; Lenggenhager et al., 2011; Lenggenhager et al., 2007; Palluel et al., 2011). Several methods were developed to measure the effect of the full-body illusion, like questionnaires (Salomon et al., 2013) and tasks which aim to determine the perceived self-location in space of the participants (Aspell et al., 2009; Lenggenhager et al., 2011; Lenggenhager et al., 2009; Lenggenhager et al., 2007; Nakul et al., 2020). Blind walking is an example for such a task (Aspell et al., 2009; Lenggenhager et al., 2011; Lenggenhager et al., 2007). In order to quantify the self-location via blind walking, participants are displaced from their initial position and are then instructed to move back to their previous position. Under the influence of the full-body illusion, participants tend to overshoot their position and move closer to the position of the avatar (Aspell et al., 2009; Lenggenhager et al., 2011; Lenggenhager et al., 2007). Other tasks that aim to determine the

perceived self-location involve the interaction with a ball, e.g., by dropping a ball and indicating when the ball would reach the ground (Lenggenhager et al., 2009) or by indicating when a rolling ball would reach the perceived self-location position of the participant (Nakul et al., 2020). There is also an approach which uses looming sounds to determine the boundaries of the peripersonal space (Noel et al., 2015), a space that is usually defined as the space around the body limited by the grasping distance (Longo & Lourenco, 2007; Rizzolatti et al., 1981; Wiesing et al., 2021). In addition to these behavioral results and effects, imaging studies were able to provide support for full-body illusion related activity in structures like the bilateral premotor cortex, intraparietal sulcus and sensorimotor cortex (Blanke et al., 2015; Ionta et al., 2011; Lenggenhager et al., 2011).

Embodied cognition

Embodied cognition represents a perspective on cognition which involves the sensorimotor system as more than a mere input-output system (Wilson, 2002). In the context of the underlying family of theories, information provided by and experiences gained with the body are taken into account. There is a growing amount of evidence for this family of theories with supporting studies coming from manifold fields of research, including but not limited to semantic processing (Bechtold et al., 2019; Gallese & Lakoff, 2005; Glenberg & Kaschak, 2003), visual guidance (Warren, 1984), peripersonal space (Longo & Lourenco, 2007) and depth perception (Wiesing et al., 2021). Especially the involvement of body-related information in perceptual processing represents a drastic shift in perspective. In prior accounts, the perceptual system was assumed to be a mere input system, which would consequently only provide us with objective unaltered perceptual input (Pylyshyn, 2003; Wilson,

2002). Recent research suggests a different perspective, indicating that our perception is able to adjust to our current situation (Proffitt, 2006; Proffitt et al., 1995). Moreover, it also takes the experiences gained with (Wiesing et al., 2021) and the information provided by our sensorimotor system into account (Longo & Lourenco, 2007; Proffitt, 2006).

In the context of depth perception, the consideration of information beyond the monocular and binocular cues is essential in order to interpret the signals correctly and to obtain an accurate distance estimation. Several studies provide compelling evidence that sensorimotor and body-related information is factored in during this process (Cont & Zimmermann, 2021; Longo & Lourenco, 2007; Ogawa et al., 2018; van der Hoort et al., 2011; Volcic et al., 2013; Wiesing et al., 2021). The size of our limbs or our grasping distance (Longo & Lourenco, 2007; Volcic et al., 2013; Wiesing et al., 2021), the size of our body in relation to objects (Ogawa et al., 2018; van der Hoort et al., 2011) and information of planned movements (Cont & Zimmermann, 2021) are all information that is used to further narrow down the interpretation of ambiguous cues and to get a more accurate estimate of the distance to an object. There is even evidence that outside of the peripersonal space, our perception takes information about the state of our body into account when judging distances, e.g., our physical fatigue (Proffitt, 2006; Warren, 1984). Depth perception should therefore be regarded as an inherently embodied process that not only uses non-visual information sources but relies on the contributions of these sources to provide accurate estimations.

Research objective Study 1

Study 1 included in this dissertation aimed at providing further evidence for the embodied perspective on depth perception described above. Depth

perception relies on the interpretation of ambiguous cues, which can only be accurate when additional sources of information are taken into account (Cont & Zimmermann, 2021; Longo & Lourenco, 2007; Ogawa et al., 2018; van der Hoort et al., 2011; Volcic et al., 2013; Wiesing et al., 2021). While many sources of information have already been shown to influence our interpretation of these cues, a potential contribution of the perceived self-location in space was not yet investigated. This is not surprising given that this internal representation of one's own position is usually directly linked to the actual location in space. The full-body illusion represents a convenient tool to dissociate this link between the perceived self-location and the actual location as it shifts the perceived self-location away from the actual location towards an avatar that was used during its induction (Aspell et al., 2009; Ionta et al., 2011; Lenggenhager et al., 2011; Lenggenhager et al., 2007; Palluel et al., 2011). By making use of this illusion, Study 1 tried to provide additional evidence for embodied accounts of perception by showing that depth perception processes take the perceived self-location into account.

Perception of a stable world

Eye and head movements

In order to perceive the area of interest with the fovea, we perform gaze shifts. Gaze shifts are realignments of the line of sight and typically comprise a leading eye and a subsequent head movement (Sidenmark & Gellersen, 2019). The eye movement performed in the context of a gaze shift is a saccade, a fast and short eye movement, that lasts between 20 and 50 milliseconds and reaches speeds of up to 600 visual degrees per second (C. M. Harris & Wolpert, 2006). Microsaccades, smooth pursuit eye movements, nystag-

mus, and converging or diverging movements represent the other types of movements the eye is able to perform (Dodge, 1903). Microsaccades are performed during fixation and fulfill the purpose of correcting for gaze drift caused by the eye muscles, e.g., due to fatigue (Di Stasi et al., 2013). Smooth pursuit saccades, as the name suggests, are utilized to track a moving object. Nystagmus is a combination of saccades and pursuit movements. Research differentiates the nystagmus based on the trigger and therefore distinguishes between the optokinetic and the vestibular nystagmus. The first is triggered when we try to fixate a moving object which covers a large portion of our visual field, e.g., the landscape we perceive while we look out of the window of a moving car or train (Büttner & Kremmyda, 2007). The purpose of the latter is to compensate for movements of our own body (Abadi, 2002), e.g., during the performance of head movements.

The muscles supporting and controlling the head are capable of letting us perform head movements on every plane, e.g., pitch, yaw and roll movements. Head movements typically follow the main sequence, defined by Zangemeister et al. (1981), which states that the average peak velocity of head movements is higher the larger the amplitude of the head movement is. Head movements differ from eye movements with regard to the control we have over them during their performance. Zimmermann (2021) was able to provide evidence for online adjustments to the trajectory of head movements. In contrast, saccades are pre-planned movements that we are unable to change online (D. Robinson, 1975). Another difference between the two types of movement is their duration. Head movements towards a gaze target usually last between 400 and 800 milliseconds (Andres & Hartung, 1989; Hoffmann et al., 2017), which is much longer than the average duration of a saccade of 30 ms to 50 ms (C. M. Harris & Wolpert, 2006). Due to this difference, it is not

surprising that the visual processing during head movements differs from that of saccades. The performance of saccades is linked to mechanisms that suppress the processing of the visual input slightly before and during the performance of a saccade (Binda & Morrone, 2018). While there is evidence that such processes are also present for head movements (Adelstein et al., 2006), the suppression applied to the processing of visual input during head movements seems to be much less pronounced. The differences regarding the suppression of visual input during eye movements compared to head movements are in line with findings in neurophysiology, which suggest that the signals of eye and head movements seem to be separately processed on a neural level (Freedman et al., 1996; Freedman & Sparks, 1997; Walton et al., 2008).

Spatial remapping

Our perception of the world relies heavily on our ability to differentiate between self- and externally-produced motion. Motion sickness is a prime example for the consequences when this ability fails (Money, 1970). The exact underlying mechanisms of this ability are a long standing topic of discussion (Gruesser, 1986; van der Steen, 1998; Wallach, 1985, 1987; Wertheim, 1994). Early accounts of visual stability focused primarily on saccades and argued that we achieve visual stability with the information conveyed by efference copies (Sperry, 1950; von Holst & Mittelstaedt, 1950). The efference copy is a signal which comprises the properties of a planned movement, allowing us to subtract this information from the visual input we perceive (Sperry, 1950; von Holst & Mittelstaedt, 1950). Single cell studies performed with monkeys suggest that we perform this subtraction by shifting the receptive fields of neurons in the lateral intraparietal area, the V3A and the frontal eye field

based on the information available about the planned saccade (Duhamel et al., 1992; Hall & Colby, 2011; Nakamura & Colby, 2002; Sommer & Wurtz, 2004). This process is called *spatial remapping* and seems to be performed prior to the movement onset and should therefore be regarded as a predictive mechanism (Melcher, 2007). Studies performed with humans are in line with these results (Ross et al., 2001) and indicate that also in the human brain spatial remapping occurs in the parietal (Medendorp et al., 2003; Merriam et al., 2003) and extrastriate cortex (Merriam et al., 2007).

Research has proposed different mechanisms for how our system performs spatial remapping. One line of research proposes a global mechanism that spatially remaps the whole visual field whenever we produce any kind of motion affecting our visual field (Bischof & Kramer, 1968; Breitmeyer et al., 1982; Bridgeman et al., 1994; Irwin et al., 1994; Sperry, 1950; von Holst & Mittelstaedt, 1950). There are three different approaches to this global mechanism that could be used: the comparison between the predicted and perceived outcome, the use of spatiotopic maps and the use of a reference object (Bridgeman, 1983). The comparison approach assumes that, a prediction of the visual outcome of each self-produced motion based on the respective efference copy is computed and then compared to the actual outcome. Any deviation will then be attributed to external motion (Bridgeman, 1983). The second approach suggests the use of a coordinate system that is built upon the information conveyed by the efference copy and the information perceived with the retina (Bischof & Kramer, 1968; Breitmeyer et al., 1982). Finally, the reference object approach suggests that the spatial remapping is performed in relation to the position of a single or several objects of interest in the environment (Bridgeman et al., 1994; Irwin et al., 1994). In contrast to these global mechanisms, local mechanisms expect only objects

of importance to be adjusted based on self-produced motion (Goldberg & Bruce, 1990; Gottlieb et al., 1998; Kusunoki et al., 2000; Sommer & Wurtz, 2004; Szinte & Cavanagh, 2011), leaving all other objects in the environment uncorrected. More recent accounts favor local mechanisms over global ones (Hall & Colby, 2011).

Most of the research in this field focused solely on eye movements (Bridge-man, 1983; Fracasso et al., 2010; Honda, 1991; Rock & Ebenholtz, 1962; Szinte & Cavanagh, 2011; Wurtz, 2018), neglecting the potential role of head movements. This represents a simplification of the underlying problem. The gaze shifts we perform often include a head movement component, especially in unrestricted environments and in the context of larger gaze shifts (Freedman & Sparks, 1997; Sidenmark & Gellersen, 2019). Due to the separation of eye and head movement signals on a neural level (Freedman et al., 1996; Freedman & Sparks, 1997; Walton et al., 2008), there are several possible ways the sensorimotor system could potentially perform the spatial remapping for eye-head gaze shifts. The sensorimotor system could perform the spatial remapping solely on the information available about the saccade or it could incorporate the information about the head movement into the spatial remapping, either as a separate remapping or within the remapping performed for the saccade. Head movements have the potential to contribute much information to remapping processes given the amount of sources that provide information about them. These include vestibular signals, neck proprioception and efference copy information (Crowell et al., 1998; L. R. Harris, 1994; Lisberger, 1984; Mergner et al., 1992). Against the background of embodied accounts of visual perception and the above described mechanisms of multisensory integration, it seems very unlikely that the available information about head movements is neglected, discarded or disadvantageous for

spatial remapping.

Serial dependence

Serial dependence is a mechanism thought to aid in the perception of our environment by smoothing out discontinuities (Cicchini et al., 2018; Fischer & Whitney, 2014). This smoothing is accomplished by biasing our perception of the perceived input in the direction of the input perceived previously. For example, the orientation of a cuboid stimulus we have perceived previously influences our perception of the orientation of a cuboid stimulus we are currently perceiving (Fischer & Whitney, 2014). The involvement of this mechanism has been shown for a variety of perceptual aspects, including body size (Alexi et al., 2018), eye gaze (Alais et al., 2018), facial identity and expression (Liberman et al., 2014; Taubert & Alais, 2016), head movements (Zimmermann, 2021), locomotion (Wiesing & Zimmermann, 2023), numerosity (Bliss et al., 2017), pulchritude (Taubert & Alais, 2016), saccades (Cont & Zimmermann, 2021), spatial position (Bliss et al., 2017; Manassi et al., 2018), visual orientation (Cicchini et al., 2017; Fischer & Whitney, 2014; Fritsche et al., 2017), and visual stability (Manassi & Whitney, 2022).

Research objective Study 2

Study 2 included in this dissertation extended the line of research on visual stability described above. Our sense of visual stability relies heavily on our ability to differentiate between externally and self-produced motion and the compensation for the latter. Study 2 aimed at investigating how we maintain our sense of visual stability in the context of head movements. As many factors can influence movements of the head and eye, e.g., fatigue and load, the sensorimotor system faces frequently changes in the relation

between planned movement and the visual consequences of the respective movement. Despite the potential value of feedback and the use of feedback in future predictions, little attention has been given to the role of experience in this context (Gruesser, 1986; Rineau et al., 2023; van der Steen, 1998; Wallach, 1985, 1987; Wertheim, 1994; Wexler, 2003). Study 2 therefore explored the role of past experiences and how these experiences shape our sense of visual stability, e.g., how and when we adjust our expectations of the consequences of our own movements based on the experiences we gain.

Space constancy

Differentiating between self- and externally-produced motion is crucial for our ability to keep an accurate spatial representation of objects in our surroundings, termed *space constancy* (Angelaki & Hess, 2005; Bridgeman, 1983; Cullen, 2019). Due to the overlapping requirements of visual stability and space constancy, both abilities are interconnected. Without accounting for the exact self-produced motion, we are unable to judge if objects in our surroundings moved or if the spatial shift of their position in relation to us was caused by the movements we performed.

Research objective Study 3

Study 3 included in this dissertation extended the line of research on space constancy in the context of gaze shifts described above. In Study 3 we investigated if spatial remapping processes take information about head movements into account or are solely performed for saccades. We have access to many sources that provide information about our head movements (Crowell et al., 1998; L. R. Harris, 1994; Lisberger, 1984; Mergner et al., 1992). Against the background of multisensory integration mechanisms, it is

unlikely that these sources of information are simply neglected or discarded. Despite the potential for the mere contribution of additional information about the performed gaze shift, e.g., in the form of the efference copy of the head movement, head movements could enhance spatial remapping also in other ways. We therefore also explored the potential benefits of the performance of eye movements during the head movement for spatial remapping in Study 3.

Overview of studies

This dissertation comprises three studies which investigated the effect of body-related information and sensorimotor experience on visual perception. All studies were in line with the ethical standards defined in the Declaration of Helsinki. Written informed consent prior to the voluntary participation was given by every participant. The following sections include summaries of the three studies. Full details are provided in the original research articles in the appendix.

Study 1

Research question and hypotheses

Depth perception relies on monocular and even more strongly on binocular cues (Nicolle et al., 1995; Servos et al., 1992). One of these binocular cues is the difference in the retinal projection of objects on the left and right eye, termed binocular disparity. Binocular disparity is highly ambiguous (Johnston, 1991; Norman et al., 1996). Far away bigger objects can produce the same retinal projection as smaller closer objects. One approach the brain

might take in order to resolve this ambiguity is to interpret the retinal projections with the help of sensorimotor knowledge. This knowledge could for example be the trajectory of the movement that would be necessary to grasp the perceived object. Based on the planned movement one can infer the position of the object by adding the vector of the movement to our known position in space (Cont & Zimmermann, 2021; Volcic et al., 2013; Wiesing et al., 2021). If depth perception does take information about potential movements and our own position into account, then changes regarding these information sources should be reflected in the way we perceive objects in depth.

As there is already evidence that changes in our perceived limb length influence depth perception (Linkenauger et al., 2015), we opted to check if changes in our perceived location in space also influence our perception of objects in depth. In order to experimentally manipulate the perceived location in space, we used the full-body illusion, an illusion that is known to shift the perceived self-location of participants (Aspell et al., 2009; Ehrsson, 2007; Lenggenhager et al., 2011; Lenggenhager et al., 2009; Lenggenhager et al., 2007; Nakul et al., 2020). This illusion is induced by tactile stimulation of the participants' back. During this stimulation, the participant is in a virtual reality setup and observes an avatar which undergoes a tactile stimulation that is congruent to what the participant is experiencing. This induction is then compared to the same sequence but with a tactile stimulation on the avatars back that is incongruent to the experienced tactile stimulation (Ehrsson, 2007; Lenggenhager et al., 2007). The resulting shift of the perceived self-location occurs on the anterior-posterior axis and manifests itself in a forward shift towards the avatar (Aspell et al., 2009; Lenggenhager et al., 2011; Lenggenhager et al., 2009; Lenggenhager et al., 2007; Nakul et al., 2020).

This shift is thought to be caused by conflicting information provided by individual senses causing the brain to find a compromise between them (Aspell et al., 2009; Ionta et al., 2011; Lenggenhager et al., 2011; Lenggenhager et al., 2007; Palluel et al., 2011). This compromise follows the known principles of multisensory integration which dictate that the information provided by the most accurate involved sense is weighted more strongly during the integration and biases the involved less accurate senses (Ernst & Banks, 2002). In the induction of the full-body illusion three senses are involved: vision, proprioception and touch (Blanke et al., 2015). Vision provides the most accurate signal out of the three and is therefore prioritized during the integration of the signal. Additionally, the other senses are biased towards the input of vision, suggesting the participant to stand closer to the shown avatar than they objectively are (Aspell et al., 2009; Ionta et al., 2011; Lenggenhager et al., 2011; Lenggenhager et al., 2007; Palluel et al., 2011).

Research on the peripersonal space, i.e., the space in which we integrate multisensory body-related signals (Noel et al., 2015), was able to provide support for a more accurate perception of objects within this space. This initial dichotomous view was abandoned due to the findings of Bufacchi and Iannetti (2018) that suggest a gradual transition rather than a simple in or out classification. Based on these findings and the already extensive research on the full-body illusion, we hypothesized that the effects the full-body illusion has on the perception of the self-location in space should influence depth perception processes. As the full-body illusion causes a forward shift of the perceived self-location and peripersonal space, we expected that a reduced distance between the presented stimuli and the boundaries of the peripersonal space is used in the calculations to resolve binocular disparity, which in turn should enhance depth perception performance.

Methods

The sample of Study 1 consisted of 20 participants who performed a visual depth and a mental imagery task in a virtual reality setup. The visual depth task required participants to judge which of two briefly presented spheres was closer to their perceived location in a two-alternative forced choice task. In each trial, one of the two spheres was manipulated with regard to its position on the anterior-posterior axis, resulting in this sphere being further away, at the same distance, or closer to the participant than the other sphere. In the mental imagery task, participants were approached by a ball rolling towards them on the floor of the virtual environment at a slight angle, starting at the opposite side of the room. Before the ball reached the position of the participant, the vision of the participant was fully blocked and they were instructed to imagine the ball rolling at the same speed as before and to indicate when they thought the ball had reached their position. This task is a modified version of the task of Nakul et al. (2020). In the original version, the ball is rolling towards the participant on the anterior-posterior axis and not at an angle. Due to the visual depth task presenting the stimuli either in the left or right visual field, we here opted to adjust the mental imagery task to be more comparable to the visual depth task in order to have a more suitable measure of the peripersonal space.

Participants performed two experimental sessions, one with congruent and one with incongruent tactile stimulation. These stimulations were followed by seven trials of the visual depth task and two trials of the mental imagery task. This sequence of tactile stimulation and trials was repeated ten times before participants filled out a questionnaire at the end of the session in order to quantify the intensity of the experienced full-body illusion (Salomon et al., 2013). It was necessary to repeat the respective tactile stim-

ulation throughout a session in order to avoid a critical decay of the potential effect of the full-body illusion on the perceived self-location (Keenaghan et al., 2020).

Results and discussion

The questionnaire data provided evidence for a meaningful difference between the experimental sessions with congruent and incongruent tactile stimulation sequences, as participants reported to have identified themselves more strongly with the avatar and experienced the tactile stimulation applied to the avatar more as if it occurred on their own physical back in congruent sessions. Based on this data we split the sample into two groups, one which experienced the full-body illusion more intensely and one which did experience the full-body illusion in a weaker form.

In order to analyze the data of the visual depth task, we fitted psychometric functions to the responses of the participants and calculated the point of subjective equality (PSE) and the just-noticeable difference (JND). The PSE gives insight into the absolute depth perception of participants, i.e., the accuracy of their depth perception judgement. The JND on the other hand provides information about the precision or sensitivity with which participants were able to differentiate between different stimuli positions.

The JND in the visual depth task was reduced in experimental sessions with a congruent tactile stimulation sequence but only for the group of participants who reported to have experienced a strong full-body illusion. This indicates that this group of participants was more sensitive to distance changes of the stimuli than the group which reported to have experienced the full-body illusion in a weaker form. The most likely explanation for this effect is a shift of the peripersonal space of the participants to the position of the

avatar (Aspell et al., 2009; Ionta et al., 2011; Lenggenhager et al., 2011; Lenggenhager et al., 2007; Noel et al., 2015; Palluel et al., 2011). This result is surprising given that an illusion is causing an improvement in depth perception judgements which pose a vital task of our visual perception but is in line with other studies using illusions which also caused improvements in perceptual tasks (Vignemont et al., 2005; Volcic et al., 2013). No effects for the PSE were found, confirming that the sessions with congruent and incongruent tactile stimulation sequences were highly comparable and only differed from one another regarding the applied tactile stimulation.

The mental imagery task was analyzed with regard to the distance between the approaching ball at the time of the participants' response and the participant. This distance quantified the perceived position of the participants in space in relation to their objective position. The data of the mental imagery task failed to provide evidence for a quantifiable shift of the perceived self-location shift of the participant induced by the full-body illusion for both groups of participants. This failure to quantify a shift towards the avatar can most likely be attributed to the modifications of the task in comparison to the original task introduced by Nakul et al. (2020). In the original task the ball was rolling towards the participant on the anterior-posterior axis, due to the angle we used in our version of the task, it seems that our version of the task was unsuitable to quantify the shift towards the avatar.

Conclusion

The full-body illusion induced shift in the perceived self-location in the group of participants which experienced the full-body illusion more intensely led to an improvement in depth perception performance reflected by lower JNDs. This provides further evidence that ambiguous binocular cues are

interpreted against the background of additional information including the perceived self-location in space.

Study 2

Research question and hypotheses

Being able to differentiate between self-produced and external motion is of high importance for visual stability, spatial constancy and movement planning (Rineau et al., 2023; Wallach, 1987; Wexler, 2003). This is highlighted by the consequence of failing to differentiate between these motion sources, i.e., motion sickness accompanied by nausea (B. Cohen et al., 2019). The differentiation between externally and self-produced motion was mostly investigated in the context of saccades with a fixed head position (Wurtz, 2018). Bigger gaze shifts on the other hand, especially in an unrestricted environment, include head movements (Freedman et al., 1996; Freedman & Sparks, 1997; Walton et al., 2008). A gaze shift starts with a saccade towards the target area, followed by a head movement to center the vision around this area (Freedman & Sparks, 1997; C. M. Harris & Wolpert, 2006; Sidenmark & Gellersen, 2019). During the head movement, the vestibulo-ocular reflex causes a saccade in the opposite direction of the head movement which stabilizes the image on the retina (Barnes, 1979). Despite this compensation of the motion caused by the head movement, we are still able to perceive the motion, while we do not perceive the motion caused by saccades (Binda & Morrone, 2018).

One way to differentiate self from externally caused motion is the efference copy, a signal based on the movement command, which is thought to be used in the mechanism compensating for the predicted motion caused by the

planned movement (Sperry, 1950; von Holst & Mittelstaedt, 1950). While this seems to be a sound mechanism for saccades (Sperry, 1950; von Holst & Mittelstaedt, 1950), head movements are influenced by additional sources of noise, like an exhaustion of the involved neck muscles (Monjo et al., 2015; Scotland et al., 2014). Additionally, there is the potential of online changes of the head movement that are not accounted for by the initial efference copy (Zimmermann, 2021). It is therefore likely that other sources of information play an important role to compensate for the motion caused by head movements, especially when the separate processing of head movements is taken into account (Freedman et al., 1996; Freedman & Sparks, 1997; Walton et al., 2008). The research question of study 2 was therefore if past experiences influence our sense of visual stability and how we use these experiences to adjust what we perceive as visually stable. We hypothesized that we use our past experiences to adjust what we perceive as visually stable, e.g., if a performed head movement results in more motion than we anticipated, we adjust our expectation and shift what we perceive as visually stable closer to this experience. This in turn means that the next time we perform the same movement, we would expect to perceive more motion than before.

Methods

The sample of study 2 consisted of 38 participants. To manipulate the relation between head movements and the associated motion, a virtual reality setup was employed. A trial within this setup consisted in the fixation of a target presented in the left or right visual field of the participant. Once the participant fixated this target, another target appeared in the opposite visual field. Participants were instructed to perform a head movement to the second target once it disappeared. During this head movement, a visual velocity gain

was applied which manipulated the behavior of the environment in relation to the head movement. In order to change this relation, the environment was virtually rotated more than the head movement required, either against or in the direction of the head movement. This manipulation effectively changed the speed of the head movement by suggesting faster or slower motion on the retinae. Importantly, this manipulation was tied to the performed head movement in the trial, i.e., if the participant performed a head movement of 10 rotational degree and a visual velocity gain of 1.1 was active, the result of the head movement was a shift of 11 rotational degree instead of 10. The task of the participants was to decide if they perceived the gaze-contingent visual motion as faster or slower than what they would judge to be unaltered in a two-alternative forced choice task. The virtual environment was void of any salient references and consisted of a grating with a spatial frequency of 0.05 c/deg (specified in rotational degree) which encircled the participant fully.

Results and discussion

The results of study 2 suggest that participants shifted their criterion of what they perceive as unaltered with regard to the relation between head movement and corresponding motion. This shift can be explained by serial dependencies, a phenomenon that uses prior experiences to align currently perceived signals with past experiences in order to smoothen the perceived perceptual input (Cicchini et al., 2018; Fischer & Whitney, 2014). When a faster gaze-contingent visual motion occurred in the previous trial, participants shifted their criterion of visual stability towards faster movements, expecting their head movements to cause faster motion on their retinae. The same pattern but in the opposite direction was observed for gaze-contingent

visual motions that were slower than anticipated. This is a clear indication that the underlying mechanism that shapes our criterion of visual stability is experience-based. Any other mechanism that does not take past experiences into account would have failed to adjust to the manipulations which, in turn, would have led to motion sickness. The found criterion shift of what visual velocity gain is perceived as visually stable also seemed to depend on the similarity between the prior experience and the current visual input. Participants showed descriptively stronger criterion shifts when the previous head movement was performed in the same direction as the current head movement.

Half of the participants were able to differentiate between the different visual velocity gains based on the direction of the respective visual velocity gain, e.g., they perceived faster visual velocity gains as faster and slower visual velocity gains as slower than the unity visual velocity gain, i.e., no manipulation of the visual consequences of the performed movement. In order to differentiate between faster and slower visual velocity gains, one has to integrate the signal about the own head movement and the perceived motion in the scene. If this integration is not performed and the participant is evaluating the visual consequences and the information available about the performed head movement independently from one another, a different response pattern emerges. This different response pattern was observed for the other half of participants. These participants did not consider the direction of the visual velocity gains, e.g., a visual velocity gain of 1.3 was perceived as equally fast as a visual velocity gain of 0.7. But they were still able to differentiate between individual visual velocity gains very well, e.g., they were able to differentiate a visual velocity gain of 1.15 from a visual velocity gain of 1.3. This pattern indicates that this subgroup of participants judged the

deviation from visual stability instead of the integrated result of their own movement and the applied visual velocity gain.

While their responses did not reflect an integration of the visual velocity gain and their own movement, their response pattern can only be achieved by subtracting the visual motion consequences of self-produced movements from the visual input. Only then one can judge the visual motion caused by the applied visual velocity gain and differentiate between different visual velocity gain speeds. The fact that this group of participants judged faster and slower visual velocity gains as equally fast makes sense given that the environment is artificially rotated more the further the applied visual velocity gain deviated from unity visual velocity gain, i.e., the environment is rotated the same amount for a visual velocity gain of 0.9 and 1.1 even though these gains are applied in different directions.

The way this subgroup of participants shifted their criterion of visual stability is more complex than for the other group. The group of participants which integrated both signals shifted their criterion towards the visual velocity gain active in the previous trial. This shift was more pronounced when the head movements of the previous and current trial were performed in the same direction. For the subgroup of participants which did not integrate the signals, a criterion shift only occurred when the error they perceived in the previous trial was similar to the error they perceived in the current trial. The error in this context denotes the deviation from visual stability. To give an example, if a participant of this subgroup performed a head movement to the left, retinal motion to the opposite direction was expected. Depending on the visual velocity gain active during the head movement, this retinal motion was either as expected, amplified or reduced. When the same participant performed once again a head movement to the left, the error signals

should match. For participants who did not integrate the signal of their own movement with the visual consequences, this match of errors only occurred when a faster visual velocity gain was active in the current trial as a faster visual velocity gain caused an error in the opposite direction of the performed head movement. This finding is in line with previous studies suggesting that serial dependencies can act directly on perception instead of influencing later cognitive processing stages (Cicchini et al., 2017).

The adjustment to the altered relation between head movement and corresponding motion was performed fast. Participants were able to adjust their head movement online to the visual velocity gain, i.e., slowing down the head movement when the visual velocity gain suggested a faster head movement. Furthermore, this also shows that information about head movements is very precise and readily available information. A single or a combination of different sources could provide this information, including the vestibular system, neck proprioceptive, or the efference copy of the respective head movement (Crowell et al., 1998; L. R. Harris, 1994; Mergner et al., 1992). As the visual velocity gain was tied to the respective head movement, only information about the head movement itself was able to provide the necessary information precisely enough to predict the caused motion and to adjust the head movement dynamics.

Conclusion

The results of the study provide strong support for an experience dependent mechanism that constantly updates our expectation of motion caused by our own head movements. Based on the underlying recalibrations the sensorimotor system is even enable to perform online adjustments of head movements, displaying a fast and very frequent processing of the related

information. This highlights once again that our perceptual system takes all available information sources into account, including recent sensorimotor experiences.

Study 3

Research question and hypotheses

Space constancy is the ability to keep an accurate spatial representation of our surroundings despite performing movements constantly (Angelaki & Hess, 2005; Bridgeman, 1983; Cullen, 2019). Gaze shifts represent the most frequent and drastic visual change that we have to account for in order to maintain space constancy. During the performance of a gaze shift additional external motion could occur in the scene, it is therefore necessary to have an accurate estimation of the visual consequences of the performed gaze shift in order to be able to still notice additional external motion (Angelaki & Hess, 2005; Bridgeman, 1983; Cullen, 2019). Electrophysiological studies suggest that these predictions occur prior to the performance of a saccade and are used to perform a spatial remapping (Duhamel et al., 1992; Hall & Colby, 2011; Nakamura & Colby, 2002; Sommer & Wurtz, 2004). This spatial remapping based on the predictions about the consequences of the performed saccade enables the distinction between externally generated and self-produced motion, given that the prediction about the self-produced motion is accurate. Up to this point this process was extensively investigated with pure saccade gaze shifts (Fracasso et al., 2010; Honda, 1991; Rock & Ebenholtz, 1962; Szinte & Cavanagh, 2011; Wurtz, 2018). Despite the possibility that this remapping is solely solved based on information about the saccade, gaze shifts including a head movement could benefit from in-

formation available about the head movement. It is a well established fact that we have access to at least three information sources regarding our head movement, including vestibular signals, neck proprioception and efference copy information (Crowell et al., 1998; L. R. Harris, 1994; Lisberger, 1984; Mergner et al., 1992).

To test this possible contribution of head movements in spatial remapping processes in the context of gaze shifts, we let participants perform either saccades or gaze shifts including an eye and head movement in a spatial localization task. As there is evidence for additional information provided by head movements in the form of different information sources, e.g., neck proprioception, we expected a facilitation of spatial remapping accuracy when the gaze shift included a head movement component. Furthermore, we also let participants perform two different types of eye-head gaze shifts. Participants were either unrestricted in their performance of the gaze shift or were required to first perform a saccade and then to fixate the target during the performance of the head movement. When participants performed a head movement while they fixated a target, less motion should have occurred on the retinae. The idea behind this manipulation was to explore the potential benefits that the performance of additional eye movements during the head movement could provide. If the performance of additional eye movements during the head movement provides additional value to the spatial remapping processes, the accuracy of these processes should have improved. In an additional experiment, we introduced artificial background motion during the performance of eye-head gaze shifts. The purpose of this experiment was to explore the use of background motion in spatial remapping processes. We expected that the introduction of additional background motion reduced spatial remapping accuracy, indicating that this source of information is taken

into account in the context of spatial remapping processes.

Methods

A virtual reality setup was used to present stimuli before and after the performance of gaze shifts that either included solely an eye movement or an eye and head movement. Study 3 included three separate experiments. The sample of Experiment 1 and 2 included 24 participants whereas the sample of Experiment 3 included 26 participants. In Experiment 1 participants performed an eye-head gaze shift that was either unrestricted or sequential. A sequential eye-head gaze shift started with the performance of a saccade, followed by the fixation of the gaze shift target and the head movement. This manipulation was intended to offer insights into space constancy when the target is stabilized on the retina. This should have reduced the overall perceived motion of the head movement.

Each trial started with the presentation of a stimulus while the participants fixated a fixation cross. They then were instructed to perform a gaze shift to a target. Once they arrived at and fixated the target location, a second stimulus was briefly presented. The position of the second stimulus varied across trials. The task of the participants was to judge if the second stimulus was to the left or right of the first stimulus. Participants underwent several conditions in which the amount of visual references during the performance of the task varied, e.g., in some conditions the background was a grating instead of the otherwise gray background without references.

Experiment 2 was identical to Experiment 1 with the exception that the gaze shift only comprised a saccade and no head movement. By letting participants perform only saccades in this experiment we had the opportunity to directly compare the spatial remapping accuracy in eye and eye-head gaze

shifts. Experiment 3 was identical to Experiment 1 with regard to the performed gaze shifts. This experiment introduced an additional manipulation in the form of scene displacements. In this experiment the background during the performance of the eye-head gaze shift was always a grating. This grating was either rotated according to the performed head movement or was rotated more than the head movement required. This introduction of additional external motion served the purpose of further investigating the contributions of the head movements as this external motion was directly linked to the performed head movement.

Results and discussion

Participants were fairly accurate in their spatial localization judgements when performing merely saccades but this was only the case as long as visual references were available. When visual references were absent, participants underestimated the amplitude of their saccade in their spatial remapping. This underestimation was not observed when gaze shifts with a head movement were performed. For these gaze shifts the amount of visual references had little influence on the accuracy of the spatial remapping, providing evidence for a facilitating contribution of head movements to spatial remapping processes. The information provided by head movements, including vestibular, neck proprioception and efference copy information (Crowell et al., 1998; L. R. Harris, 1994; Lisberger, 1984; Mergner et al., 1992), were able to counteract the underestimation that would occur when only information about the saccade was available.

The execution of gaze shifts which included a head movement component had only an impact on spatial remapping accuracy in Experiment 3 but not in Experiment 1. This discrepancy between experiments was resolved

by taking the fixation performance of the gaze shift target during the head movement into account. The difference between unrestricted and sequential gaze shifts regarding the fixation duration was much larger in Experiment 3. The more time participants spend performing eye movements during the head movement of the gaze shift the higher the accuracy of the spatial remapping was, indicating that eye movements performed during the head movement component of gaze shifts play a crucial role in spatial remapping processes. Finally, the artificial background motion introduced in Experiment 3 was able to provide support for the inclusion of background motion information in spatial remapping processes. Additional artificial motion during the head movement component led to less accurate spatial remappings.

Conclusion

The information provided by head movements is especially important when visual references are absent during the performance of gaze shifts as the provided information is able to counterbalance the otherwise occurring underestimation of the performed saccade. The performance of saccades and the perceived background motion during the head movement component of a gaze shift also seem to play crucial roles in our ability to accurately perform spatial remappings of objects perceived prior to a gaze shift. To conclude, the sensorimotor system takes all available sources of information into account to maintain accurate spatial representations of objects in our surroundings despite the gaze shifts we perform frequently.

General Discussion

The studies in this dissertation investigated the role of bodily information and sensorimotor experiences on visual perception. Study 1 extended a line of research on the role of bodily information on depth perception by manipulating the perceived self-location of participants with the full-body illusion, an illusion that utilizes the mechanisms of multisensory integration. This enabled us to test if the perceived self-location influences depth perception sensitivity by shifting the peripersonal space, i.e., the space in which we are able to interact with our environment, closer to the stimuli (Longo & Lourenco, 2007; Rizzolatti et al., 1981; Wiesing et al., 2021). Study 1 showed that our ability to differentiate between objects at varying distances is enhanced when our perceived self-location in space is closer to the respective objects. Study 2 expanded our knowledge on visual stability in the context of head movements. By manipulating the gaze contingent visual velocity during head movements, we altered the visual consequences of head movements and were therefore able to test if sensorimotor experiences influence what we perceive as visually stable. The results of Study 2 confirmed that our sense of visual stability is constantly updated based on sensorimotor experiences we gain during the performance of head movements. Study 3 provided further insights into the interaction of eye and head movements in the context of space constancy and spatial remapping. By manipulating the way how gaze shifts were performed, Study 3 revealed that while the spatial remapping of objects works well when we only perform saccades, the accuracy of the remapping relies strongly on the presence of spatial references. The information provided about performed head movements can compensate for the accuracy reduction when spatial references are absent. Moreover, the perfor-

mance of eye movements and the background motion perceived during the head movement component of a gaze shift even further facilitate accurate spatial remapping. Overall, Study 3 revealed that spatial remapping processes rely on several information sources and that the absence of individual sources of information leads to less accurate spatial representations.

Integration of bodily information in visual perception

There is a growing amount of evidence for the integration of body-related information during visual perception processes (Longo & Lourenco, 2007; Ogawa et al., 2018; van der Hoort et al., 2011; Volcic et al., 2013; Wiesing et al., 2021). Study 1 provides evidence that the perceived self-location in space is accounted for in depth perception processes. Participants under the influence of the full-body illusion showed enhanced depth perception sensitivity in a two-alternative forced choice task in which they had to judge which of two stimuli is closer to their perceived location in space. As the baseline condition only differed from the experimental condition with regard to the congruency of the tactile stimulation, the found effect can only be attributed to the influence of the full-body illusion. The induction of the full-body illusion requires tactile stimulation that is congruent to the visual input of an avatar or object that the participant can identify with (Ehrsson, 2007; Lenggenhager et al., 2007). The baseline condition includes all of these tactile and visual inputs with the exception that the visual input is incongruent to the felt tactile stimulation and therefore fails to induce the full-body illusion (Aspell et al., 2009; Lenggenhager et al., 2011).

The fact that an illusion can enhance depth perception sensitivity is surprising. The reason for this enhancement lies most likely in the ambiguity of the cues depth perception has to rely on in order to estimate the distance

towards objects (Nicolle et al., 1995; Servos et al., 1992). We especially rely on binocular cues (Nicolle et al., 1995; Servos et al., 1992), as many of the monocular cues, like accommodation and motion parallax are unreliable sources of information to accurately estimate the distance to objects (Ferris, 1972; Fisher & Ciuffreda, 1988). While binocular cues are a more suitable source of information to estimate distances, they are also not reliable enough on their own. Binocular disparity, the difference in the retinal projection of objects on the left and right eye (Johnston, 1991; Norman et al., 1996), is for example highly ambiguous. Far away bigger objects can create the same projection as smaller closer ones. Due to their ambiguity and unreliability, these cues are unlikely to be interpreted independently from other sources of information. Prime candidates for such information are bodily information sources (Longo & Lourenco, 2007; Ogawa et al., 2018; van der Hoort et al., 2011; Volcic et al., 2013; Wiesing et al., 2021; Zimmermann, 2021). Research in this field already provided evidence for the integration of information about the size of our limbs (Longo & Lourenco, 2007; Volcic et al., 2013), our grasping distance (Wiesing et al., 2021), the size of our body (Ogawa et al., 2018; van der Hoort et al., 2011) and planned movements (Cont & Zimmermann, 2021) in such processes. Study 1 is able to extend this list of utilized information sources by providing evidence for the integration of the perceived self-location in these depth perception processes.

Effects of full-body illusion are most often explained by a forward shift on the anterior-posterior axis (Aspell et al., 2009; Lenggenhager et al., 2011; Lenggenhager et al., 2007). This forward shift of self-location is also known to be accompanied with a shift of the peripersonal space (Noel et al., 2015). This space is usually defined as the space in which we can manipulate objects, typically restricted by our grasping distance (Longo & Lourenco, 2007;

Rizzolatti et al., 1981; Wiesing et al., 2021), and is associated with an enhanced perception of objects (Bufacchi & Iannetti, 2018; Noel et al., 2015). The presumed shift of the peripersonal space caused by the full-body illusion in Study 1 therefore facilitated the processing of the objects by suggesting to the participants that the objects are closer to them than they objectively are. This is a striking example for the integration of different information sources during the processing of sensory input and is in line with research supporting an embodied view on perception (Longo & Lourenco, 2007; Ogawa et al., 2018; Proffitt, 2006; Proffitt et al., 1995; van der Hoort et al., 2011; Volcic et al., 2013; Warren, 1984; Wiesing et al., 2021; Zimmermann, 2021). Recent findings suggest that perceptual processes account for situation-dependent factors like fatigue (Proffitt, 2006; Proffitt et al., 1995), sensorimotor experiences (Wiesing et al., 2021), information about our motor system (Longo & Lourenco, 2007; Ogawa et al., 2018; van der Hoort et al., 2011; Volcic et al., 2013), and planned movements (Zimmermann, 2021). This line of research highlights that we tend to integrate all available information sources according to the principles of multisensory integration, in order to create the most actionable representation of our surroundings that we can achieve.

The role of experience in visual perception

Besides the utilization of bodily and situational information (Longo & Lourenco, 2007; Ogawa et al., 2018; Proffitt, 2006; Proffitt et al., 1995; van der Hoort et al., 2011; Volcic et al., 2013; Warren, 1984; Wiesing et al., 2021; Zimmermann, 2021), there is also an accumulation of evidence for the influence of past experiences on visual perception (Alais et al., 2018; Alexi et al., 2018; Bliss et al., 2017; Cicchini et al., 2017; Cicchini et al., 2018; Cont & Zimmermann, 2021; Fischer & Whitney, 2014; Fritsche et al., 2017; Liber-

man et al., 2014; Manassi et al., 2018; Manassi & Whitney, 2022; Taubert & Alais, 2016; Wiesing & Zimmermann, 2023; Zimmermann, 2021). The underlying mechanism is called serial dependence and its suggested purpose is to smooth out discontinuities in the visual input we perceive (Cicchini et al., 2018; Fischer & Whitney, 2014). Study 2 provides evidence for the occurrence of this mechanism in the domain of visual stability. We introduced additional scene displacements during the performance of head movements, which led to visual velocity gains faster or slower than the unity visual velocity gain. Experiences with these altered visual velocity gains shifted the criterion, i.e., the visual velocity gain that is perceived as unity visual velocity gain, participants use to judge visual stability. In our study visual velocity gains were amplifiers of the motion caused by head movements, e.g., a visual velocity gain faster than unity gain caused the environment to rotate against the direction of the head movement, creating a faster gaze-contingent visual motion. Experiences with visual velocity gains different than unity gain led to a shift of the criterion of perceived visual stability in the direction of the visual velocity gain, i.e., participants perceived faster visual velocity gains as visually stable when they perceived a faster visual velocity gain during the previous movement. This indicates that the prediction of the anticipated visual consequences caused by head movements is constantly updated based on past experiences.

Head movements differ from saccades with regard to the visual processing occurring during their execution. During saccades the perception of retinal motion is suppressed (Binda & Morrone, 2018; Castet et al., 2002) while retinal motion during head movements is freely accessible. This difference was also observed in Study 2 as participants were able to adjust their head movements online to the active visual velocity gain, e.g., participants slowed

their movement down when a visual velocity gain faster than the unity visual velocity gain was applied. Smooth pursuit eye movements seem to be more similar to head movements with regard to the recalibration processes found in Study 2. Haarmeier et al. (2001) were able to show that experiences with altered scene displacement velocity influenced the perception of the active scene displacement velocity. Luna et al. (2021) additionally showed that these recalibrations were spatially limited by the receptive fields of the corresponding motion-processing neurons. These findings further support the conclusions of Study 2 as the same pattern can be observed in the domain of eye movements.

Most of the research dedicated to serial dependencies focuses solely on isolated perceptual processes, e.g., the effect of the previous orientation of a bar on the perceived orientation of a currently perceived bar (Bliss et al., 2017; Cicchini et al., 2017; Fischer & Whitney, 2014; Fritsche et al., 2017; Liberman et al., 2014; Taubert & Alais, 2016). The results of Study 2 suggest an interaction between the sensory and motor system as the gaze-contingent visual velocity has to be predicted based on the performed movement. Despite the differences in the response patterns, all participants were able to accurately predict or subtract the self-produced motion from the visually perceived motion. Moreover, the found serial dependency effects were only observed when the prior experiences shared features like the direction of the head movement with the currently performed movements. This additionally highlights a requirement for serial dependencies to occur. When past experiences differ drastically from the currently made one, the information and experiences gained in the past are not used to bias the current one (Cicchini et al., 2017). Overall, the found effect shows an involvement of motor information in the processing and calculating of serial dependencies and is

therefore in line with other studies which included motor components in their task (Alais et al., 2018; Cont & Zimmermann, 2021; Wiesing & Zimmermann, 2023; Zimmermann, 2021).

The sub group of participants in Study 2 which subtracted their self-produced motion from their judgments pointed towards an aspect of serial dependencies, also mentioned by Cicchini et al. (2017). The serial dependencies were dependent on the visual input and targeted this input also directly instead of being a mechanism involved in later stages of processing. As Cicchini et al. (2017) point out, there are several different serial dependency mechanisms, including positive, negative, perceptual and post-perceptual mechanisms. The response pattern of the subgroup that subtracted the self-produced motion from their judgment, is best described by positive perceptual serial dependencies. For the group that integrated all available signals, we can only conclude that the involved serial dependencies were positive. To determine at what stage of processing serial dependency processes play a role for this group of participants further research is required. Future experiments could try to disentangle the different possible serial dependency mechanism that could play a role for this group of participants by varying the aspect the task question aims at. In Study 2, participants were asked to judge the visual contingent motion. In order to judge this motion they had to integrate all motion signals before making their judgment. If they were asked to judge the motion of the environment independently of the self-caused motion or to judge the self-produced motion independently from the environment, potential further insights into the role and different mechanisms of serial dependencies in the context of visual stability could be gained. The introduction of different task questions in sequential trials could also aid in determining the level that the serial dependency mechanisms at

play target, as suggested by Cicchini et al. (2017). If for example the previous task question aimed at the perceptual level, e.g. the motion perceived on the retinae, and serial dependencies can be found in the responses to the task question used in Study 2 that aimed at an integrative judgment, we could conclude that the serial dependencies act on a perceptual level. The same procedure could be performed with a task question that aims at the post-perceptual level.

The influence of gaze shifts on spatial representations

Our ability to keep an accurate spatial representation of objects in the context of self-produced motion is highly dependent on information provided about our movements and body in space (Bridgeman, 1983; Duhamel et al., 1992; Wurtz, 2018). Gaze shifts represent the most frequent and drastic spatial shift of objects on our retina. Even though gaze shifts include a head movement component, especially larger ones, studies investigating how we compensate for gaze shifts primarily focused on pure eye movement gaze shifts with a fixed head position (Fracasso et al., 2010; Honda, 1991; Rock & Ebenholtz, 1962; Szinte & Cavanagh, 2011; Wurtz, 2018). These studies came to the conclusion that we are able to accurately compensate for pure eye gaze shifts but show a slight overcompensation for the saccade we perform. Study 3 investigated the contribution of head movements and eye movements performed during the head movement component of the gaze shift on spatial remapping. Participants performed a spatial localization task in which they were required to judge the position of a stimulus presented after the performance of a gaze shift with a stimulus presented prior to the gaze shift. We manipulated the presence of visual references and the performance of the gaze shift. Participants either performed saccades or gaze shifts that

included both, an eye and head movement component. Furthermore, gaze shifts with an eye and head movement were either performed without any restrictions or required the participant to perform the saccade first, fixate the gaze shift target and then perform the head movement while maintaining fixation. In an additional experiment, participants performed gaze shifts including an eye and head movement component with either a static or a rotating environment to explore the possibility that background motion is used in spatial remapping processes.

Study 3 revealed that participants were able to accurately remap stimuli when they performed saccades. However, this performance was highly dependent on the amount of available spatial references. In the absence of spatial references, participants undercompensated for the saccades they performed. This undercompensation was absent when participants performed gaze shifts that included a head movement component. This indicates the importance of head movements in the context of gaze shifts and spatial remapping. Eye movements on their own provide only enough information for accurate remappings when visual references are present. The additional information provided by head movements seems to be necessary to compensate for the absence of such references. These information sources include vestibular, neck proprioception and efference copy information (Crowell et al., 1998; L. R. Harris, 1994; Lisberger, 1984; Mergner et al., 1992). Vestibular signals offer a high temporal resolution with a latency of 14 ms (Lisberger, 1984) and are probably processed together with the information provided by the neck proprioception as the combination of both signals is required for posture, balance and vestibular spinal reflexes (Crowell et al., 1998; L. R. Harris, 1994; Mergner et al., 1992). The efference copy information might be especially useful in the context of spatial remapping as it is an information source avail-

able prior to the performance of the head movement and studies were able to provide evidence that at least some of the spatial remapping occurs prior to the performance of the gaze shift (Duhamel et al., 1992; Nakamura & Colby, 2002). Both vestibular and neck proprioception information is only available once the head movement is already initiated. Since head movements last up to 800 ms (Andres & Hartung, 1989; Hoffmann et al., 2017) and can be changed online (Zimmermann, 2021), the value of the information provided by the vestibular system and neck proprioception increases during the performance of the head movement. It is therefore likely that a combination of all information sources is used to achieve the most accurate spatial remapping possible (Calvert et al., 2004; Ernst & Banks, 2002; Gelder & Bertelson, 2003; Stein & Stanford, 2008).

There is evidence for spatial representation in eye-centered coordinates, especially in the motor domain (Batista et al., 1999; Y. E. Cohen & Andersen, 2000; Vetter et al., 1999). While it is very likely that the sensorimotor system codes spatial representations of objects or actions to some extent in eye-centered coordinates, the most likely approach involves the spatial representations in several coordinate systems, including head-centered coordinate systems. The reasoning behind this assumption is: if additional information sources are available which could further improve the spatial representation of an object, there is no reason for a discarding or neglecting of this information (Pouget et al., 2002). Colby and Duhamel (1996) discovered single neurons that represent spatial positions of objects either in relation to the fovea, the head or even the arm. These different spatial representations are most likely selected and used depending on the effector that is most important for the spatial representation of the object (Colby & Duhamel, 1996). These results imply that the sensorimotor system is in fact using all avail-

able information sources as head- or arm-centered coordinate systems require information beyond the information the eye can provide.

Differences in remapping accuracy were also observed between unrestricted gaze shifts and gaze shifts, which required participants to perform the saccade first and fixate during the head movement. When participants performed eye movements during the head movement component of the gaze shift, they performed the spatial remapping more accurately. This was unexpected given that saccades are known to undershoot their target by roughly 10 % (Becker, 1972; Henson, 1979; D. A. Robinson, 1973) and that movements are always subject to noise (van Beers et al., 2004). While there are several possible types of eye movements that could be performed during the head movement component of a gaze shift, a form of the nystagmus seems to be the most likely one. There are two types of nystagmus, the optokinetic and the vestibular nystagmus. We perform the optokinetic nystagmus when we perceive motion in a large portion of our visual field (Büttner & Kremmyda, 2007). The vestibular nystagmus is performed to compensate for self-produced motion (Abadi, 2002). Given that participants perform active head movements in Study 3, the vestibular nystagmus is most likely engaged to compensate for motion so that the gaze gets stabilized. To accomplish this task, small eye movements are necessary and each of these has the potential to once again adjust the spatial remapping allowing us to achieve a more accurate spatial representation of the prior to the gaze shift perceived stimulus.

Study 3 also revealed that background motion is used as a source of information for the spatial remapping of objects. When the background did not remain static during the performance of gaze shifts including an eye and head movement but moved against the gaze shift direction, participants overcompensated for the gaze shifts they performed. Moreover, this effect was

modulated by the performance of eye movements during the head movement component of the gaze shift. When participants were unrestricted in their execution of eye movements this overestimation was less pronounced than if they were required to fixate the gaze shift target during the head movement component. This indicates once more that the eye movements performed during the head movement component of the gaze shift serve an important purpose in improving the spatial remapping accuracy. Furthermore, this interpretation is also in line with research performed in the field of locomotion (Wiesing & Zimmermann, 2023). While we found worse spatial remapping accuracy when less eye movements were performed, it remains unclear if this difference in spatial remapping accuracy can solely be explained by the performance of less eye movements or if a different kind of eye movements were performed depending on the restriction of the performed gaze shift. Moreover, Study 3 provided evidence for the usage of background motion information in the context of spatial remapping processes. If this background motion information is perceived via the eyes, e.g., by the performance of a specific type of eye movement, or by the information available about the head movements remains unclear. Follow-up studies could explore this ambiguity further.

Conclusion

The results of the studies presented in this dissertation provide evidence for the consideration of bodily information and sensorimotor experiences in visual perception processes. Study 1 was able to provide such support in the context of depth perception by showing that the involved processes account for the perceived self-location in space. Study 2 and 3 focused on our ability to differentiate and account for different sources of perceived motion. Study

2 provided support for the utilization of past sensorimotor experiences during current visual perception processes and Study 3 suggests that all available bodily and motor information is used in order to create the most actionable representation of the world. This dissertation focused on the information sources used during visual processes and the way they are accounted for. The provided insights support an embodied perspective on visual perception and improve our understanding of the integration of different information sources in the context of visual processes even further.

References

- Abadi, R. V. (2002). Mechanisms underlying nystagmus. *Journal of the Royal Society of Medicine*, 95(5), 231–234. <https://doi.org/10.1177/014107680209500504>
- Adelstein, B. D., Li, L., Jerald, J. J., & Ellis, S. R. (2006). Suppressed perception of head-referenced image motion during head movement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(26), 2678–2682. <https://doi.org/10.1177/154193120605002602>
- Alais, D., Kong, G., Palmer, C., & Clifford, C. (2018). Eye gaze direction shows a positive serial dependency. *Journal of Vision*, 18(4), 11–11. <https://doi.org/10.1167/18.4.11>
- Alexi, J., Cleary, D., Dommissie, K., Palermo, R., Kloth, N., Burr, D., & Bell, J. (2018). Past visual experiences weigh in on body size estimation. *Scientific reports*, 8(1), 1–8. <https://doi.org/10.1038/s41598-017-18418-3>
- Andres, R. O., & Hartung, K. J. (1989). Prediction of head movement time using fitts' law. *Human Factors*, 31(6), 703–713. <https://doi.org/10.1177/001872088903100606>
- Angelaki, D. E., & Hess, B. J. M. (2005). Self-motion-induced eye movements: Effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6(12), 966–976. <https://doi.org/10.1038/nrn1804>
- Aspell, J. E., Lenggenhager, B., & Blanke, O. (2009). Keeping in touch with one's self: Multisensory mechanisms of self-consciousness. *PLOS ONE*, 4(8), 1–10. <https://doi.org/10.1371/journal.pone.0006488>
- Barnes, G. R. (1979). Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets. *The Journal of Physiology*, 287(1), 127–147. <https://doi.org/10.1113/jphysiol.1979.sp012650>
- Barrett, K. E., Barman, S. M., S., B., & H.L., B. (2012). *Chapter 10. hearing and equilibrium*.
- Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science*, 285(5425), 257–260. <https://doi.org/10.1126/science.285.5425.257>

- Bechtold, L., Ghio, M., Antoch, G., Turowski, B., Wittsack, H.-J., Tettamanti, M., & Bellebaum, C. (2019). How words get meaning: The neural processing of novel object names after sensorimotor training. *NeuroImage*, 197, 284–294. <https://doi.org/10.1016/j.neuroimage.2019.04.069>
- Becker, W. (1972). The control of eye movements in the saccadic system. *Bibliotheca ophthalmologica: supplementa ad ophthalmologica*, 82, 233–243.
- Binda, P., & Morrone, M. C. (2018). Vision during saccadic eye movements. *Annual Review of Vision Science*, 4(1), 193–213. <https://doi.org/10.1146/annurev-vision-091517-034317>
- Bischof, N., & Kramer, E. (1968). Untersuchungen und überlegungen zur richtungswahrnehmung bei willkürlichen sakkadischen augenbewegungen. *Psychologische Rundschau*, (3), 185–218.
- Blanke, O., Slater, M., & Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron*, 88(1), 145–166. <https://doi.org/10.1016/j.neuron.2015.09.029>
- Bliss, D. P., Sun, J. J., & D’Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, 7(1), 14739. <https://doi.org/10.1038/s41598-017-15199-7>
- Breitmeyer, B. G., Kropfl, W., & Julesz, B. (1982). The existence and role of retinotopic and spatiotopic forms of visual persistence. *Acta psychologica*, 52(3), 175–196. [https://doi.org/https://doi.org/10.1016/0001-6918\(82\)90007-5](https://doi.org/https://doi.org/10.1016/0001-6918(82)90007-5)
- Bridgeman, B. (1983). Mechanisms of space constancy. In A. Hein & M. Jeannerod (Eds.), *Spatially oriented behavior* (pp. 263–279). Springer New York. https://doi.org/10.1007/978-1-4612-5488-1_15
- Bridgeman, B., Van der Heijden, A., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, 17(2), 247–258. <https://doi.org/10.1017/S0140525X00034361>
- Bufacchi, R. J., & Iannetti, G. D. (2018). An action field theory of peripersonal space. *Trends in Cognitive Sciences*, 22(12), 1076–1090. <https://doi.org/10.1016/j.tics.2018.09.004>

- Büttner, U., & Kremmyda, O. (2007). Smooth pursuit eye movements and optokinetic nystagmus. *40*. <https://doi.org/10.1159/000100350>
- Calvert, G., Spence, C., Stein, B. E., et al. (2004). *The handbook of multi-sensory processes*. MIT press.
- Castet, E., Jeanjean, S., & Masson, G. S. (2002). Motion perception of saccade-induced retinal translation. *Proceedings of the National Academy of Sciences*, *99*(23), 15159–15163. <https://doi.org/10.1073/pnas.232377199>
- Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, *17*(14), 6–6. <https://doi.org/10.1167/17.14.6>
- Cicchini, G. M., Mikellidou, K., & Burr, D. C. (2018). The functional role of serial dependence. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1890), 20181722. <https://doi.org/10.1098/rspb.2018.1722>
- Cohen, B., Dai, M., Yakushin, S. B., & Cho, C. (2019). The neural basis of motion sickness. *Journal of neurophysiology*, *121*(3), 973–982. <https://doi.org/10.1152/jn.00674.2018>
- Cohen, Y. E., & Andersen, R. A. (2000). Reaches to sounds encoded in an eye-centered reference frame. *Neuron*, *27*(3), 647–652. [https://doi.org/10.1016/S0896-6273\(00\)00073-8](https://doi.org/10.1016/S0896-6273(00)00073-8)
- Colby, C. L., & Duhamel, J.-R. (1996). Spatial representations for action in parietal cortex. *Cognitive Brain Research*, *5*(1), 105–115. [https://doi.org/https://doi.org/10.1016/S0926-6410\(96\)00046-8](https://doi.org/https://doi.org/10.1016/S0926-6410(96)00046-8)
- Cont, C., & Zimmermann, E. (2021). The motor representation of sensory experience. *Current Biology*, *31*(5), 1029–1036.e2. <https://doi.org/10.1016/j.cub.2020.11.032>
- Crowell, J. A., Banks, M. S., Shenoy, K. V., & Andersen, R. A. (1998). Visual self-motion perception during head turns. *Nature neuroscience*, *1*(8), 732–737. <https://doi.org/10.1038/3732>
- Cullen, K. E. (2019). Vestibular processing during natural self-motion: Implications for perception and action. *Nature Reviews Neuroscience*, *20*(6), 346–363. <https://doi.org/10.1038/s41583-019-0153-1>

- Di Stasi, L. L., McCamy, M. B., Catena, A., Macknik, S. L., Cañas, J. J., & Martinez-Conde, S. (2013). Microsaccade and drift dynamics reflect mental fatigue. *European Journal of Neuroscience*, *38*(3), 2389–2398. <https://doi.org/10.1111/ejn.12248>
- Dodge, R. (1903). Five types of eye movement in the horizontal meridian plane of the field of regard. *American journal of physiology-legacy content*, *8*(4), 307–329. <https://doi.org/10.1152/ajplegacy.1903.8.4.307>
- 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*(5040), 90–92. <https://doi.org/10.1126/science.1553535>
- Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science*, *317*(5841), 1048–1048. <https://doi.org/10.1126/science.1142175>
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. <https://doi.org/10.1038/415429a>
- Ferris, S. H. (1972). Motion parallax and absolute distance. *Journal of experimental psychology*, *95*(2), 258. <https://doi.org/10.1037/h0033605>
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature neuroscience*, *17*(5), 738–743. <https://doi.org/10.1038/nn.3689>
- Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and apparent distance. *Perception*, *17*(5), 609–621. <https://doi.org/10.1068/p170609>
- Fracasso, A., Caramazza, A., & Melcher, D. (2010). Continuous perception of motion and shape across saccadic eye movements. *J Vis*, *10*(13), 14. <https://doi.org/10.1167/10.13.14>
- Freedman, E. G., Stanford, T. R., & Sparks, D. L. (1996). Combined eye-head gaze shifts produced by electrical stimulation of the superior colliculus in rhesus monkeys. *Journal of Neurophysiology*, *76*(2), 927–952. <https://doi.org/10.1152/jn.1996.76.2.927>

- Freedman, E. G., & Sparks, D. L. (1997). Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys. *Journal of Neurophysiology*, 77(5), 2328–2348. <https://doi.org/10.1152/jn.1997.77.5.2328>
- Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, 27(4), 590–595. <https://doi.org/10.1016/j.cub.2017.01.006>
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, 22(3-4), 455–479. <https://doi.org/10.1080/02643290442000310>
- Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in cognitive sciences*, 7, 460–7. <https://doi.org/10.1016/j.tics.2003.08.014>
- Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. *Psychology of learning and motivation*, 43, 93–126.
- Goldberg, M. E., & Bruce, C. J. (1990). Primate frontal eye fields. iii. maintenance of a spatially accurate saccade signal. *Journal of neurophysiology*, 64(2), 489–508. <https://doi.org/10.1152/jn.1990.64.2.489>
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391(6666), 481–484. <https://doi.org/10.1038/35135>
- Gruesser, O.-J. (1986). Interaction of efferent and afferent signals in visual perception a history of ideas and experimental paradigms. *Acta Psychologica*, 63(1), 3–21. [https://doi.org/10.1016/0001-6918\(86\)90039-9](https://doi.org/10.1016/0001-6918(86)90039-9)
- Haarmeier, T., Bunjes, F., Lindner, A., Berret, E., & Thier, P. (2001). Optimizing visual motion perception during eye movements. *Neuron*, 32(3), 527–535. [https://doi.org/10.1016/S0896-6273\(01\)00486-X](https://doi.org/10.1016/S0896-6273(01)00486-X)
- Hall, N. J., & Colby, C. L. (2011). Remapping for visual stability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1564), 528–539. <https://doi.org/10.1098/rstb.2010.0248>
- Harris, C. M., & Wolpert, D. M. (2006). The main sequence of saccades optimizes speed-accuracy trade-off. *Biological Cybernetics*, 95(1), 21–29. <https://doi.org/10.1007/s00422-006-0064-x>

- Harris, L. R. (1994). Visual motion caused by movements of the eye, head and body. *Visual detection of motion*, 1, 397–435. <https://doi.org/10.1007/BF02521358>
- Henson, D. (1979). Investigation into corrective saccadic eye movements for refixation amplitudes of 10 degrees and below. *Vision Research*, 19(1), 57–61. [https://doi.org/10.1016/0042-6989\(79\)90121-4](https://doi.org/10.1016/0042-6989(79)90121-4)
- Hill, A. L. (1972). Direction constancy. *Perception & Psychophysics*, 11(2), 175–178. <https://doi.org/10.3758/BF03210370>
- Hillier, S., Immink, M., & Thewlis, D. (2015). Assessing proprioception: A systematic review of possibilities. *Neurorehabilitation and Neural Repair*, 29(10), 933–949. <https://doi.org/10.1177/1545968315573055>
- Hoffmann, E. R., Chan, A. H., & Heung, P. (2017). Head rotation movement times. *Human factors*, 59(6), 986–994. <https://doi.org/10.1177/0018720817701000>
- Hohwy, J., & Paton, B. (2010). Explaining away the body: Experiences of supernaturally caused touch and touch on non-hand objects within the rubber hand illusion. *PLOS ONE*, 5(2), 1–10. <https://doi.org/10.1371/journal.pone.0009416>
- Honda, H. (1991). The time courses of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. *Vision Research*, 31(11), 1915–1921. [https://doi.org/10.1016/0042-6989\(91\)90186-9](https://doi.org/10.1016/0042-6989(91)90186-9)
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1), 106. <https://doi.org/10.1113/jphysiol.1962.sp006837>
- Ionta, S., Gassert, R., & Blanke, O. (2011). Multi-sensory and sensorimotor foundation of bodily self-consciousness – an interdisciplinary approach. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00383>
- Irwin, D. E., McConkie, G. W., Carlson-Radvansky, L. A., & Currie, C. (1994). A localist evaluation solution for visual stability across saccades. *Behavioral and Brain Sciences*, 17(2), 265–266. <https://doi.org/10.1017/S0140525X00034439>

- Johnson, C. A., Wall, M., & Thompson, H. S. (2011). A history of perimetry and visual field testing. *Optometry and Vision Science*, 88(1). <https://doi.org/10.1097/OPX.0b013e3182004c3b>
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision research*, 31(7-8), 1351–1360. [https://doi.org/10.1016/0042-6989\(91\)90056-B](https://doi.org/10.1016/0042-6989(91)90056-B)
- Keenaghan, S., Bowles, L., Crawford, G., Thurlbeck, S., Kentridge, R. W., & Cowie, D. (2020). My body until proven otherwise: Exploring the time course of the full body illusion. *Consciousness and Cognition*, 78, 102882. <https://doi.org/10.1016/j.concog.2020.102882>
- Khan, S., & Chang, R. (2013). Anatomy of the vestibular system: A review [3]. *NeuroRehabilitation*, 32, 437–443. <https://doi.org/10.3233/NRE-130866>
- Kusunoki, M., Gottlieb, J., & Goldberg, M. E. (2000). The lateral intraparietal area as a salience map: The representation of abrupt onset, stimulus motion, and task relevance. *Vision research*, 40(10-12), 1459–1468. [https://doi.org/10.1016/S0042-6989\(99\)00212-6](https://doi.org/10.1016/S0042-6989(99)00212-6)
- Laeng, B., & Endestad, T. (2012). Bright illusions reduce the eye's pupil. *Proceedings of the National Academy of Sciences*, 109(6), 2162–2167. <https://doi.org/10.1073/pnas.1118298109>
- Lee, e. a., S. C. (2011). *Vestibular system anatomy*. Retrieved August 30, 2012, from emedicine.medscape.com/article/883956-%20overview%20aw2aab6c10
- Lenggenhager, B., Halje, P., & Blanke, O. (2011). Alpha band oscillations correlate with illusory self-location induced by virtual reality. *European Journal of Neuroscience*, 33(10), 1935–1943. <https://doi.org/10.1111/j.1460-9568.2011.07647.x>
- Lenggenhager, B., Mouthon, M., & Blanke, O. (2009). Spatial aspects of bodily self-consciousness. *Consciousness and Cognition*, 18(1), 110–117. <https://doi.org/10.1016/j.concog.2008.11.003>
- Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: Manipulating bodily self-consciousness. *Science*, 317(5841), 1096–1099. <https://doi.org/10.1126/science.1143439>

- Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, *24*(21), 2569–2574. <https://doi.org/10.1016/j.cub.2014.09.025>
- Linkenauger, S. A., Bülthoff, H. H., & Mohler, B. J. (2015). Virtual arm's reach influences perceived distances but only after experience reaching. *Neuropsychologia*, *70*, 393–401. <https://doi.org/10.1016/j.neuropsychologia.2014.10.034>
- Lisberger, S. (1984). The latency of pathways containing the site of motor learning in the monkey vestibulo-ocular reflex. *Science*, *225*(4657), 74–76. <https://doi.org/10.1126/science.6610214>
- Longo, M. R., & Lourenco, S. F. (2007). Space perception and body morphology: Extent of near space scales with arm length. *Experimental Brain Research*, *177*(2), 285–290. <https://doi.org/10.1007/s00221-007-0855-x>
- Luna, R., Serrano-Pedraza, I., Gegenfurtner, K. R., Schütz, A. C., & Souto, D. (2021). Achieving visual stability during smooth pursuit eye movements: Directional and confidence judgements favor a recalibration model. *Vision Research*, *184*, 58–73. <https://doi.org/10.1016/j.visres.2021.03.003>
- Macaluso, E., Noppeney, U., Talsma, D., Vercillo, T., Hartcher-O'Brien, J., & Adam, R. (2016). The curious incident of attention in multisensory integration: Bottom-up vs. top-down. *Multisensory Research*, *29*(6-7), 557–583. <https://doi.org/10.1163/22134808-00002528>
- Manassi, M., Liberman, A., Kosovicheva, A., Zhang, K., & Whitney, D. (2018). Serial dependence in position occurs at the time of perception. *Psychonomic Bulletin & Review*, *25*(6), 2245–2253. <https://doi.org/10.3758/s13423-018-1454-5>
- Manassi, M., & Whitney, D. (2022). Illusion of visual stability through active perceptual serial dependence. *Science Advances*, *8*(2), eabk2480. <https://doi.org/10.1126/sciadv.abk2480>
- Medendorp, W. P., Goltz, H. C., Vilis, T., & Crawford, J. D. (2003). Gaze-centered updating of visual space in human parietal cortex. *Journal of Neuroscience*, *23*(15), 6209–6214. <https://doi.org/10.1523/JNEUROSCI.23-15-06209.2003>

- Melcher, D. (2007). Predictive remapping of visual features precedes saccadic eye movements. *Nature Neuroscience*, 10(7), 903–907. <https://doi.org/10.1038/nn1917>
- Mergner, T., Rottler, G., Kimmig, H., & Becker, W. (1992). Role of vestibular and neck inputs for the perception of object motion in space. *Experimental Brain Research*, 89(3), 655–668. <https://doi.org/10.1007/BF00229890>
- Merriam, E. P., Genovese, C. R., & Colby, C. L. (2003). Spatial updating in human parietal cortex. *Neuron*, 39(2), 361–373. [https://doi.org/10.1016/S0896-6273\(03\)00393-3](https://doi.org/10.1016/S0896-6273(03)00393-3)
- Merriam, E. P., Genovese, C. R., & Colby, C. L. (2007). Remapping in human visual cortex. *Journal of Neurophysiology*, 97(2), 1738–1755. <https://doi.org/10.1152/jn.00189.2006>
- Money, K. E. (1970). Motion sickness. *Physiological reviews*, 50(1), 1–39.
- Monjo, F., Terrier, R., & Forestier, N. (2015). Muscle fatigue as an investigative tool in motor control: A review with new insights on internal models and posture–movement coordination. *Human Movement Science*, 44, 225–233. <https://doi.org/10.1016/j.humov.2015.09.006>
- Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences*, 99(6), 4026–4031. <https://doi.org/10.1073/pnas.052379899>
- Nakul, E., Orlando-Dessaints, N., Lenggenhager, B., & Lopez, C. (2020). Measuring perceived self-location in virtual reality. *Scientific reports*, 10, 6802. <https://doi.org/10.1038/s41598-020-63643-y>
- Nicolle, J. M.-T. P.-D., Servos, P., & GOOdale, M. (1995). Adapting to monocular vision: Grasping with one eye. *Exp Brain Res*, 104, 107–114. <https://doi.org/10.1007/BF00229860>
- Niessen, M. H. M., Veeger, D. H. E. J., & Janssen, T. W. J. (2009). Effect of body orientation on proprioception during active and passive motions. *American Journal of Physical Medicine & Rehabilitation*, 88(12). <https://doi.org/10.1097/PHM.0b013e3181c1eac1>

- Noel, J.-P., Pfeiffer, C., Blanke, O., & Serino, A. (2015). Peripersonal space as the space of the bodily self. *Cognition*, *144*, 49–57. <https://doi.org/10.1016/j.cognition.2015.07.012>
- Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (1996). The visual perception of three-dimensional length. *Journal of experimental psychology: Human Perception and Performance*, *22*(1), 173. <https://doi.org/10.1037/0096-1523.22.1.173>
- Ogard, W. K. (2011). Proprioception in sports medicine and athletic conditioning. *Strength & Conditioning Journal*, *33*(3). <https://doi.org/10.1519/SSC.0b013e31821bf3ae>
- Ogawa, N., Narumi, T., & Hirose, M. (2018). Object size perception in immersive virtual reality: Avatar realism affects the way we perceive. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 647–648. <https://doi.org/10.1109/VR.2018.8446318>
- Palluel, E., Aspell, J. E., & Blanke, O. (2011). Leg muscle vibration modulates bodily self-consciousness: Integration of proprioceptive, visual, and tactile signals. *Journal of Neurophysiology*, *105*(5), 2239–2247. <https://doi.org/10.1152/jn.00744.2010>
- Pfeiffer, C., Schmutz, V., & Blanke, O. (2014). Visuospatial viewpoint manipulation during full-body illusion modulates subjective first-person perspective. *Exp Brain Res*, *232*, 4021–4033. <https://doi.org/10.1007/s00221-014-4080-0>
- Pouget, A., Ducom, J.-C., Torri, J., & Bavelier, D. (2002). Multisensory spatial representations in eye-centered coordinates for reaching. *Cognition*, *83*, B1–B11. [https://doi.org/10.1016/S0010-0277\(01\)00163-9](https://doi.org/10.1016/S0010-0277(01)00163-9)
- Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on psychological science*, *1*(2), 110–122. <https://doi.org/10.1111/j.1745-6916.2006.00008.x>
- Proffitt, D. R., Bhalla, M., Gossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic bulletin & review*, *2*, 409–428. <https://doi.org/10.3758/BF03210980>
- Proske, U. (2006). Kinesthesia: The role of muscle receptors. *Muscle & Nerve*, *34*(5), 545–558. <https://doi.org/10.1002/mus.20627>

- Proske, U., & Gandevia, S. C. (2012). The proprioceptive senses: Their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews*. <https://doi.org/10.1152/physrev.00048.2011>
- Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. MIT press.
- Rineau, A.-L., Bringoux, L., Sarrazin, J.-C., & Berberian, B. (2023). Being active over one's own motion: Considering predictive mechanisms in self-motion perception. *Neuroscience & Biobehavioral Reviews*, 105051. <https://doi.org/10.1016/j.neubiorev.2023.105051>
- Rizzolatti, G., Scandolara, C., Matelli, M., & Gentilucci, M. (1981). Afferent properties of periarculate neurons in macaque monkeys. ii. visual responses. *Behavioural Brain Research*, 2(2), 147–163. [https://doi.org/10.1016/0166-4328\(81\)90053-X](https://doi.org/10.1016/0166-4328(81)90053-X)
- Robinson, D. (1975). Oculomotor control signals. basic mechanisms of ocular motility and their clinical implications eds p. bach-y-rita and g. lennerstrand.
- Robinson, D. A. (1973). Models of the saccadic eye movement control system. *Kybernetik*, 14, 71–83. <https://doi.org/10.1007/BF00288906>
- Rock, I., & Ebenholtz, S. (1962). Stroboscopic movement based on change of phenomenal rather than retinal location. *The American journal of psychology*, 75(2), 193–207. <https://doi.org/10.2307/1419603>
- Ross, J., Morrone, M., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, 24(2), 113–121. [https://doi.org/10.1016/S0166-2236\(00\)01685-4](https://doi.org/10.1016/S0166-2236(00)01685-4)
- Salomon, R., Lim, M., Pfeiffer, C., Gassert, R., & Blanke, O. (2013). Full body illusion is associated with widespread skin temperature reduction. *Frontiers in Behavioral Neuroscience*, 7. <https://doi.org/10.3389/fnbeh.2013.00065>
- Scotland, S., Adamo, D. E., & Martin, B. J. (2014). Sense of effort revisited: Relative contributions of sensory feedback and efferent copy. *Neuroscience Letters*, 561, 208–212. <https://doi.org/10.1016/j.neulet.2013.12.041>

- Servos, P., Goodale, M. A., & Jakobson, L. S. (1992). The role of binocular vision in prehension: A kinematic analysis. *Vision Research*, 32(8), 1513–1521. [https://doi.org/10.1016/0042-6989\(92\)90207-Y](https://doi.org/10.1016/0042-6989(92)90207-Y)
- Sidenmark, L., & Gellersen, H.-W. (2019). Eye, head and torso coordination during gaze shifts in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27, 1–40. <https://doi.org/10.1145/3361218>
- Sommer, M. A., & Wurtz, R. H. (2004). What the brain stem tells the frontal cortex. ii. role of the sc-md-fef pathway in corollary discharge. *Journal of neurophysiology*, 91(3), 1403–1423. <https://doi.org/10.1152/jn.00740.2003>
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of comparative and physiological psychology*, 43(6), 482. <https://doi.org/10.1037/h0055479>
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4), 255–266. <https://doi.org/10.1038/nrn2331>
- Szinte, M., & Cavanagh, P. (2011). Spatiotopic apparent motion reveals local variations in space constancy. *Journal of Vision*, 11(2), 4–4. <https://doi.org/10.1167/11.2.4>
- Taubert, J., & Alais, D. (2016). Serial dependence in face attractiveness judgements tolerates rotations around the yaw axis but not the roll axis. *Visual Cognition*, 24(2), 103–114. <https://doi.org/10.1080/13506285.2016.1196803>
- Tuthill, J. C., & Azim, E. (2018). Proprioception. *Current Biology*, 28(5), R194–R203. <https://doi.org/10.1016/j.cub.2018.01.064>
- van der Steen, F. A. M. (1998). An earth-stationary perceived visual scene during roll and yaw motions in a flight simulator [6]. *Journal of Vestibular Research*, 8, 411–425. <https://doi.org/10.3233/VES-1998-8602>
- van der Hoort, B., Guterstam, A., & Ehrsson, H. H. (2011). Being barbie: The size of one’s own body determines the perceived size of the world. *PLOS ONE*, 6(5), 1–10. <https://doi.org/10.1371/journal.pone.0020195>

- van Beers, R. J., Haggard, P., & Wolpert, D. M. (2004). The role of execution noise in movement variability. *Journal of Neurophysiology*, *91*(2), 1050–1063. <https://doi.org/10.1152/jn.00652.2003>
- Vetter, P., Goodbody, S. J., & Wolpert, D. M. (1999). Evidence for an eye-centered spherical representation of the visuomotor map. *Journal of Neurophysiology*, *81*(2), 935–939. <https://doi.org/10.1152/jn.1999.81.2.935>
- Vignemont, F. D., Ehrsson, H., & Haggard, P. (2005). Bodily illusions modulate tactile perception. *Current Biology*, *15*(14), 1286–1290. <https://doi.org/10.1016/j.cub.2005.06.067>
- Volcic, R., Fantoni, C., Caudek, C., Assad, J. A., & Domini, F. (2013). Visuo-motor adaptation changes stereoscopic depth perception and tactile discrimination. *Journal of Neuroscience*, *33*(43), 17081–17088. <https://doi.org/10.1523/JNEUROSCI.2936-13.2013>
- von Holst, E., & Mittelstaedt, H. (1950). Das reafferenzprinzip. *Naturwissenschaften*, *37*(20), 464–476. <https://doi.org/10.1007/BF00622503>
- Wald, G. (1945). Human vision and the spectrum. *Science*, *101*(2635), 653–658. <https://doi.org/10.1126/science.101.2635.653>
- Wallach, H. (1985). Learned stimulation in space and motion perception. *American Psychologist*, *40*, 399–404. <https://doi.org/10.1037/0003-066X.40.4.399>
- Wallach, H. (1987). Perceiving a stable environment when one moves. *Annual review of psychology*, *38*(1), 1–29. <https://doi.org/10.1146/annurev.ps.38.020187.000245>
- Walton, M. M. G., Bechara, B., & Gandhi, N. J. (2008). Effect of reversible inactivation of superior colliculus on head movements. *Journal of Neurophysiology*, *99*(5), 2479–2495. <https://doi.org/10.1152/jn.01112.2007>
- Warren, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 683–703. <https://doi.org/10.1037/0096-1523.10.5.683>
- Wertheim, A. H. (1994). Motion perception during selfmotion: The direct versus inferential controversy revisited. *Behavioral and Brain Sciences*, *17*(2), 293–311. <https://doi.org/10.1017/S0140525X00034646>

- Wexler, M. (2003). Voluntary head movement and allocentric perception of space. *Psychological Science*, 14(4), 340–346. <https://doi.org/10.1111/1467-9280.14491>
- Wiesing, M., Kartashova, T., & Zimmermann, E. (2021). Adaptation of pointing and visual localization in depth around the natural grasping distance. *Journal of Neurophysiology*, 125(6), 2206–2218. <https://doi.org/10.1152/jn.00012.2021>
- Wiesing, M., & Zimmermann, E. (2023). Serial dependencies between locomotion and visual space. *Scientific Reports*, 13(1), 3302. <https://doi.org/10.1038/s41598-023-30265-z>
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*, 9, 625–636. <https://doi.org/10.3758/BF03196322>
- Wurtz, R. H. (2018). Corollary discharge contributions to perceptual continuity across saccades. *Annual review of vision science*, 4, 215–237. <https://doi.org/10.1146/annurev-vision-102016-061207>
- Zangemeister, W. H., Lehman, S., & Stark, L. (1981). Simulation of head movement trajectories: Model and fit to main sequence. *Biological Cybernetics*, 41(1), 19–32. <https://doi.org/10.1007/BF01836124>
- Zimmermann, E. (2021). Sensorimotor serial dependencies in head movements. *Journal of Neurophysiology*, 126(3), 913–923. <https://doi.org/10.1152/jn.00231.2021>

Affidavit

Eidesstattliche Erklärung gemäß § 5 der Promotionsordnung vom 15.06.2018 der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf:

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorliegenden oder ähnlichen Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den

04.10.2023

Datum



Manuel Bayer

Acknowledgements

I want to thank my supervisor Prof. Dr. Eckart Zimmermann for providing me with a never ending stream of new research ideas and the support throughout all the projects. I also want to thank my mentor Prof. Dr. Christian Bellebaum for all the pieces of advice and support. I am very grateful for all lab members and student assistants I had the opportunity to work with, with special thanks going to Dr. Antonella Pomé, Clara Fritz and Sandra Tyralla. I immensely enjoyed working on projects and solving daily tasks together. I want to thank Dr. Benjamin Weissmüller for erasing the weakling inside of me by punching me in the face repeatedly. A very special thanks to Dr. Laura Bechtold, who not only saved me during my bachelor but exposed me to the world of science, paving the way for my whole career. I also want to thank my girlfriend Lena Leusch for all the emotional support she has given me throughout my studies. My utmost gratitude goes to my mother who raised me pretty much on her own while suffering from countless strokes of fate and hardships. I cannot put in words how proud I am to be your son and how thankful I am for all that you have given up for me to offer me the life I had.

Appendix

Original article of study 1

Bayer, M., Betka, S., Herbelin, B., Blanke, O., & Zimmermann, E. (2023). The full-body illusion changes visual depth perception. *Scientific Reports*, 13(1), 10569. <https://doi.org/10.1038/s41598-023-37715-8>

I was the main author of this article. I contributed to the development of the paradigm, created the experimental setup, conducted the data acquisition and analysis, contributed to the interpretation of the results, wrote the initial draft of the manuscript and cooperated with all co-authors on the following versions of the manuscript.

Original article of study 2

Bayer, M., & Zimmermann, E. (2023). Serial dependencies in visual stability during self-motion. *Journal of Neurophysiology*, 130 (2), 447–457. <https://doi.org/10.1152/jn.00157.20231051>

I was the main author of this article. I contributed to the development of the paradigm, created the experimental setup, supervised the data acquisition, conducted the data analysis, contributed to the interpretation of the results, wrote the initial draft of the manuscript and cooperated with the co-author on the following versions of the manuscript.

Manuscript of study 3

Bayer, M., & Zimmermann, E. (2023). Head movements contribute to spatial updating across gaze shifts. *Manuscript submitted for publication in Proceeding of the royal society B: biological sciences.*

I was the main author of this article. I contributed to the development of the paradigm, created the experimental setup, supervised the data acquisition, conducted the data analysis, contributed to the interpretation of the results, wrote the initial draft of the manuscript and cooperated with the co-author on the following versions of the manuscript.



OPEN The full-body illusion changes visual depth perception

Manuel Bayer^{1✉}, Sophie Betka², Bruno Herbelin², Olaf Blanke^{2,3} & Eckart Zimmermann¹

Knowing where objects are relative to us implies knowing where we are relative to the external world. Here, we investigated whether space perception can be influenced by an experimentally induced change in perceived self-location. To dissociate real and apparent body positions, we used the full-body illusion. In this illusion, participants see a distant avatar being stroked in virtual reality while their own physical back is simultaneously stroked. After experiencing the discrepancy between the seen and the felt location of the stroking, participants report a forward drift in self-location toward the avatar. We wondered whether this illusion-induced forward drift in self-location would affect where we perceive objects in depth. We applied a psychometric measurement in which participants compared the position of a probe against a reference sphere in a two-alternative forced choice task. We found a significant improvement in task performance for the right visual field, indicated by lower just-noticeable differences, i.e., participants were better at judging the differences of the two spheres in depth. Our results suggest that the full-body illusion is able to facilitate depth perception at least unilaterally, implying that depth perception is influenced by perceived self-location.

Our perception of depth is constructed from monocular and binocular cues with a preponderance of the latter in natural vision^{1,2}. Retrieving spatial information from binocular disparity, i.e., the difference in the retinal projection of objects on the left and right eye, is inherently ambiguous^{3,4}. A small but close object can generate the same retinal projection as a bigger object located far away. In order to resolve this ambiguity, the brain might rely on sensorimotor knowledge like the distance of locomotion required to reach an object or the size of an arm movement that would be necessary to grasp the object^{5,6}. The coordinates of a movement can be used to interpret the visuospatial location of the corresponding object⁷. If spatial perception is calibrated by action, then the perceived location of objects should also depend on our perceived location in space.

A convenient experimental tool to dissociate between the subjective feeling of where we are and the physical location of our body is provided by the full-body illusion^{8,9}. The illusion can be created by presenting a video stream in a head-mounted display, which shows the participants' back being stroked by the experimenter, while the physical back of the participants is also stroked synchronously⁹. The discrepancy between the visual and the tactile location of stroking produces a shift in the perceived self-location of the observer, as if participants were standing in front of their physical body. In other versions of the full-body illusion, instead of the participants' back, an avatar is shown¹⁰. This induction of the full-body illusion is then usually compared to an asynchronous condition with a delay between the felt and visual stroking⁹. The delay between the visual stroking on the avatar's back and the felt stroking on the participant's back should impede the induction of the full-body illusion while still providing an identical visual stimulus.

Full-body illusions are descendants of the previously discovered rubber hand illusion¹¹. Both illusions are driven by a conflict between senses, forcing the brain to adopt a compromise between the diverging information sources. The resulting perception follows the principles of multisensory integration, according to which the more uncertain sense is biased toward the more certain one¹². In full-body illusions three senses are involved, including vision, proprioception and touch¹³. Since the variance in somatosensory signals is higher than in vision, participants feel like standing where they see the avatar in the full-body illusion, thus shifting self-location from an embodied self-location (centered on the upper body; trunk and/or face, as tested without the full body illusion¹⁴) towards the position of the avatar^{9,10,15–17}. Multisensory integration during the full-body illusion might be implemented by trimodal neurons, which are activated by the combined presence of three sensory signals^{13,18,19}. Indeed, imaging studies have shown that full-body illusions involve bilateral premotor cortex, intraparietal sulcus and sensorimotor cortex^{10,13,15}.

¹Department of Experimental Psychology, Heinrich-Heine-University, Düsseldorf, Germany. ²Laboratory of Cognitive Neuroscience, NeuroX Institute & Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland. ³Department of Clinical Neuroscience, Geneva University Hospital, Geneva, Switzerland. ✉email: manuel.bayer@hhu.de

Furthermore, electrophysiological studies found that PMC and IPS host tri- and bimodal neurons with visual and somatosensory receptive fields in the arms and the trunk (for a review, see¹³). These trunk-centered receptive fields cover the whole body²⁰ and are therefore well suited to bring forth the full-body illusion. Illusory self-identification with a virtual body is also associated with physiological and nociceptive changes; for instance, the skin conductance responds to a threat directed towards the virtual body²¹. The changes in touch, pain perception and physiology that occur during illusory self-identification indicate that states of illusory self-identification alter the way humans process stimuli from their body^{22–24}.

Full-body illusions can be estimated by subjective measures, i.e., questionnaires, and by objective measures^{17,25,26}. In questionnaires, participants report that they locate themselves closer to or at the position of the visual body or avatar. Several objective measures have provided evidence that, after the induction of the full-body illusion, participants estimate themselves to be at a different location in space than their physical body^{26,27}. This shift of the perceived self-location was measured by changes in behavior or by perceptual effects which reflect the felt position of the body. An instance of an objective measure is a mental imagery task (i.e., mental ball dropping task). Lenggenhager, Mouthon and Blanke²⁸ instructed their participants to imagine dropping a ball and to indicate when they think it reached the ground while being in a prone position on a bench. The estimated time before participants indicated that the ball reached the ground was shorter after the induction of the full-body illusion. In a comparable study in which a ball is approaching the participant, Nakul et al.²⁶ demonstrated that participants judged the arrival of the ball earlier after the induction of the full-body illusion, compatible with a forward drift in self-location. Several other studies used blind walking tasks, during which participants were displaced and had to walk back to their initial position. Under the full-body illusion participants tended to overshoot their position, consistent with the experience of being displaced forward in space^{9,10,17}.

Another approach to quantify the effects of the full-body illusion is the measurement of the shift of the peripersonal space, i.e., the space in which we integrate multisensory body-related signals around the body²⁹. For instance, Noel et al.²⁹ presented a looming sound and a tactile vibration to the participants, who were asked to respond as soon as they perceive the tactile stimulation. This procedure enables the measurement of the boundaries of the peripersonal space by presenting stimuli in the front and the back and makes it possible to measure the peripersonal space shifts at several distances between the real physical body of the participant and the full-body illusion related avatar.

In order to investigate if the experimentally induced drift in self-location during the full-body illusion alters depth perception we used a psychometric measurement. In this visual depth task, two spheres were presented repeatedly immediately after illusion induction. Participants were asked to judge which of the spheres was closer to them. Since depth perception is more accurate within the peripersonal space³⁰ and more accurate for closer objects in general, we expected the experimentally induced self-location drift to influence the participants' task performance. If the peripersonal space drifts toward the avatar and therefore also toward the stimuli, participants might be able to judge the distance of the stimuli more precisely, reflected by lower just-noticeable differences (JNDs). As the task only requires the participants to judge a relative distance, we did not expect to find any effect on absolute perception, which would manifest itself in point of subjective equality (PSE) differences.

Results

Questionnaire. Participants performed a questionnaire to quantify the intensity of the illusion. This questionnaire comprised five items, covering the following aspects of the full-body illusion: perceived self-location, self-identification and illusory touch experiences²². There were significant differences between the synchronous and asynchronous condition for the second (self-identification) and fourth item (illusory touch experience) of the questionnaire (item 2: $t(19) = 2.04$, $p = 0.028$, item 4: $t(19) = 1.78$, $p = 0.045$, see Fig. 1). In both cases, participants were showing more agreement to these statements in the synchronous compared to the asynchronous condition, in line with previous results^{9,17,22}. This result confirms that the present setup was able to successfully induce the full-body illusion. For the other items we did not find any significant differences between the synchronous and asynchronous condition (item 1: $t(19) = -0.53$, $p = 0.700$, item 3: $t(19) = 0.10$, $p = 0.459$, item 5 failed the test for normality, we therefore opted to perform a wilcoxon signed-rank test: $W = 53.5$, $p = 0.475$).

Based on the responses in the questionnaire we computed the average response across all items for each participant. In this calculation item 5 was mirrored, as more agreement indicated a weaker illusion. Based on this average response we split the participants in a high quotient ($M = 1.25$, $SD = 2.41$) and low quotient group ($M = -5.02$, $SD = 4.11$) via a median split. This split was performed to account for the variance in the response to the full-body illusion. Participants with a strong full body illusion (questionnaire quotient ≥ -1.00) were classified as high quotient and participants with a weak full body illusion (questionnaire quotient < -1.00) were classified as low quotient.

Psychometric measures. To quantify the influence of the full-body illusion on visual depth discrimination, we compared JNDs from the synchronous versus asynchronous sessions. To capture the effect of the full-body illusion, we split the participants into two groups according to the subjective strength of the full-body illusion that was estimated by the questionnaire. Figure 2b shows JND differences in cm between the synchronous and asynchronous (asynchronous–synchronous) condition for all participants of the high quotient group for trials in which the probe was presented in the right visual field. If data points lie on the dashed line, there is no difference between the two conditions. One can see that all data points (except one) lay to the left of the dashed line, thus showing lower JNDs in the synchronous than in the asynchronous sessions. Across all participants we found a significant difference in JNDs between synchronous and asynchronous stroking sessions for the high quotient group in trials in which the probe sphere was presented in the right visual field (paired t -test, $t(10) = 4.16$, $p = 0.008$). Figure 2a shows JNDs for all participants from sessions in which the probe was presented in the

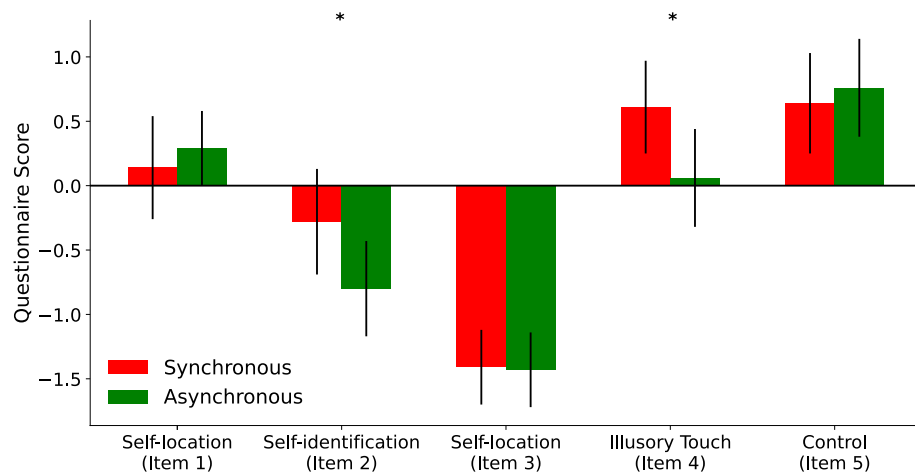


Figure 1. Mean scores of each full-body illusion questionnaire item for the synchronous (red) and asynchronous condition (green). Error bars indicate the standard error. There were significant differences between the two conditions for the second and fourth item. Item 2 asked the participants to what extent they felt as if the avatar they saw was their own and item 4 asked to what extent the stroking felt as if it was located on the avatar. There were no significant differences between the synchronous and asynchronous condition for the first, third and fifth item. Item 1 asked the participants to what extent they felt as if they were slightly above or below the seen avatar, item 3 asked the participants to what extent they felt as if their own body shifted towards the seen avatar and item 5 asked the participants to what extent they felt as if nothing changed. *Indicates $p < 0.05$.

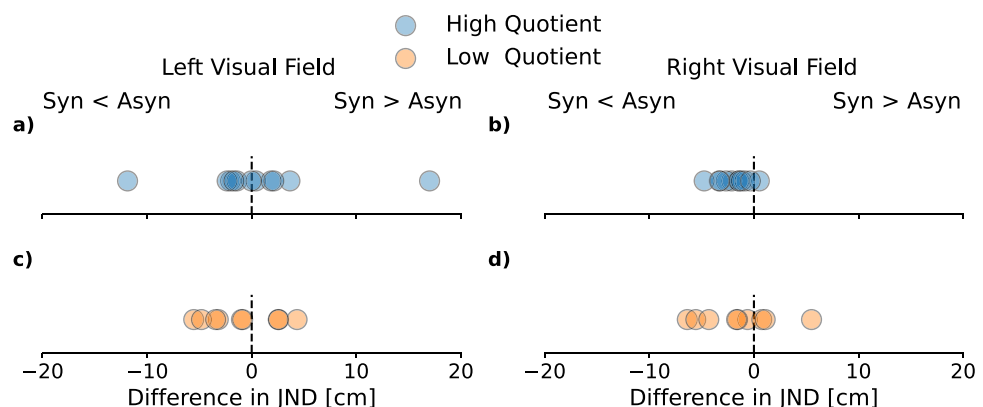


Figure 2. Precision of localization in the visual depth task, as quantified by JNDs. A lower JND indicates a higher sensitivity, i.e., higher precision in solving the task. Differences between the synchronous and asynchronous condition were split by the side of the visual field in which stimuli were presented (left/right) and the questionnaire group (low/high quotient). (a) JND differences for the high quotient group from trials in which the probe was presented in the left visual field. Data points represent single subject data. The dashed line indicates equality between the synchronous and asynchronous condition. Values to the left of the dashed line represent participants who had a lower JND in the synchronous (syn) compared to the asynchronous condition (asyn) and vice versa. (b) JND differences for the high quotient group from trials in which the probe was presented in the right visual field. One can see that all data points (except one) lay to the left of the dashed line, thus showing lower JNDs in the synchronous compared to the asynchronous condition. Across all participants we found a significant difference in JNDs between synchronous and asynchronous stroking sessions for the high quotient group in trials in which the probe was presented in the right visual field. (c) JND differences for the low quotient group in which the probe was presented in the left visual field. (d) JND differences for the low quotient group in which the probe was presented in the right visual field. Data are relatively evenly distributed around the dashed line, indicating no JND difference between the synchronous and asynchronous condition.

left visual field. In this condition, data points can be found on both sides of the dashed line. A paired t -test did not reveal a significant difference ($t(10) = -0.25$, $p = 0.999$).

No JND differences were observed for the low quotient group. Results for the low quotient group are shown in Fig. 2c and d. Data points lie to both sides of the dashed line for trials in which the probe was presented in the left or right visual field. Paired t -tests did not reveal a significant difference, neither for the left ($t(8) = 0.89$, $p = 0.999$), nor for the right visual field ($t(8) = 1.15$, $p = 0.999$).

To check if the full-body illusion biased depth perception we also analyzed PSEs, by comparing the synchronous against the asynchronous condition with the same median split as for the JNDs (Fig. 3). This analysis showed no significant difference between the PSE of the high quotient group for trials in which the probe was in the left ($t(10) = 0.05$, $p = 0.999$, see Fig. 3a) or in the right visual field ($t(10) = 0.25$, $p = 0.999$, see Fig. 3b). The same was the case for the low quotient group in the left ($t(8) = -0.18$, $p = 0.999$, see Fig. 3c) and in the right visual field ($t(8) = 0.14$, $p = 0.999$, see Fig. 3d).

These results are in line with our expectations as significant differences for the PSE would imply changes on the absolute depth perception of the participants. If the PSE of the synchronous condition was higher than the one of the asynchronous condition, participants would have perceived the probe as closer in depth after the induction of the full-body illusion.

Psychometric functions of example participants are shown in Fig. 4. Figure 4a and b represent one participant of the high quotient group and Fig. 4c and d one participant of the low quotient group. One can see that in Fig. 4b the red line is steeper than the green line, which indicates that the JND in the synchronous is lower than in the asynchronous condition in trials in which the probe was presented on the right. For the same participant this effect is less pronounced in trials in which the probe was presented on the left (see Fig. 4a). For the participant of the low quotient group there are little to no differences in steepness of the two curves for both sides of the visual field (see Fig. 4c,d).

Mental imagery task. Based on previous data the mental imagery task was carried out to quantify changes in self-location^{15,26}. We calculated the distance between the position of the participant and the ball at the time of the participant's response. Figure 5 shows the means of all participants for the synchronous and asynchronous condition split into the high quotient and low quotient group and the side of the visual field the ball was approaching the participants from. Positive values indicate that participants responded too early, i.e. the ball was still in front of them, while negative values indicate a response that was too late, i.e. the ball had already passed their position and was behind them. We would have expected positive and higher values in the synchronous compared to the asynchronous condition if participants experienced a drift in self-location during the full-body illusion. In some cases, the mean response of the participants led to positive values (see Fig. 5c), but the differences between the individual means was small compared to the variance in the data reflected by the size of the standard error. There was no significant difference for the high quotient group between the synchronous and asynchronous condition, neither for trials in which the ball approached from the left ($t(10) = -0.64$, $p = 0.999$, see Fig. 5a) nor in trials in which the ball approached from the right visual field ($t(10) = -0.55$, $p = 0.999$, see Fig. 5b). The same pattern applies to the low quotient group as there was no significant difference between the synchronous and asynchronous condition independently of the visual field the ball was approaching from (left: $t(8) = -0.63$, $p = 0.999$, see Fig. 5c, right: $t(8) = -1.44$, $p = 0.756$, see Fig. 5d).

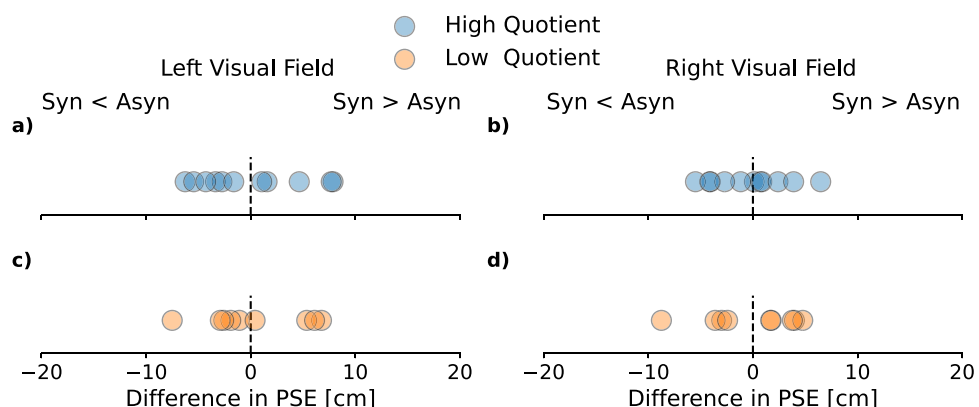


Figure 3. PSE differences between the synchronous and asynchronous condition split by the side of the visual field in which stimuli were presented (left/right) and the questionnaire group (low/high quotient). (a) PSE differences for the high quotient group from trials in which the probe was presented in the left visual field. Data points represent single subject data. The dashed line indicates equality between the synchronous and asynchronous condition. Values to the left of the dashed line represent participants who had a lower PSE in the synchronous (syn) compared to the asynchronous condition (asyn) while values to the right of the line show participants who had a higher PSE in the synchronous compared to the asynchronous condition. (b) PSE differences for the high quotient group from trials in which the probe was presented in the right visual field. (c) PSE differences for the low quotient group from trials in which the probe was presented in the left visual field. (d) PSE differences for the low quotient group from trials in which the probe was presented in the right visual field.

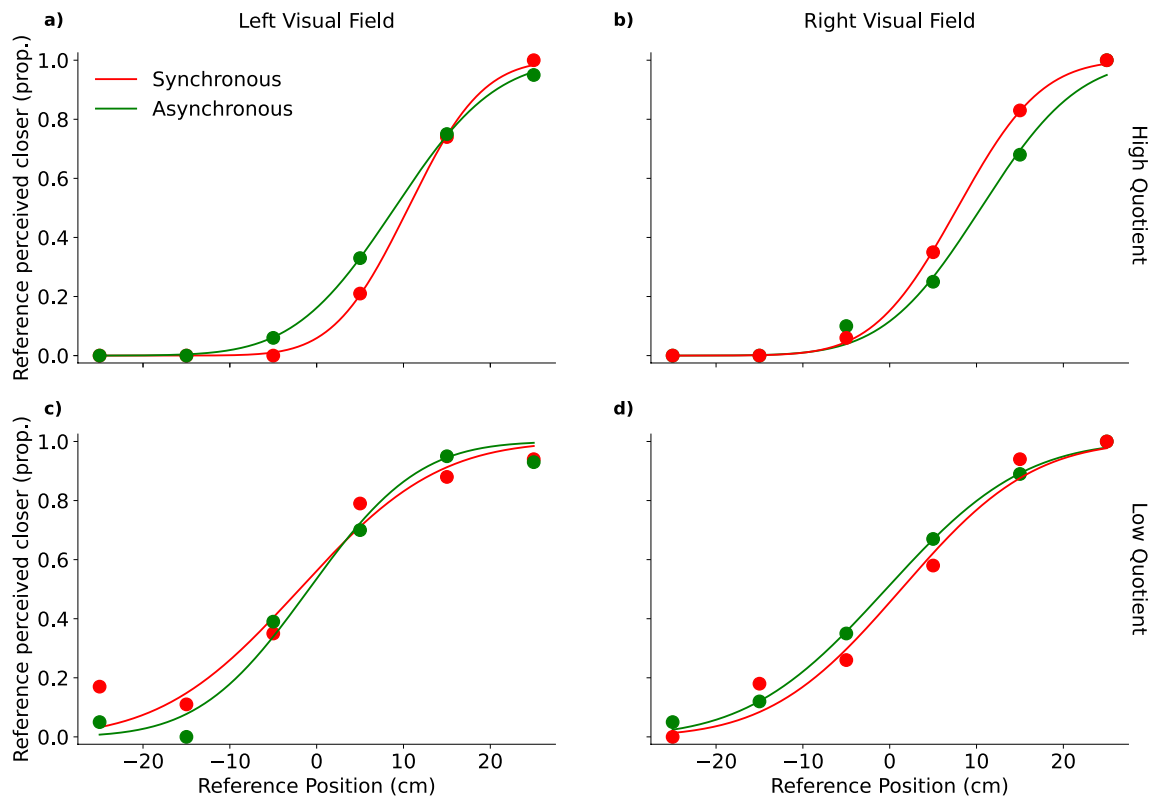


Figure 4. Representative psychometric functions of individual participants. (a) Data of a participant from the high quotient group. The ordinate indicates the proportion of cases the reference was chosen as the closer stimulus and the abscissa shows the position of the reference stimulus in relation to the probe stimulus. Higher values indicate that the reference stimulus was closer to the participant on the anterior–posterior axis and negative values indicate that the reference was further away from the participant than the probe stimulus. The lines represent the fitted cumulative gaussian functions for the synchronous (red) and asynchronous (green) condition and the dots represent the data points used for the fitting procedure. (b) One can see that the red curve is steeper than the green curve, indicating that this participant was more sensitive in the task after the induction of the full-body illusion in trials in which the probe was presented in the right visual field. (c, d) Little to no differences can be observed for this participant from the low questionnaire group when comparing the synchronous against the asynchronous condition for the JND.

Discussion

We investigated whether shifts of perceived self-location can influence depth perception. To manipulate perceived self-location, we induced the full-body illusion. Participants saw an avatar being stroked on its back and synchronously felt the same stroking movement on their own physical back. The full-body illusion caused participants to perceive themselves closer to the avatar's position due to a perceived self-location drift in the direction of the avatar^{9,10,17,26,28}.

After the full-body illusion was induced, participants compared the distance of two spheres that were presented in front of them. We measured the discrimination performance after the induction and compared it to the discrimination performance in an asynchronous condition where the full-body illusion should have not been induced. The discrimination sensitivity in this task was increased when participants experienced the full-body illusion, as if they would stand closer to the discrimination targets. Except for the synchronicity of the stroking, both conditions were identical. Any improvement in discrimination can thus only be related to the synchrony of stroking that generates the multisensory conditions for the full-body illusion. Moreover, the improvement in discrimination was only present for the group of participants which experienced the full-body illusion more intensively, further providing support for a direct link between the full-body illusion and the change in depth perception.

We used objective psychometric measurements which are far less susceptible to cognitive influences than subjective measures, like verbal reports³¹. However, psychometric measures are not free of higher-level cognitive biases. It has been shown that in experiments participants can deliberately shift the bias of a psychometric function without changing its slope³². In our study, the slope of the function increased without concomitant change in the bias. Increasing the slope through conscious or unconscious strategies is far more unlikely than changing the bias. The latter only requires, in case of uncertainty, to respond more often with one of the two answers. The slope of the psychometric function can be changed in two directions. An individual participant can decrease the slope by simply responding more randomly. However, to increase the slope of the psychometric function,

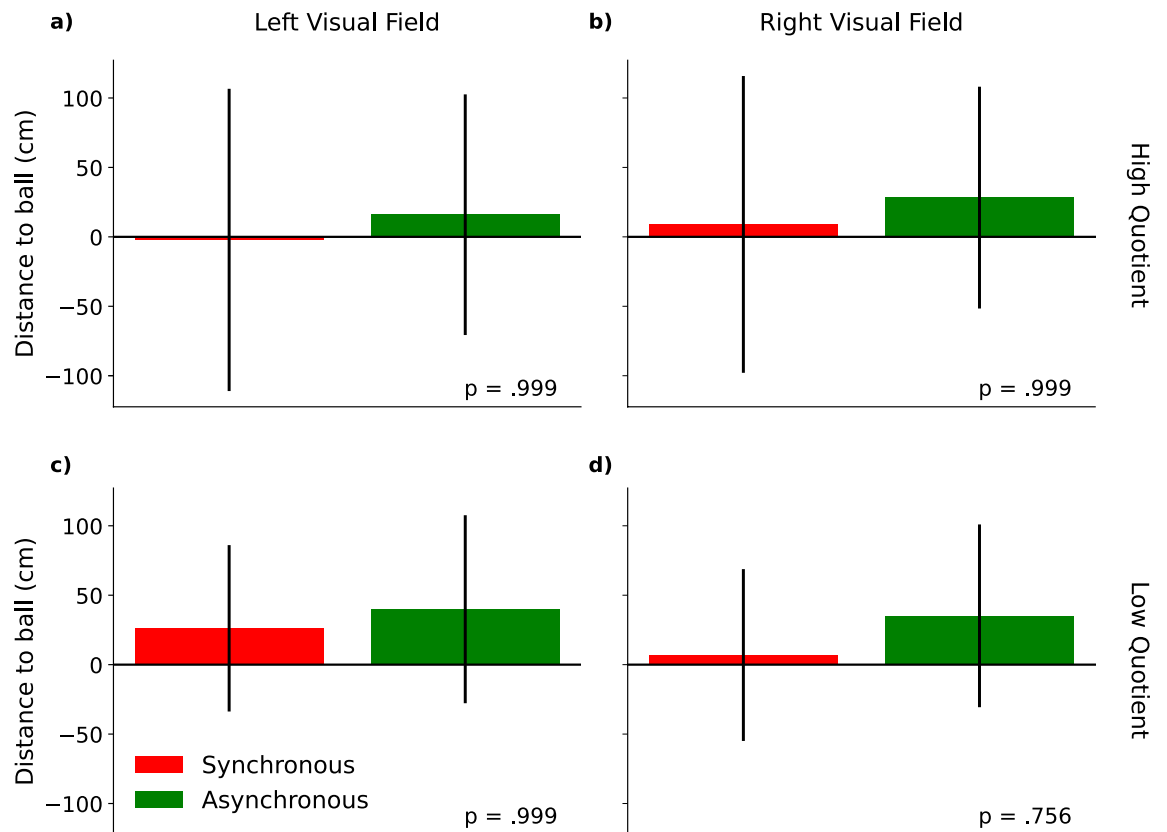


Figure 5. Mean distances between the ball and the participants in the mental imagery task. (a) The mean distances at the time of the participants' response for the synchronous (green) and asynchronous condition (red) for the high quotient group from trials in which the ball approached the participant in the left visual field. Error bars indicate the standard error. Positive values indicate that participants responded too early, the ball was still in front of them. Negative values indicate a late response, e.g. the ball was already behind them. Participants were accurate in their judgement of the position of the approaching ball and no difference between the conditions were found. (b–d) The same pattern was observed for the high quotient group in trials in which the ball approached the participant in the right visual field and for the low quotient group irrespective of the visual field the ball was approaching from.

the participant must actually know the correct response. There is no reason to suppose that a group of observers can systematically guess the correct answers to pretend a signature of better discrimination performance.

How can an illusion yield better discrimination performance in depth perception? *Prima facie*, the effect might seem paradoxical, since objectively measured visual discrimination performance increases, following the induction of an illusory change in position. Improvements in discrimination through illusory changes of body-parts have been shown previously: Vignemont et al.²³ demonstrated that the artificial elongation of individual fingers can lead to an increase in tactile discrimination sensitivity. For the present study, the projection on the two retinæ produces ambiguous information about the position of objects in depth. A small but close object is connected to the same retinal image as a bigger object located far away^{3,4}. Previous studies have shown that binocular depth perception is calibrated to the “natural grasping distance”, i.e. the distance that is usually chosen for manual interaction^{5,6}. Planning interactions with the external world requires that we have an implicit knowledge about our own position with regard to the location of objects in the world. Put differently, executing a motor plan means to minimize the distance between start and desired end location of an effector. In order to perform a goal directed hand movement, the brain must know the current hand location that is the starting position of the hand movement.

Linkenauger et al.³³ were able to show that we utilize the length of our limbs and the associated grasping distance to judge distances. In their study they used a virtual reality setup in which participants performed grasping movements while being represented by an avatar with various arms lengths. The mere manipulation of the avatars arm length did not have an effect on the participants' depth perception, only after they were able to collect experiences with their manipulated arm length their judgements of distances changed. This result further provides support for the importance of grasping distance and our experiences in interacting with our environment for our perception of depth. The full-body illusion modifies the internal spatial representation of the body with regard to external space and shifts self-location towards the seen position of the avatar^{9,10,17,26}. Changes in the felt position in space through the full-body illusion may affect the internal representation of the natural grasping or walking distance and thereby the calibration of visual depth.

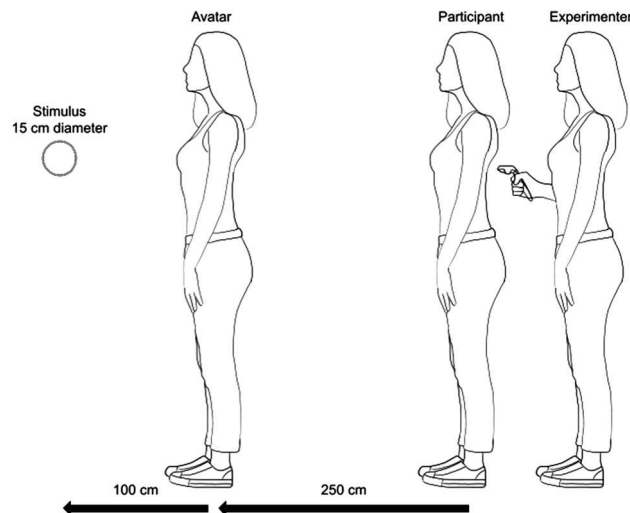


Figure 6. Graphical illustration of the experimental setup. The displayed avatar was only present during the full-body illusion induction periods and was adjusted to the participant's gender and height. It was positioned 250 cm in front of the participant in the VR environment. The probe stimulus in each trial of the visual depth task was presented 100 cm in front of the avatar and 350 cm in front of the participant.

Although we did not anticipate lateralized effects on space perception, we now discuss several factors that might contribute to it. The realism of the used avatars and the representation of the body of the participant could have had an impact on the results as well. A study of Ebrahimi et al.³⁴ for example provided support for the necessity of a realistic representation of the participants body in the virtual environment in the context of depth perception. Lugin et al.³⁵ on the other hand showed that the level of anthropomorphism did not influence virtual body ownership. Participants in our study were required to give their responses with a VR controller held in their dominant hand. Since the controller followed their movement, it could have been regarded as a form of representation of their own body. As this was only the case for one side of their body, this could have impacted their depth perception unilaterally. Motor actions or the mere visual perception of an object representing their right hand might have caused recalibration processes of the peripersonal space in the respective visual field. An alternative explanation is that the full-body illusion distorts the left and right visual field unequally or only affects one. Another reason for only finding a lateralized effect might be the lack of power. While we aimed to have a sample comparable to the ones used in previous studies investigating the full-body illusion, a bigger sample size could have potentially enabled us to find the effect also for the right side of the visual field.

Our results in the mental imagery task can probably be explained by our modification of the task. The original task used by Nakul et al.²⁶ used trajectories on the anterior-posterior axis while we used angled trajectories, which might have interfered with the task's ability to quantify the full-body illusion related forward shift. The idea behind the angled trajectories was to make the task more similar to the visual depth task in which stimuli also appear in an angle relative to the participant's view.

In conclusion, our results show that depth perception is partially calibrated by signals that mediate our perception of where we are located in space. The full-body illusion influences the interpretation of early visual processing at least unilaterally by modifying the internal representation of the body's position in relation to external objects.

Methods

Participants. A total of 20 participants took part in the experiment. The sample included 13 females and seven males (M_{age} : 22.90, SD_{age} : 3.58). The sample size was determined based on the sample sizes used in similar studies. The Edinburgh Handedness Inventory was used to quantify the handedness of the participants (M : 77.38, SD : 31.32). Two participants reported that they are left handed. All participants had normal or corrected-to-normal vision. Every participant gave written informed consent prior to the experiment in accordance with the declaration of Helsinki, participated voluntarily and received either course credit or 10 € for each hour of participation as monetary compensation. This study was approved by the local ethics committee of the mathematical and natural science faculty at the Heinrich Heine University.

Setup. After participants gave their written consent, they were equipped with a head-mounted display. The HTC Vive with Dual AMOLED 3.6" screens, a resolution of 1080 × 1200 pixels per eye (2160 × 1200 pixels combined), a refresh rate of 90 Hz and a field of view of 110 degrees was used. The participants were standing upright during the experiment.

The full-body illusion was induced by presenting a virtual avatar in front of the participant, while the experimenter stroked the participant's back. Simultaneously with the physical stroking, the participant saw the virtual avatar being stroked in the head-mounted display⁹. The physical stroking was performed by the experimenter with a tracked hand-held VR controller (see Fig. 6).

Stimulus presentation was generated by a custom program created with Unreal Engine (version 4.25, <https://www.unrealengine.com>) and was conducted on a Windows 10 desktop computer (Alienware Aurora R8, Intel Core i7-8700K @3.7GHz, 16 GB RAM, NVIDIA GeForce GTX 1080 graphics card). The virtual environment was run using SteamVR (version 1.17, <https://store.steampowered.com/app/250820/SteamVR/>) with the SteamVR 1.0 tracking system. Previous research has shown that the system provides suitable tracking of head and hand positions for research purposes if tracking loss is prevented³⁶. There were no salient visual reference points in the virtual environment (see Fig. 7).

Visual depth task. In the visual depth task participants had to judge the distance of two spheres relative to their perceived self-location in the VR world by indicating which of the two is closer to them with a press on the touchpad of the VR controller. The two spheres had a diameter of 15 cm and were each shown for 200 ms (100 cm apart from each other and 8.13 rotational degree to each side of the central line of sight). The spheres were presented consecutively with an inter stimulus interval of 200–250 ms. We refer to the first presented sphere in a trial as the probe and to the second sphere as the reference. The probe was always presented at a distance of 350 cm in front of the participant, while the position in depth of the reference was systematically varied on the anterior-posterior axis in six equidistant increments (5, 15 or 25 cm further away or closer to the participant), which were presented equiprobably across trials. The side of the first presented sphere and the side of the sphere with the variable position was counterbalanced across trials and randomized in order.

Mental imagery task. In the mental imagery task a red ball appeared at the end of the room at an angle of approximately 15° to the left or right from the view direction of the participant²⁶. The ball rolled on the floor toward the participant's viewpoint for 2 s at a constant velocity. Before the ball arrived at the location of the participant, a black screen was displayed. Participants were instructed to imagine the ball continuing just as before and to indicate when the ball would reach the position between their feet by touching the touchpad of the VR controller, which they were holding with their dominant hand.

Procedure. Experiments utilizing or investigating the full-body illusion usually compare two conditions against each other, one in which the stroking is synchronous to what the participants see on the avatars back and one with asynchronous stroking. The full-body illusion is thought to be successfully induced in the synchronous condition, while the asynchronous one serves as a baseline. In the asynchronous condition, the visual information followed the tactile stimulation, by using a delay between the felt and visual stroking seen in the head-mounted display. We used a delay of 500 ms, which is the most common delay used²⁶. Figure 7 shows the virtual environment the participants were experiencing in the head-mounted display during the induction of the full-body illusion. The avatar's appearance was adjusted and scaled to match the height and the gender of the participant.



Figure 7. Virtual environment of the experiment during the induction of the full-body illusion from the view point of the participant. The environment only comprised a grey surface to keep potential references to a minimum during the different tasks.

The synchronous and asynchronous stroking conditions were tested in separate sessions. The order of all conditions was counterbalanced across participants. During each run the participant performed the visual depth task, the mental imagery task and underwent stroking segments.

During the stroking period, the participant handed the VR controller to the experimenter, who performed a short calibration for the motion tracking. Then, an avatar appeared 250 cm in front of the participant facing straight away. In the following period the stroking was applied to the whole back of the participant, who was instructed to focus on the avatar, which was either stroked synchronously or asynchronously. The duration of this period was 60 seconds in the first block of each individual run and 30 seconds in the following blocks. The avatar was only visible during this period.

At the beginning of a session, the participant performed 10 training trials for the visual depth task and mental imagery task respectively. In the following period participants performed a total of 240 trials of the visual depth task and 40 trials of the mental imagery task split across 20 blocks split into two individual runs. Each block

began with the induction of the full-body illusion. Depending on the condition of the respective run, this induction was either synchronous or asynchronous. Afterwards, twelve visual depth task and two mental imagery task trials were performed, which took approximately seven seconds. The order of the two tasks was randomized throughout the whole experiment.

After the last block, an additional stroking period was performed to re-induce the full-body illusion before participants were asked to fill out a questionnaire²². The questionnaire was displayed on a Dell Monitor (1920 × 1080 pixel). The questionnaire comprised five different items and participants used a computer mouse to indicate their agreement on a seven-point Likert scale. Item 1 and 3 were chosen to determine changes in perceived self-location. Item 1 inquired participants to what extent they felt as if they were slightly above or below the seen avatar and item 3 to what extent they felt as if their own body shifted toward the seen avatar. Item 2 covered the aspect of self-identification by inquiring the participants to what extent they felt as if the body they saw was their own. Item 4 was intended to quantify the extent of illusory touch experiences by inquiring the participants to what extent the stroking felt as if it was located on the avatar. Item 5 served as a control item and inquired participants to what extent they felt as if nothing changed. Since the original items from Salomon et al.²² were in English they were translated into German.

Statistical analyses. Statistical analyses were performed in python, using the package scipy³⁷ (version 1.7.3, <https://scipy.org>). All reported p-values are Bonferroni corrected when performed in the context of multiple comparisons.

Visual depth task. From the participants' responses in the visual depth task we calculated psychometric functions by fitting cumulative gaussian functions to the average data of each reference sphere position. To estimate discrimination performance in depth, we determined the JND by selecting the variance of the psychometric function. We also determined the PSE, given by the mean of the psychometric function, to estimate the bias in depth perception. The mean number of responses per stimulus level for the computed psychometric functions was 19.62 (SD = 5.17). These calculations were performed for the synchronous and asynchronous condition and for each side of the visual field the probe was presented on individually (left or right visual field), resulting in four values for both, the JND and PSE, for each participant.

Mental imagery task. For the mental imagery task, we calculated the distance between the ball and the real position of the participant in the VR world for each trial. These values were then compared between the synchronous and asynchronous sessions.

Questionnaire. We estimated the success of inducing the full body illusion by analyzing the subjective measure, i.e. the questionnaire that participants filled out after each session. Since the illusion is elicited when the experimenter strokes the participants' back synchronously to the stroking seen in the head-mounted display, we compared item scores from sessions with synchronous vs. sessions with asynchronous stroking.

Data availability

The datasets analysed during the current study are available from the corresponding author on reasonable request.

Received: 24 March 2023; Accepted: 26 June 2023

Published online: 29 June 2023

References

1. Servos, P., Goodale, M. A. & Jakobson, L. S. The role of binocular vision in prehension: A kinematic analysis. *Vis. Res.* **32**, 1513–1521. [https://doi.org/10.1016/0042-6989\(92\)90207-Y](https://doi.org/10.1016/0042-6989(92)90207-Y) (1992).
2. Nicolle, J.M.-T.P.-D., Servos, P. & Goodale, M. Adapting to monocular vision: Grasping with one eye. *Exp. Brain Res.* **104**, 107–114 (1995).
3. Johnston, E. B. Systematic distortions of shape from stereopsis. *Vis. Res.* **31**, 1351–1360 (1991).
4. Norman, J. F., Todd, J. T., Perotti, V. J. & Tittle, J. S. The visual perception of three-dimensional length. *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 173 (1996).
5. Volcic, R., Fantoni, C., Caudek, C., Assad, J. A. & Domini, F. Visuomotor adaptation changes stereoscopic depth perception and tactile discrimination. *J. Neurosci.* **33**, 17081–17088. <https://doi.org/10.1523/JNEUROSCI.2936-13.2013> (2013).
6. Wiesing, M., Kartashova, T. & Zimmermann, E. Adaptation of pointing and visual localization in depth around the natural grasping distance. *J. Neurophysiol.* **125**, 2206–2218. <https://doi.org/10.1152/jn.00012.2021> (2021).
7. Cont, C. & Zimmermann, E. The motor representation of sensory experience. *Curr. Biol.* **31**, 1029–1036.e2. <https://doi.org/10.1016/j.cub.2020.11.032> (2021).
8. Ehrsson, H. H. The experimental induction of out-of-body experiences. *Science* **317**, 1048–1048. <https://doi.org/10.1126/science.1142175> (2007).
9. Lenggenhager, B., Tadi, T., Metzinger, T. & Blanke, O. Video ergo sum: Manipulating bodily self-consciousness. *Science* **317**, 1096–1099. <https://doi.org/10.1126/science.1143439> (2007).
10. Lenggenhager, B., Halje, P. & Blanke, O. Alpha band oscillations correlate with illusory self-location induced by virtual reality. *Eur. J. Neurosci.* **33**, 1935–1943. <https://doi.org/10.1111/j.1460-9568.2011.07647.x> (2011).
11. Botvinick, M. & Cohen, J. Rubber hands 'feel' touch that eyes see. *Nature* **391**, 756–756. <https://doi.org/10.1038/35784> (1998).
12. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
13. Blanke, O., Slater, M. & Serino, A. Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* **88**, 145–166. <https://doi.org/10.1016/j.neuron.2015.09.029> (2015).
14. van der Veer, A. H., Alsmith, A. J. T., Longo, M. R., Wong, H. Y. & Mohler, B. J. Where am i in virtual reality?. *PLOS ONE* **13**, 1–10. <https://doi.org/10.1371/journal.pone.0204358> (2018).

15. Ionta, S., Gassert, R. & Blanke, O. Multi-sensory and sensorimotor foundation of bodily self-consciousness: An interdisciplinary approach. *Front. Psychol.* **2**, 383. <https://doi.org/10.3389/fpsyg.2011.00383> (2011).
16. Palluel, E., Aspell, J. E. & Blanke, O. Leg muscle vibration modulates bodily self-consciousness: Integration of proprioceptive, visual, and tactile signals. *J. Neurophysiol.* **105**, 2239–2247. <https://doi.org/10.1152/jn.00744.2010> (2011).
17. Aspell, J. E., Lenggenhager, B. & Blanke, O. Keeping in touch with one's self: Multisensory mechanisms of self-consciousness. *PLOS ONE* **4**, 1–10. <https://doi.org/10.1371/journal.pone.0006488> (2009).
18. Hohwy, J. & Paton, B. Explaining away the body: Experiences of supernaturally caused touch and touch on non-hand objects within the rubber hand illusion. *PLOS ONE* **5**, 1–10. <https://doi.org/10.1371/journal.pone.0009416> (2010).
19. Pfeiffer, C., Schmutz, V. & Blanke, O. Visuospatial viewpoint manipulation during full-body illusion modulates subjective first-person perspective. *Exp. Brain Res.* **232**, 4021–4033. <https://doi.org/10.1007/s00221-014-4080-0> (2014).
20. Duhamel, J.-R., Colby, C. L. & Goldberg, M. E. *Congruent Representations of Visual and Somatosensory Space in Single Neurons of Monkey Ventral Intra-parietal Cortex (Area VIP)* 223–236 (Oxford University Press, 1991).
21. Preuss, N. & Ehrsson, H. H. Full-body ownership illusion elicited by visuo-vestibular integration. *J. Exp. Psychol. Hum. Percept. Perform.* **45**, 209. <https://doi.org/10.1037/xhp0000597> (2019).
22. Salomon, R., Lim, M., Pfeiffer, C., Gassert, R. & Blanke, O. Full body illusion is associated with widespread skin temperature reduction. *Front. Behav. Neurosci.* <https://doi.org/10.3389/fnbeh.2013.00065> (2013).
23. Vignemont, F. D., Ehrsson, H. & Haggard, P. Bodily illusions modulate tactile perception. *Curr. Biol.* **15**, 1286–1290. <https://doi.org/10.1016/j.cub.2005.06.067> (2005).
24. Pamment, J. & Aspell, J. Putting pain out of mind with an 'out of body' illusion. *Eur. J. Pain* **21**, 334–342. <https://doi.org/10.1002/ejp.927> (2017).
25. Blanke, O. Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* **13**, 556–571. <https://doi.org/10.1038/nrn3292> (2012).
26. Nakul, E., Orlando-Dessaints, N., Lenggenhager, B. & Lopez, C. Measuring perceived self-location in virtual reality. *Sci. Rep.* **10**, 6802. <https://doi.org/10.1038/s41598-020-63643-y> (2020).
27. Serino, A. *et al.* Bodily ownership and self-location: Components of bodily self-consciousness. *Conscious. Cogn.* **22**, 1239–1252. <https://doi.org/10.1016/j.concog.2013.08.013> (2013).
28. Lenggenhager, B., Mouthon, M. & Blanke, O. Spatial aspects of bodily self-consciousness. *Conscious. Cogn.* **18**, 110–117. <https://doi.org/10.1016/j.concog.2008.11.003> (2009).
29. Noel, J.-P., Pfeiffer, C., Blanke, O. & Serino, A. Peripersonal space as the space of the bodily self. *Cognition* **144**, 49–57. <https://doi.org/10.1016/j.cognition.2015.07.012> (2015).
30. Naceri, A., Chellali, R. & Hoinville, T. Depth perception within peripersonal space using head-mounted display. *Presence Teleoper. Virtual Environ.* **20**, 254–272. https://doi.org/10.1162/PRES_a_00048 (2011).
31. Loomis, J. M. *et al.* Visual perception of egocentric distance in real and virtual environments. *Virtual Adapt. Environ.* **11**, 21–46 (2003).
32. Morgan, M., Dillenburger, B., Raphael, S. & Solomon, J. A. Observers can voluntarily shift their psychometric functions without losing sensitivity. *Attent. Percept. Psychophys.* **74**, 185–193. <https://doi.org/10.3758/s13414-011-0222-7> (2012).
33. Linkenauger, S. A., Bühlhoff, H. H. & Mohler, B. J. Virtual arm's reach influences perceived distances but only after experience reaching. *Neuropsychologia* **70**, 393–401. <https://doi.org/10.1016/j.neuropsychologia.2014.10.034> (2015).
34. Ebrahimi, E., Hartman, L. S., Robb, A., Pagano, C. C. & Babu, S. V. Investigating the effects of anthropomorphic fidelity of self-avatars on near field depth perception in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 1–8. <https://doi.org/10.1109/VR.2018.8446539> (2018).
35. Lugin, J.-L., Latt, J. & Latoschik, M. E. Avatar anthropomorphism and illusion of body ownership in vr. In *2015 IEEE Virtual Reality (VR)*, 229–230. <https://doi.org/10.1109/VR.2015.7223379> (2015).
36. Niehorster, D. C., Li, L. & Lappe, M. The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research. *i-Perception* <https://doi.org/10.1177/2041669517708205> (2017).
37. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for scientific computing in python. *Nat. Methods* **17**, 261–272. <https://doi.org/10.1038/s41592-019-0686-2> (2020).

Acknowledgements

This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement Number 757184-moreSense) to E.Z.

Author contributions

M.B., E.Z. and O.B. conceived the experiment, M.B. conducted the experiment and analysed the results. All authors reviewed the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

RESEARCH ARTICLE

Sensory Processing

Serial dependencies in visual stability during self-motion

 Manuel Bayer and Eckart Zimmermann

Institute for Experimental Psychology, Heinrich-Heine-University Düsseldorf, Germany

Abstract

Every time we move our head, the brain must decide whether the displacement of the visual scene is the result of external or self-produced motion. Gaze shifts generate the biggest and most frequent disturbance of vision. Visual stability during gaze shifts is necessary for both, dissociating self-produced from external motion and retaining bodily balance. Here, we asked participants to perform an eye-head gaze shift to a target that was briefly presented in a head-mounted display. We manipulated the velocity of the scene displacement across trials such that the background moved either too fast or too slow in relation to the head movement speed. Participants were required to report whether they perceived the gaze-contingent visual motion as faster or slower than what they would expect from their head movement velocity. We found that the point of visual stability was attracted to the velocity presented in the previous trial. Our data reveal that serial dependencies in visual stability calibrate the mapping between motor-related signals coding head movement velocity and visual motion velocity. This process is likely to aid in visual stability as the accuracy of this mapping is crucial to maintain visual stability during self-motion.

NEW & NOTEWORTHY We report that visual stability during self-motion is maintained by serial dependencies between the current and the previous gaze-contingent visual velocity that was experienced during a head movement. The gaze-contingent scene displacement velocity that appears normal to us thus depends on what we have registered in the recent history of gaze shifts. Serial dependencies provide an efficient means to maintain visual stability during self-motion.

head movement; self-motion; serial dependence; visual stability

INTRODUCTION

When we walk through a clouded street, a plethora of different stimuli impinge on our sensory receptors. The spatial localization of these stimuli is complicated by our movements changing the position of the receptors relative to the external world. To interpret stimuli in the external world it is thus of fundamental importance to differentiate between movements of the own body and stimulus motion. At the same time, movement planning and keeping the body in the right posture also depend on distinguishing self-produced from external motion (1–3). The weight the brain gives to this discrimination can be painfully realized in situations that aggravate this process, like reading in a car or wearing a poorly calibrated head-mounted display (4). Instances of motion sickness illustrate that the brain predicts the regular coupling of a certain head movement with a certain motion velocity. Another striking example demonstrating the intimate connection of self- and visual motion is the experience

of illusory vection, i.e., stationary observers perceive self-motion when they see whole field motion (5).

Visual stability during self-motion is a long-standing, unresolved problem in neuroscience (1, 6–9). How the brain accomplishes visual stability during gaze shifts has mostly been studied during saccade eye movements (10). However, vision during head-fixed saccades changes differently than over the course of a head movement. During saccades, motion produced by the retinal displacement is omitted (11). Over the course of a head movement, which on average lasts 400–800 ms (12, 13), the motion produced by the gaze shift can be clearly seen. A gaze shift mostly contains both, an eye and a head component. The two components are coded as a combined gaze shift. Downstream of the superior colliculus level, they are separated into different saccade and head rotation commands (14–16). The dynamic of the gaze shift starts with a saccade displacing the line of sight toward a desired target. Then, the head follows the preceding saccade to recenter vision. Although the head is moving, the

vestibulo-ocular reflex triggers an eye movement in the direction opposite to the head movement, such that objects of interest are stabilized on the retina (17).

To judge visual stability under this circumstance, the retinal image motion must be transformed in suprarretinal, i.e., head-centered coordinates to evaluate the contribution of eye movements to the perceived displacement of the scene (18, 19). Subtracting the estimated head movement velocity from the visual motion in head-centered coordinates will reveal whether motion occurred in the external world or if it was produced by the own body (20). Humans rely on three sources to retrieve the velocity of the head movement: vestibular signals, which report the position and movement of the head in space, neck proprioception, representing a head-on-trunk signal, and an efference copy, consisting of a copy of the motor command (21, 22).

The vestibular system contains the semicircular canals transducing angular accelerations and the otolith organs transducing linear accelerations (20). Both subsystems provide information about head rotation velocity, which is accessible almost immediately. The fastest pathway conveying vestibular information initiates a vestibulo-ocular reflex at a minimum duration of only 14 ms (23). Similarly, as for saccade eye movements, the main sequence describes a linear relationship between head movement amplitudes and their peak velocities (24). An efference copy that encodes the desired displacement vector of the head movement should thus be usable to predict the displacement speed of the visual scene. However, the use of the efference is limited by a potentially variable state of the neck muscles (25, 26) and the possibility of head movement control being changed online (27).

The judgment of visual stability across head movements is a remarkable accomplishment, given the variety of signals and the corresponding coordinate transformations. It is thus not surprising that discrimination sensitivity for relative motion between the world and the observer is reduced during head movements (20, 28–30). How can the sensorimotor system ensure that the signals remain mapped accurately onto each other? Recalibration is necessary to guarantee that the gaze-contingent motion follows the predicted lawful transformation.

Research has now established that visual input is constantly biased toward previously registered features. Serial dependencies in vision have been shown for many perceptual attributes, like visual orientation (31–33), numerosity (34), spatial position (34, 35), facial identity and expression (36, 37), eye gaze (38), pulchritude (37), and body size (39). Their role is likely to smoothen out discontinuities in vision (31, 40). Serial dependencies also exist between errors produced by movements and visual space perception, which have been shown for saccades (41), head movements (27), and locomotion (42).

The biggest discontinuities in vision are produced when gaze shifts displace the entire visual scene. Here, we wondered if the estimate of visual stability during head movements is calibrated by serial dependencies between the current and the past displacement speeds of the visual scenes. Retaining an accurate prediction of the visual scene displacement during head movements might be solved by taking into account sensory input of the previous

movements. Wearing head-mounted displays, participants were asked to perform head movements to the remembered position of a target. During the execution of the head movement, the visual environment in the display moved in the opposite direction to mimic the visual displacement occurring in a natural gaze shift. Using a head-mounted display allows to change the ratio between the head movement velocity and the velocity of the visual displacement. After each performed head movement, participants were asked if the movement was perceived as faster or slower than what they expected.

Experimentally dissociating motor performances from their contingent sensory consequences yields a stimulus that is inherently ambiguous. If the dissociation is strong enough to be detected, the sensorimotor system can either attribute the error to a distorted internal mapping or to a movement of the external world. Depending on how participants attribute the error, they can perceive the velocity of the visual displacement as a modification of the consequences of their own movement or as external motion. In the latter case, participants would not be able to distinguish faster from slower visual displacements.

METHODS

Participants

Forty-six participants were tested. Due to poor task performance, the data of eight subjects had to be excluded. The final sample included 23 females and 15 males (M_{age} : 23.56, SD_{age} : 4.22). Every participant gave written informed consent before the experiment in accordance with the Declaration of Helsinki, participated voluntarily, and received either course credit or monetary compensation. This study was approved by the local ethics committee of the mathematical and natural science faculty at Heinrich-Heine University Düsseldorf.

Setup

Stimuli were presented by a custom program created with Unreal Engine (v. 4.26), running on a Windows 10 desktop computer (Alienware Aurora R8, Intel Core i7 8700 @3.2 GHz, 16 GB RAM, NVIDIA GeForce RTX 2080 graphics card). The used head-mounted display was an HTC Vive Pro with two dual AMOLED 3.5 in. screens, a resolution of $1,440 \times 1,600$ pixels per eye ($2,880 \times 1,600$ pixels combined), a refresh rate of 90 Hz, and a field of view of 110° . The built-in 120 Hz eye tracker was used to record the eye position and a custom Python app, working at a mean sample rate of 429.24 (SE = 2.28 Hz), was used to record the head position and orientation. The virtual environment was run using SteamVR (v. 1.25.3) with the SteamVR 2.0 tracking system. Previous research has shown that the system provides suitable tracking of head and hand positions for research purposes if tracking loss is prevented (43, 44). The virtual environment was a Gabor patch with a spatial frequency of 0.05 c/deg (specified in rotational degree), which fully encircled the participant.

Procedure

At the beginning of each experimental session, the participant performed a training. During the training, participants

were allowed to explore the virtual environment by freely moving their gaze. We manipulated the visual velocity gain by changing the velocity of the gaze-contingent motion shown in the head-mounted display. If, for example, an increased visual velocity gain was applied, the environment rotated against the direction of the head movement performed by the participant with higher velocity. After participants felt comfortable in the environment, the experimenter exposed them to different visual velocity gains, ranging from 0.7 to 1.3. The main experiment used visual velocity gains between 0.65 and 1.35. These were identified as the boundaries of the acceptable range that did not cause nausea in a pilot study.

During this initial training, participants also saw a red sphere in the center of their view. This sphere served as a marker of the central field of view and was therefore bound to the orientation of the head. In the main experiment, participants were required to fixate the stimuli with their head by aligning their center field of view with them. Once participants were able to distinguish visual velocity gains, the initial training was finished and the next training started. In this training, participants performed 50 trials of the experimental task with feedback.

A single trial of this task started with a 15×15 cm fixation cross presented 2.5 m in front of the position of the participant. Once the participant fixated the fixation cross with their head (center field of view), a head movement target cross appeared either 1.44 m to the left or right of the fixation cross (30°). The head movement target cross disappeared 255 ms after onset and the fixation cross disappeared 255 ms after the head movement target cross disappeared. Once both crosses disappeared, the participant was instructed to perform a head movement to the position of the target cross. Once they thought to have reached this position, they had to press the trigger of the virtual reality (VR) controller they were holding with their dominant hand. After the trigger press, they were asked if the gaze-contingent visual motion was perceived as faster or slower than what they expected. Participants responded with a press on the touchpad of the VR controller. Pressing the upper part of the touchpad implied the movement was perceived as faster, while the lower part implied a slower movement. The key bindings and questions were always presented slightly above their field of view after the trigger press. If the participant failed to fixate the center fixation cross with their head until both crosses disappeared, the trial was restarted. If the head movement was performed in the wrong direction or away from the head movement target cross, the trial was automatically aborted and excluded from analysis.

During training, participants also saw the red sphere as an indicator for their center field of view. After they gave their response in the task, they received feedback if their judgment was correct. In this training, only the increased visual velocity gains 1.12, 1.23, and 1.35 and the decreased visual velocity gains 0.65, 0.77, and 0.88 were used, whereas in the main experiment also trials with unity visual velocity gain were included. After the training was completed, the percentage of correct responses was displayed. If this value was below 60%, the training was repeated once before the main experiment started.

The main experiment consisted of 628 trials and was split into two runs. In this part of the experiment, the participants

did not see the red sphere in the center of their field of view and received no response feedback. Besides that, the task was identical to the task described for the training with feedback.

Head and Eye Movements

We first analyzed the head position data, which were recorded in rotational degrees and determined the start and end position of the head movements. We used moving averages with the 10 following visual velocity values to smoothen the data to more accurately determine the start and end. We then determined the time point at which the peak velocity of the respective trial was reached. From that time point, we checked the individual velocity values in the direction of the trial start and end. The boundaries of the head movement were set where the velocity values were below $3^\circ/\text{s}$. Trials which had not enough data points to compute the moving average or failed to meet the velocity criteria were excluded. Due to these criteria, 1.04% of trials were excluded across all participants.

We also analyzed the end position of the single trial head movements in relation to the achieved peak velocity within the respective trial (see Fig. 1). To test if this relationship is statistically significant, we performed a one-sided t test versus 0. Participants moved their heads further to the sides the higher the achieved peak velocity was during the trial, $t(37) = 8.16$, $P < 0.001$.

Statistical Analyses

Head movement analyses.

We calculated the head movement peak velocities and amplitudes for each participant for each individual visual velocity gain level for trial n and trial $n - 1$. We fitted linear functions to the resulting data points and tested these single-subject slopes for significance with one sample t tests.

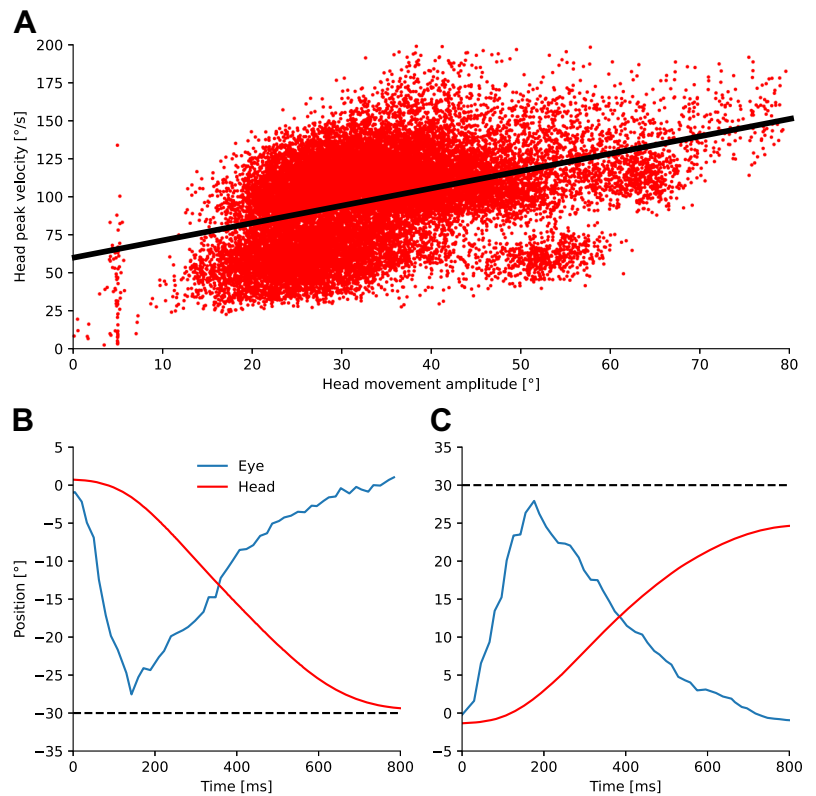
Psychometric analyses.

We calculated the point of visual stability (PVS) of each participant for each serial dependence gain level, e.g., one PVS for all trials, which were preceded by a visual velocity gain of 1.12. This was further split up into trials which were preceded by a trial with a head movement in the same direction (congruent) or in the opposite direction (incongruent).

We then fitted a linear function to the PVSs across the individual serial dependence gain levels for each participant. We fitted a cumulative Gaussian distribution to the responses across all visual velocity gain levels. For one group of participants, data were better described by two psychometric functions. To these data, we fitted one psychometric function to the decreased visual velocity gains (0.65–1.00) and one to increased visual velocity gains (1.00–1.35). This way we had one PVS for decreased visual velocity gains (decreased PVS) and one for increased visual velocity gains (increased PVS).

We also estimated the 25% and 75% consistent response levels. Participants with data better described by a single psychometric function were confident that a visual velocity gain of 0.89 was slower and a visual velocity gain of 1.12 was faster than what they expected. The same calculation was performed for the participants with data better described by

Figure 1. Head movement main sequence and example trials. **A:** average head movement main sequence, illustrating the relationship between head movement amplitude and peak velocity. Data were collapsed across trials and participants (fit function: $y = 59.88 + 1.14x$). **B:** gaze shift traces for the head movement (shown in red) and the eye movement component (shown in blue) from an example trial in which the target was presented on the left. The dashed line represents the target position. The eye moves to the target first, followed by the rotation of the head. During the head movement, the eye moves back to recenter its position. **C:** gaze shift traces from an example trial in which the target was presented on the right. Same conventions as in **B**.

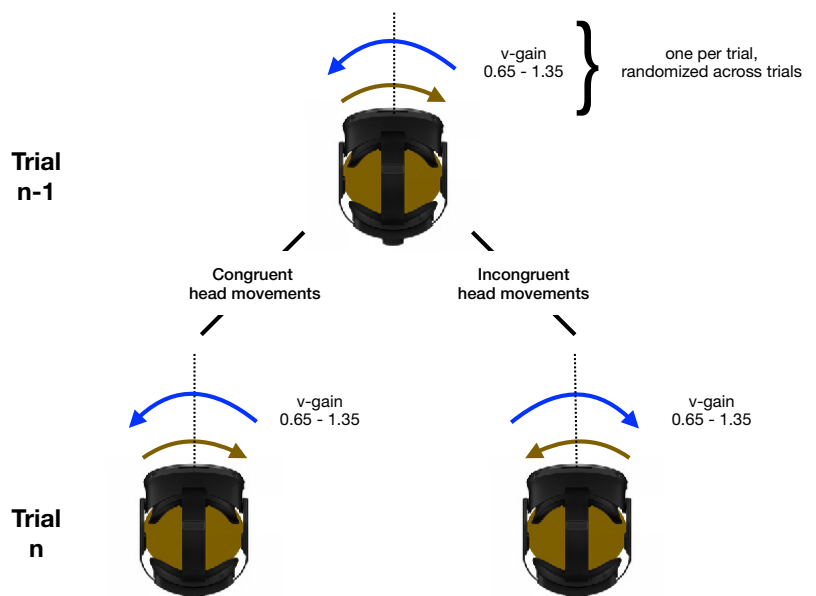


two psychometric functions but split into the increased and decreased PVS. For the increased PVS, participants of this group were confident that a visual velocity gain of 1.05 was slower and a gain of 1.24 was faster than what they expected. For the decreased PVS, participants of this group were confident that a visual velocity gain of 0.95 was slower and a visual velocity gain of 0.77 was faster than what they expected.

To analyze the putative occurrence of serial dependencies, we calculated the PVSs in relation to the visual velocity gain

(7 values) and the head movement direction (2 values) in the previous trial. This way we calculated 14 PVS values for each participant (see Fig. 2). Hereinafter, congruent refers to trial pairs that included two trials in which head movements were performed in the same direction and incongruent to trial pairs in which the first head movement was performed in the opposite direction to the second head movement. We then fitted linear functions to the single subject PVS values split by congruency, resulting in one slope value for

Figure 2. Serial dependence combinations. Illustration of different combinations of visual velocity gain and head movement direction. The brown arrows indicate the direction the head movement was performed while the blue arrows indicate the perceived relative motion. If, for example, a visual velocity gain above unity gain was active during the trial, the environment moved faster across the retina than if unity gain was active.



congruent and incongruent trial pairs for each participant. These slopes were then tested with a one-sample t test for significance. In the case of the group of participants with data better described by two psychometric functions, we performed these analysis steps independently for the increased and decreased PVS. Statistical analyses were performed in Python with the package `scipy` (45, v. 1.7.3). Data and used code are publicly available (<https://doi.org/10.17605/OSF.IO/24BQ9>).

RESULTS

Head Movements

We first analyzed head movement dynamics. Participants had to move their head to the remembered position of the target. Head movement dynamics followed the main sequence for head movements (24) with an average peak velocity of $98.97^\circ/\text{s}$ ($98.97^\circ/\text{s} \pm \text{SE } 0.20^\circ/\text{s}$) and a mean amplitude of 34.25° ($34.25^\circ \pm \text{SE } 0.07^\circ$). Head movements were performed with a mean velocity of $45.05^\circ/\text{s}$ ($45.05^\circ \pm \text{SE } 0.01$) and lasted 780.25 ms ($780.25 \pm \text{SE } 1.51 \text{ ms}$) on average.

We analyzed head movement peak velocities and amplitudes, separately for trial pairs with head movements in the same direction (termed “congruent”) and with head movements in the opposite direction (termed “incongruent,” see Fig. 3). We fitted linear functions to the peak velocities and head movement amplitudes of the individual visual velocity gains on the single subject level. We found a negative relationship between the visual velocity gain and the peak velocity, indicating that the manipulation of the visual velocity gain modified the gaze shift dynamics online. Participants had lower peak velocities the higher the visual velocity gain was [one-sample paired t test: $t(37) = -3.34$, $P = 0.002$]. Depending on how the difference between the expected and the actual visual velocity gain is interpreted, a higher visual velocity gain can either signify that the head movement is faster than expected or that the scene moved in the opposite direction as the gaze shift. In both cases, slowing down the head movement would be the necessary compensation. Consistent with this interpretation, participants also performed smaller head movements the higher the visual velocity gain was compared with unity gain [one-sample

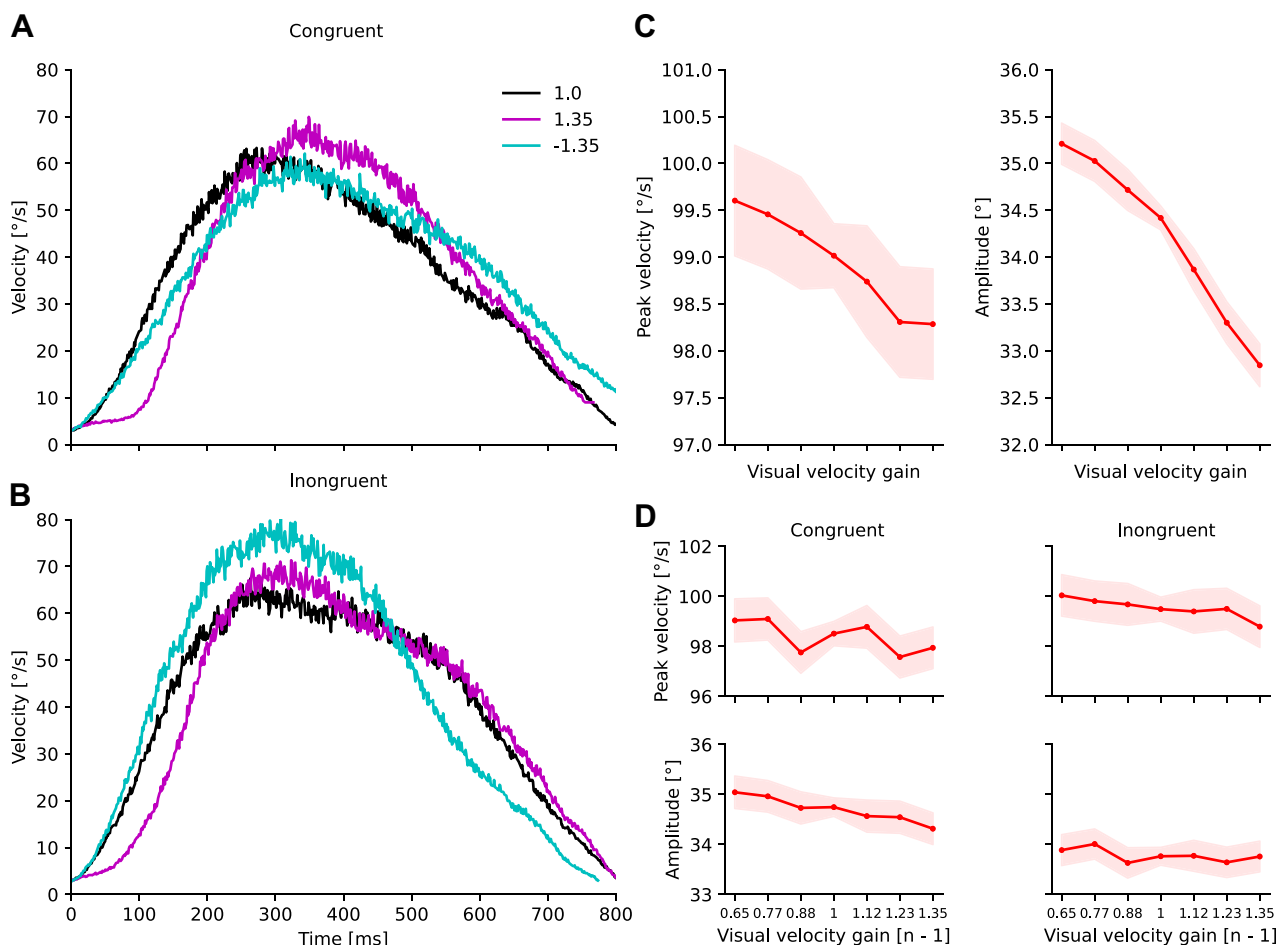


Figure 3. Head movement peak velocities and amplitudes in relation to the gain in trial n . **A:** velocity profile of three example trials in which congruent head movements were performed with either an increased (shown in magenta), a decreased (shown in cyan), or a unity visual velocity gain (shown in black). **B:** velocity profiles from incongruent head movements. Same conventions as in **A**. **C:** head movement peak velocities and amplitudes as a function of the visual velocity gain in trial n . One can see that there is a strong linear relationship between the visual velocity gain in trial n and the head movement amplitude and a slightly weaker relationship for peak velocity. **D:** head movement peak velocities and amplitudes as a function of the visual velocity gain in trial $n-1$. There were little to no relationships between the visual velocity gain of the previous trial and the amplitudes and peak velocities of trial n .

paired t test: $t(37) = -4.66$, $P < 0.001$]. These data show that the difference between expected and actual gaze-contingent visual velocity is processed to be readily available already during the head movement, thus underscoring the importance of visual stability during gaze shifts.

We then analyzed head movement dynamics as a function of the visual velocity gain in trial $n - 1$ split by the congruency of the head movement (see Fig. 3). The visual velocity gain in trial $n - 1$ did not change the peak velocities in trial n , neither for congruent [one-sample paired t test: $t(37) = -1.29$, $P = 0.411$] nor for incongruent trial pairs [one-sample paired t test: $t(37) = -1.70$, $P = 0.194$]. In contrast, the amplitude of the head movement was lower the higher the visual velocity gain was in trial $n - 1$ in congruent trial pairs [one-sample paired t test: $t(37) = -3.39$, $P = 0.003$]. No such effect was found for the head movement amplitude in incongruent trial pairs [one-sample paired t test: $t(37) = -0.87$, $P = 0.777$].

Psychometric Data

We next estimated the visual velocities at which participants experienced visual stability. To this end, we calculated psychometric functions, based on the average responses

indicating whether the displacement of the visual stimulus appeared faster or slower than expected. We selected all n trials in which the corresponding $n - 1$ trial had the same visual velocity gain and calculated separate psychometric functions from the average responses in the n trials. Psychometric functions from four example observers can be seen in Fig. 4A. We also separated trial pairs with congruent and incongruent head movements. This procedure yielded 2×7 psychometric functions for each participant, reflecting the factors of congruency and visual velocity gain in trial $n - 1$.

The velocity where the cumulative Gaussian function reached 50% was chosen as the point of visual stability (PVS). Psychometric functions from four example observers can be seen in Fig. 4A. In our analysis, 19 participants showed average responses that were better described by two separate psychometric functions, one for decreased visual velocity gains and one for increased visual velocity gains (see Fig. 4B). Such an average response would occur if participants interpret a change in the visual velocity gain as stimulus movement in the external world. An increased visual velocity gain means that the stimulus moves in the opposite direction than the head movement. A decreased visual

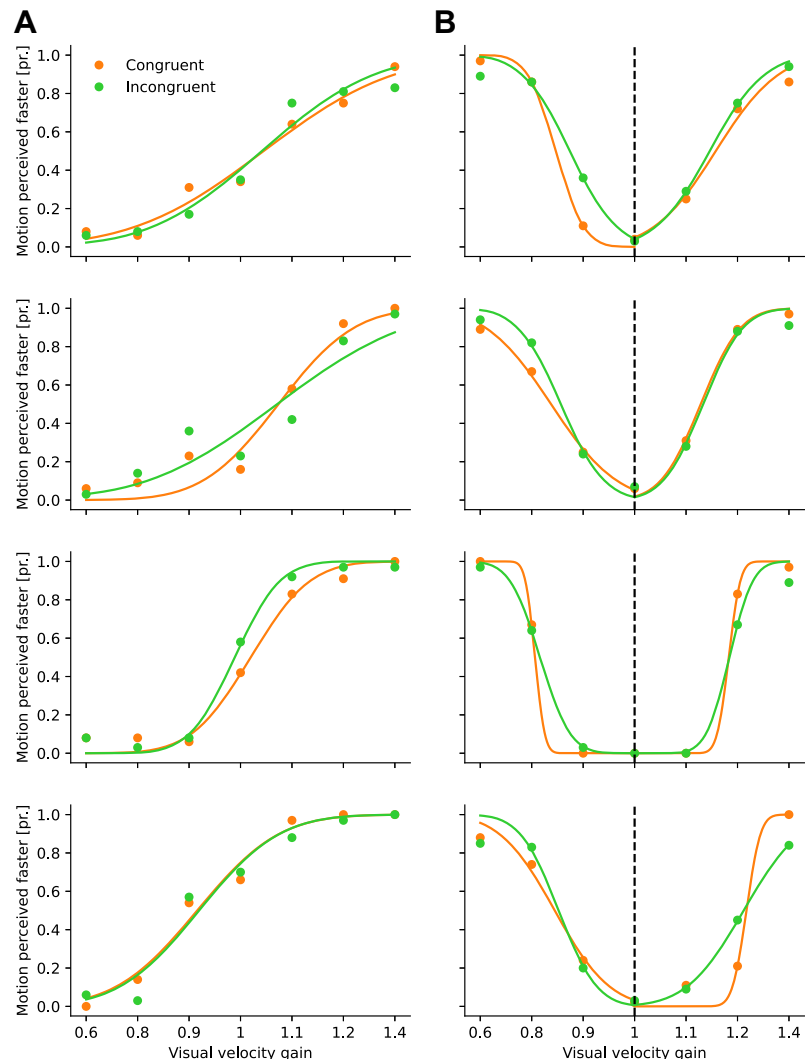


Figure 4. Example psychometric response functions. A: psychometric responses of four example participants. Data are split by the congruency of the head movement. B: four example participants whose data were better described by two individual psychometric functions.

velocity gain equals stimulus movement in the direction of the head movement. Under this interpretation, changes in the visual velocity gain would result in the perception of faster or slower motion, instead, they generate motion stimuli with the same velocity but in different directions.

Participants with data better described by a single psychometric function were accurate, with a PVS close to the unity gain ($M = -0.02$, $SD = 0.13$) and precise, with a just-noticeable difference (JND) of 0.22 ($SD = 0.14$). Similarly, participants with data better described by two individual psychometric functions had an average PVS of 0.17 ($SD = 0.06$) and a JND of 0.13 ($SD = 0.07$) for increased visual velocity gains and a PVS of -0.18 ($SD = 0.07$) and JND of 0.14 ($SD = 0.09$) for decreased visual velocity gains.

We also aimed to analyze whether the manipulation of the visual velocity gain would recalibrate the PVS. Such a shift in the PVS would become apparent in the head movements immediately following the head movement during which the manipulation of the visual velocity gain was applied. Figure 5A shows average PVSs for all $n - 1$ visual velocity gains for four example participants whose data were better described by a single psychometric function. One can see a clear linear relationship between the $n - 1$ visual velocity gain and the bias in the velocity estimate of the visual scene displacement. To estimate serial dependencies in visual stability between head movements, we fitted linear regression functions to the individual PVSs in dependence of the $n - 1$ visual velocity gains from each observer. Average slopes of the linear fits are shown in Fig. 5B. We found a significant increase in the estimate of the visual velocity in dependence of the previous visual velocity in trial $n - 1$ for congruent [slopes tested with a one-sample t test: $t(18) = 2.51$, $P = 0.011$] and incongruent trial pairs [$t(17) = 1.86$, $P = 0.040$].

We also analyzed serial dependencies in the group of participants which were better described by two psychometric functions. The same statistical procedure as described for the group of participants which was better described by a single psychometric function was used, but unsigned slopes were computed. Data from four example participants of this group can be seen in Fig. 4B. We found significant serial dependencies in congruent head movement trial pairs for increased trial n visual velocity gains [$t(18) = 2.09$, $P = 0.026$] and in incongruent head movement trial pairs for decreased trial n visual velocity gains [$t(18) = -2.63$, $P = 0.008$]. No significant relationship was observed between the visual velocity gain in trial $n - 1$ and trial n for increased trial n visual velocity gains in incongruent trial pairs [$t(18) = 0.83$, $P = 0.208$], nor for decreased trial n visual velocity gains in congruent trial pairs [$t(18) = -0.34$, $P = 0.371$; see Fig. 5C].

DISCUSSION

In this study, participants performed gaze shifts to the remembered location of a target. The gaze shift pattern consisted of a saccade to the target position, followed by a movement of the head, during which the eye slowly drifted in the opposite direction. The latter saccadic behavior exemplifies the vestibulo-ocular reflex that counteracts the head movement and stabilizes external stimuli on the retina (17) in combination with the optokinetic nystagmus, a sawtooth movement of the eye, which accomplishes the

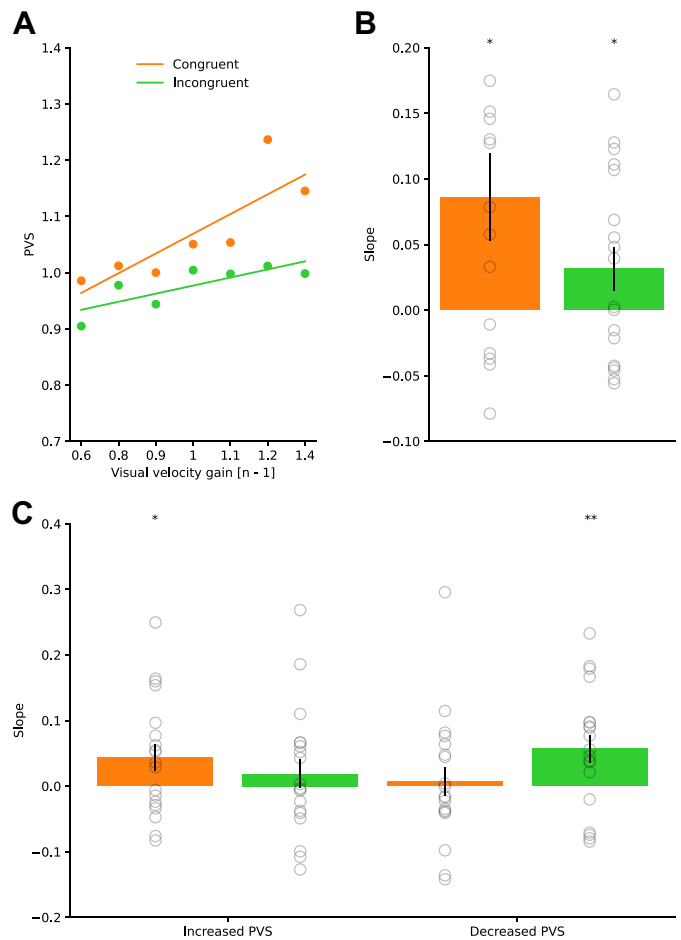


Figure 5. Point of visual stability (PVS) slopes for $n - 1$ split by head movement congruency. A: points of visual stability (PVSs) as a function of visual velocity gain in trial $n - 1$ of an example participant for trial pairs including congruent and incongruent head movements. Data are fitted by a linear function (fit function for congruent trial pairs: $y = 0.77 + 0.30x$; fit function for incongruent trial pairs: $y = 0.85 + 0.12x$). One can see that there is a clear linear relationship between the visual velocity gain of trial $n - 1$ and the PVS of trial n in congruent trial pairs. In incongruent trial pairs, this relationship is weaker. B: mean slopes for the psychometric response group split by congruency. Error bars indicate the standard error. The transparent circles represent single subject data. C: mean slopes for the increased visual velocity gains and decreased visual velocity gains split by congruency. Although there is a similar, even though smaller relationship between the visual velocity gains in trial $n - 1$ and the increased visual velocity gain PVS in congruent trial pairs, the opposite can be observed for the decreased visual velocity PVS. ** P value of below 0.01; * P value of below 0.05.

retinal stabilization of moving images. Visual stability during head movements requires the sensorimotor system to decide in each gaze shift whether the displacement of the visual scene was produced in the external world or if it was self-produced (46–49). We found that serial dependencies between the visual velocity during the previous and the current head movement determine at which visual velocity the scene displacement is considered as self or externally produced.

Although participants performed the gaze shift, we changed the velocity of the displacement in the virtual environment. When observers in the preceding trial experienced

a displacement faster than unity gain, i.e., that was too fast compared with their head movement velocity, their criterion of visual stability shifted toward faster movements and vice versa for scene displacements that were too slow. Put more concretely, if participants saw in the previous trial a background motion with a velocity slower than unity gain, they would perceive unity gain as too fast in trial n . Our data show that the neural mapping between head movement amplitudes and expected visual velocity is constantly recalibrated by serial dependencies. Deviations from this mapping, which we created artificially by manipulating the visual velocity, gain change behavior already online during the ongoing head movement. This is also reflected by changes in the head movement dynamics as participants adjusted their head movements to the active visual velocity gain. In the following movement, the point of visual stability was shifted in the direction of the visual velocity gain. This indicates that the sensorimotor system adjusts the prediction of gaze-contingent visual motion based on recent experience.

Judging external motion during self-motion requires internally estimating the movement velocity of the respective effector. For head movements, this estimate can be derived by integrating vestibular, proprioceptive, and efferent copy information (20–22). From this result, an estimate of the external motion must be subtracted. As the vestibulo-ocular reflex stabilizes the visual scene on the retina, the gaze-contingent scene displacement cannot be inferred directly from translation on the retina but must be retrieved indirectly by taking into account one or both of two possible signals. The first is the velocity of motion parallax following the angular deviation created by the head rotation. Second, the frequency of the change in eye position during optokinetic nystagmus informs about the scene displacement velocity when the retinal image is stabilized. Electrophysiological studies revealed that in the dorsal medial superior temporal area (50) and in the ventral intraparietal sulcus (51), motion parallax and extraretinal information, i.e., efference copy information about the eye position, interact synergistically. The subtraction of the relative external motion estimate from the head movement estimate yields an estimate of the absolute external motion. These computations involve the integration and comparison of several signal sources over coordinate transformations. Recalibrating the calculation of the head-centered motion is mandatory given the acute importance of visual stability for maintaining body posture, walking direction, and interpreting the visual scene correctly. Miscalibrations would be highly maladaptive, potentially inducing motion sickness in cases without an actual sensory conflict.

Recent research has revealed that sensory processing is constantly recalibrated by integrating information from previous and current percepts (for a review, see Ref. 52). Our results demonstrate that serial dependencies smoothen out the most drastic interruption of vision produced by gaze shifts. From our data, we can infer which signal from the previous head movement drives these serial dependencies and at which level they operate.

Two general response patterns appeared in our data. About half of the participants reported whether the visual velocity of the scene displacement was faster or slower than

their head movement. If the experimentally applied change in the scene displacement velocity exceeded $\sim 10\%$ of the head movement velocity, these participants could reliably discriminate external- from self-produced displacement. The second set of participants was more likely to report scene displacement velocities as faster, the stronger the visual velocity was changed experimentally, irrespective of the direction of the velocity change. This separate pattern of responses most likely stems from the ambiguous nature of the manipulation. Physically, the background is either moved against the direction of the head rotation (to create a visual velocity gain faster than unity) or it is moved in the same direction as the head rotation (to create a visual velocity gain slower than unity). Observers could either integrate the background motion with their head rotation into a head-centered perception and see the displacement as faster or slower than their head movement. Alternatively, participants could subtract their head rotation velocity from the velocity of the motion on the retina. In the latter case, they would retrieve an estimate of the physical background motion in the virtual environment and perceive, depending on the trial, leftward or rightward motion instead of slower and faster motion. Any discernible background motion in the environment during a head rotation, irrespective of its direction, is deviating from visual stability and should be considered as faster than what is usually expected. These judgments would result in a response pattern better described by two psychometric functions. It is likely that these participants interpreted the artificial change of the visual velocity gain as movement of the visual background in the external world. In this view, a decreased visual velocity gain would imply a movement of the background in the direction of the head movement and an increased visual velocity gain a movement against the direction of the head movement. Increased and decreased visual velocity gains thus produce motion stimuli with identical velocities but with opposite directions. In addition, there is also the possibility that participants interpreted the instruction differently such that one group of the participants evaluated the perceived motion relative to the external world and the other relative to their own head rotation.

Comparing these motion stimuli to the head movements would therefore yield sign-free velocity estimates. The results of the participants with a response pattern better described by two psychometric functions reveal which signal from the previous trial was taken into account and at which level it acted in the current movement. For head movements with a congruent direction in two consecutive trials, we found serial dependencies when the visual velocity gain was increased. By contrast, when two head movements with incongruent directions were performed, serial dependencies occurred only when the gain in the current trial was decreased. The most likely explanation for this result is that retino-centered visual motion in trial $n - 1$ influenced the calculation of head-centered motion in trial n under the condition that the direction of these two signals match. Figure 6 shows an illustration of this explanation. In congruent head movements (see Fig. 6A), the direction of retino-centered visual motion in trial $n - 1$ and head-centered motion in trial n match when the visual velocity gain in trial n is increased. In incongruent movements (see Fig. 6B), the opposite holds true: the

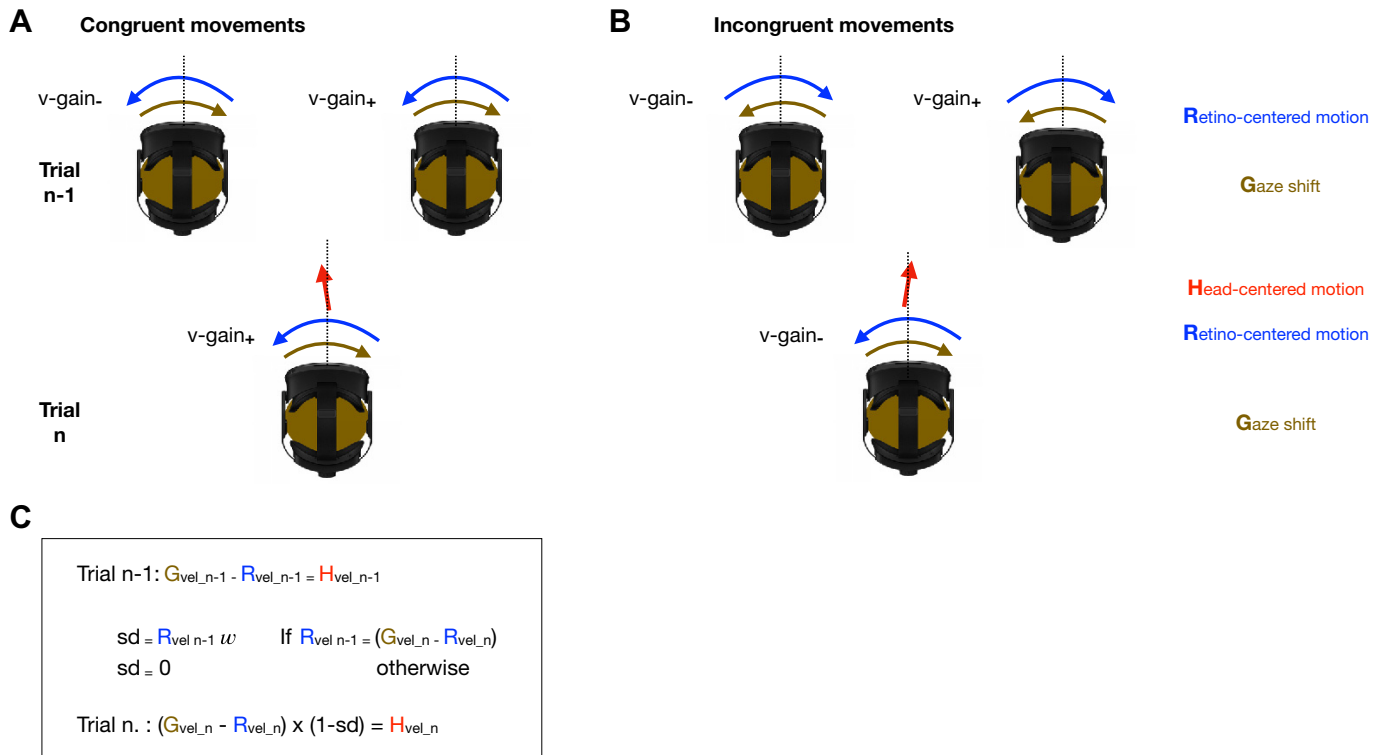


Figure 6. Serial dependence effect on visual stability. **A:** graphical illustration of the influences of the visual velocity gain applied during head movements that were performed consecutively in the same direction. Brown arrows show the direction of the head movements. Blue arrows show the direction of the gaze-contingent motion that was presented in the head-mounted displays. The red arrow shows the resulting head-centered motion. **B:** in trial pairs with incongruent head movements, the motion on the retina in trial $n - 1$ influenced the estimate of the head-centered velocity in trial n but only when the motion directions matched. When the visual scene moves with a decreased visual velocity gain, it can be either described as gaze-contingent motion that is too slow or as external motion in the same direction as the head movement. For incongruent head movement trial pairs, the direction of the retino-centered motion in trial $n - 1$ and head-centered motion in trial n match, when the visual velocity gain is decreased. **C:** formal description of how serial dependencies act on the calculation of head-centered motion.

direction of retino-centered visual motion in trial $n - 1$ and head-centered motion in trial n match when the visual velocity gain in trial n is decreased. That serial dependencies act only when motion directions match, reflects the tuning property of serial dependencies (40). Integration across dissimilar features would lead to erroneous calibration.

Recalibration of the predicted visual motion on the retina contingent with self-produced actions has been shown for other movement types. Haarmer et al. (53) manipulated the velocity of the background displacement in smooth pursuit eye movements. In test trials, they asked participants to judge whether the background was shifted leftward or rightward. Their estimates were biased by the velocity manipulation, indicating that the prediction of the self-generated motion on the retina is constantly recalibrated. Luna et al. (54) replicated the recalibration effect and showed that it is spatially specific for the visual field, thus limiting the receptive field size of the putative motion-processing neurons to maximally 10° .

Our analysis revealed that every gaze shift we perform serves to recalibrate visual stability. Serial dependencies act directly on perception (33) and their functional role is to smoothen out discontinuities in perception (31, 40). Presenting a continuously changing display that appeared as unchaining to observers, Manassi and Whitney (55)

could show that serial dependencies aim to stabilize vision over time. The most urgent need for visual stability in everyday perception occurs across gaze shifts. Unlike in saccades, where gaze-induced retinal motion is omitted by either an active (for a review, see 11) or a passive mechanism (56), during head movements, vision is consciously accessible. Average head movements mostly last long enough that retinal input can be evaluated and compared against an internal estimate of head movement velocity. Such an estimate derives from an efference copy, a vestibular signal or proprioception, informing predictively but also online about the head movement trajectory.

In conclusion, we report that visual stability during self-motion is maintained by attractive serial dependencies between the current and the previous gaze-contingent visual velocity that was experienced during a head movement. The gaze-contingent scene displacement velocity that appears normal to us thus depends on what we have registered in the recent history of gaze shifts. Serial dependencies provide an efficient means to maintain visual stability during self-motion.

DATA AVAILABILITY

Data and used code: <https://doi.org/10.17605/OSF.IO/24BQ9>.

GRANTS

This work was supported by HORIZON EUROPE European Research Council Grant 757184-moreSense.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

M.B. and E.Z. conceived and designed research; M.B. performed experiments; M.B. analyzed data; M.B. and E.Z. interpreted results of experiments; M.B. prepared figures; M.B. drafted manuscript; E.Z. edited and revised manuscript; M.B. and E.Z. approved final version of manuscript.

REFERENCES

- Wallach H. Perceiving a stable environment when one moves. *Annu Rev Psychol* 38: 1–27, 1987. doi:10.1146/annurev.ps.38.020187.000245.
- Wexler M. Voluntary head movement and allocentric perception of space. *Psychol Sci* 14: 340–346, 2003. doi:10.1111/1467-9280.14491.
- Rineau A-L, Bringoux L, Sarrazin J-C, Berberian B. Being active over one's own motion: considering predictive mechanisms in self-motion perception. *Neurosci Biobehav Rev* 146: 105051, 2023. doi:10.1016/j.neubiorev.2023.105051.
- Cohen B, Dai M, Yakushin SB, Cho C. The neural basis of motion sickness. *J Neurophysiol* 121: 973–982, 2019. doi:10.1152/jn.00674.2018.
- Keshavarz B, Riecke BE, Hettinger LJ, Campos JL. Vection and visually induced motion sickness: how are they related? *Front Psychol* 6: 472, 2015. doi:10.3389/fpsyg.2015.00472.
- Wallach H. Learned stimulation in space and motion perception. *Am Psychol* 40: 399–404, 1985. doi:10.1037//0003-066x.40.4.399.
- Gruesser OJ. Interaction of efferent and afferent signals in visual perception. A history of ideas and experimental paradigms. *Acta Psychologica* 63: 3–21, 1986. doi:10.1016/0001-6918(86)90039-9.
- Wertheim AH. Motion perception during selfmotion: the direct versus inferential controversy revisited. *Behav Brain Sci* 17: 293–311, 1994. doi:10.1017/S0140525X00034646.
- van der Steen FAM. An earth-stationary perceived visual scene during roll and yaw motions in a flight simulator. *J Vestib Res* 8: 411–425, 1998. doi:10.3233/VES-1998-8602.
- Wurtz RH. Corollary discharge contributions to perceptual continuity across saccades. *Annu Rev Vis Sci* 4: 215–237, 2018. doi:10.1146/annurev-vision-102016-061207.
- Binda P, Morrone MC. Vision during saccadic eye movements. *Annu Rev Vis Sci* 4: 193–213, 2018. doi:10.1146/annurev-vision-091517-034317.
- Andres RO, Hartung KJ. Prediction of head movement time using Fitts' law. *Hum Factors* 31: 703–713, 1989. doi:10.1177/001872088903100606.
- Hoffmann ER, Chan AHS, Heung PT. Head rotation movement times. *Hum Factors* 59: 986–994, 2017. doi:10.1177/0018720817701000.
- Freedman EG, Sparks DL. Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys. *J Neurophysiol* 77: 2328–2348, 1997. doi:10.1152/jn.1997.77.5.2328.
- Freedman EG, Stanford TR, Sparks DL. Combined eye-head gaze shifts produced by electrical stimulation of the superior colliculus in rhesus monkeys. *J Neurophysiol* 76: 927–952, 1996. doi:10.1152/jn.1996.76.2.927.
- Walton MMG, Bechara B, Gandhi NJ. Effect of reversible inactivation of superior colliculus on head movements. *J Neurophysiol* 99: 2479–2495, 2008. doi:10.1152/jn.01112.2007.
- Barnes GR. Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets. *J Physiol* 287: 127–147, 1979. doi:10.1113/jphysiol.1979.sp012650.
- Colby CL, Goldberg ME. Space and attention in parietal cortex. *Annu Rev Neurosci* 22: 319–349, 1999. doi:10.1146/annurev.neuro.22.1.319.
- Duhamel JR, Bremmer F, Ben Hamed S, Graf W. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389: 845–848, 1997. doi:10.1038/39865.
- Harris LR. Visual motion caused by movements of the eye, head and body. In *Visual detection of motion*, edited by Smith AT, Snowden RJ. London: Academic, 1994, 397–435.
- Crowell JA, Banks MS, Shenoy KV, Andersen RA. Visual self-motion perception during head turns. *Nat Neurosci* 1: 732–737, 1998. doi:10.1038/3732.
- Mergner T, Rottler G, Kimmig H, Becker W. Role of vestibular and neck inputs for the perception of object motion in space. *Exp Brain Res* 89: 655–668, 1992. doi:10.1007/BF00229890.
- Lisberger S. The latency of pathways containing the site of motor learning in the monkey vestibulo-ocular reflex. *Science* 225: 74–76, 1984. doi:10.1126/science.6610214.
- Zangemeister WH, Lehman S, Stark L. Simulation of head movement trajectories: model and fit to main sequence. *Biol Cybern* 41: 19–32, 1981. doi:10.1007/BF01836124.
- Scotland S, Adamo DE, Martin BJ. Sense of effort revisited: relative contributions of sensory feedback and efferent copy. *Neurosci Lett* 561: 208–212, 2014. doi:10.1016/j.neulet.2013.12.041.
- Monjo F, Terrier R, Forestier N. Muscle fatigue as an investigative tool in motor control: a review with new insights on internal models and posture-movement coordination. *Hum Mov Sci* 44: 225–233, 2015. doi:10.1016/j.humov.2015.09.006.
- Zimmermann E. Sensorimotor serial dependencies in head movements. *J Neurophysiol* 126: 913–923, 2021. doi:10.1152/jn.00231.2021.
- Choudhury BP, Crossey AD. Slow-movement sensitivity in the human field of vision. *Physiol Behav* 26: 125–128, 1981. doi:10.1016/0031-9384(81)90087-1.
- Johnson CA, Scobey RP. Effects of reference lines on displacement thresholds at various durations of movement. *Vision Res* 22: 819–821, 1982. doi:10.1016/0042-6989(82)90013-X.
- Snowden RJ. Sensitivity to relative and absolute motion. *Perception* 21: 563–568, 1992. doi:10.1068/p210563.
- Fischer J, Whitney D. Serial dependence in visual perception. *Nat Neurosci* 17: 738–743, 2014. doi:10.1038/nn.3689.
- Fritzsche M, Mostert P, de Lange FP. Opposite effects of recent history on perception and decision. *Curr Biol* 27: 590–595, 2017. doi:10.1016/j.cub.2017.01.006.
- Cicchini GM, Mikellidou K, Burr D. Serial dependencies act directly on perception. *J Vis* 17: 6, 2017. doi:10.1167/17.14.6.
- Bliss DP, Sun JJ, D'Esposito M. Serial dependence is absent at the time of perception but increases in visual working memory. *Sci Rep* 7: 14739, 2017. doi:10.1038/s41598-017-15199-7.
- Manassi M, Liberman A, Kosovicheva A, Zhang K, Whitney D. Serial dependence in position occurs at the time of perception. *Psychon Bull Rev* 25: 2245–2253, 2018. doi:10.3758/s13423-018-1454-5.
- Liberman A, Fischer J, Whitney D. Serial dependence in the perception of faces. *Curr Biol* 24: 2569–2574, 2014. doi:10.1016/j.cub.2014.09.025.
- Taubert J, Alais D. Serial dependence in face attractiveness judgments tolerates rotations around the yaw axis but not the roll axis. *Visual Cognition* 24: 103–114, 2016. doi:10.1080/13506285.2016.1196803.
- Alais D, Kong G, Palmer C, Clifford C. Eye gaze direction shows a positive serial dependency. *J Vis* 18: 11, 2018. doi:10.1167/18.4.11.
- Alexi J, Cleary D, Dommissie K, Palermo R, Kloth N, Burr D, Bell J. Past visual experiences weigh in on body size estimation. *Sci Rep* 8: 215, 2018. doi:10.1038/s41598-017-18418-3.
- Cicchini GM, Mikellidou K, Burr DC. The functional role of serial dependence. *Proc Biol Sci* 285: 20181722, 2018. doi:10.1098/rspb.2018.1722.
- Cont C, Zimmermann E. The motor representation of sensory experience. *Curr Biol* 31: 1029–1036.e2, 2021. doi:10.1016/j.cub.2020.11.032.
- Wiesing M, Zimmermann E. Serial dependencies between locomotion and visual space. *Sci Rep* 13: 3302, 2023. doi:10.1038/s41598-023-30265-z.
- Niehorster DC, Li L, Lappe M. The accuracy and precision of position and orientation tracking in the HTC Vive virtual reality system for scientific research. *Iperception* 8: 2041669517708205, 2017. doi:10.1177/2041669517708205.

44. Verdelet G, Desoche C, Volland F, Farnè A, Coudert A, Hermann R, Truy E, Gaveau V, Pavani F, Salemm R. *Assessing Spatial and Temporal Reliability of the Vive System as a Tool for Naturalistic Behavioural Research*. In: *2019 International Conference on 3D Immersion (IC3D)*. Brussels: Belgium, 2019. doi:10.1109/IC3D48390.2019.8975994.
45. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat I, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17: 261–272, 2020 [Erratum in *Nat Methods* 17: 352, 2020]. doi:10.1038/s41592-019-0686-2.
46. Jaekl P, Jenkin M, Harris LR. Perceiving a stable world during active rotational and translational head movements. *Exp Brain Res* 163: 388–399, 2005. doi:10.1007/s00221-004-2191-8.
47. Vaina LM, Beardsley SA, Rushton SK. *Optic Flow and Beyond*. Dordrecht: Springer Science & Business Media, 2004.
48. Redlick FP, Jenkin M, Harris LR. Humans can use optic flow to estimate distance of travel. *Vision Res* 41: 213–219, 2001. doi:10.1016/S0042-6989(00)00243-1.
49. Lappe M, Bremmer F, van den Berg AV. Perception of self-motion from visual flow. *Trends Cogn Sci* 3: 329–336, 1999. doi:10.1016/S1364-6613(99)01364-9.
50. Manning TS, Britten KH. Retinal stabilization reveals limited influence of extraretinal signals on heading tuning in the medial superior temporal area. *J Neurosci* 39: 8064–8078, 2019. doi:10.1523/JNEUROSCI.0388-19.2019.
51. Sunkara A, DeAngelis GC, Angelaki DE. Role of visual and non-visual cues in constructing a rotation-invariant representation of heading in parietal cortex. *eLife* 4: e04693, 2015. doi:10.7554/eLife.04693.
52. Pascucci D, Tanrikulu ÖD, Ozkirci A, Houborg C, Ceylan G, Zerr P, Rafiei M, Kristjánsson Á. Serial dependence in visual perception: a review. *J Vis* 23: 9, 2023. doi:10.1167/jov.23.1.9.
53. Haarmeier T, Bunjes F, Lindner A, Berret E, Thier P. Optimizing visual motion perception during eye movements. *Neuron* 32: 527–535, 2001. doi:10.1016/S0896-6273(01)00486-X.
54. Luna R, Serrano-Pedraza I, Gegenfurtner KR, Schütz AC, Souto D. Achieving visual stability during smooth pursuit eye movements: directional and confidence judgements favor a recalibration model. *Vision Res* 184: 58–73, 2021. doi:10.1016/j.visres.2021.03.003.
55. Manassi M, Whitney D. Illusion of visual stability through active perceptual serial dependence. *Sci Adv* 8: eabk2480, 2022. doi:10.1126/sciadv.abk2480.
56. Castet E, Jeanjean S, Masson GS. Motion perception of saccade-induced retinal translation. *Proc Natl Acad Sci USA* 99: 15159–15163, 2002. doi:10.1073/pnas.232377199.

Head movements contribute to spatial updating across gaze shifts

Manuel Bayer and Eckart Zimmermann

Heinrich-Heine-University Düsseldorf, Germany

Author Note

Manuel Bayer  <https://orcid.org/0000-0002-3717-2813> Correspondence concerning this article should be addressed to Manuel Bayer, Department of Experimental Psychology, Heinrich-Heine-University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany. E-mail: manuel.bayer@hhu.de

Abstract

Keeping visual space constant across movements of the eye and head is a yet not fully understood feature of perception. To understand the mechanisms that update the internal coordinates of space, research has mostly focused on eye movements. However, in natural vision head movements are an integral part of gaze shifts that enlarge the field of vision. Here, we directly compared spatial updating for eye and head movements. In a virtual reality environment, participants had to localize the position of a stimulus across the execution of a gaze shift. We found that performing head movements increased the accuracy of spatial localization. By manipulating the speed of the visual scene displacement that a head movement produced, we found that spatial updating takes into account the sensorimotor contingencies of vision. Traditional accounts of perception during gaze shifts assume that self-produced changes of vision are suppressed. In direct contrast to this theory, we find that self-produced changes in vision are analyzed by the sensorimotor system and used to monitor the displacement vector of the head. We conclude that head movements contribute to stabilizing visual space across gaze shifts and that contingencies of head movements, rather than being canceled, facilitate the updating.

Keywords: space constancy, self-motion, head movement, eye movement

Head movements contribute to spatial updating across gaze shifts

Obtaining information fast and efficiently is crucial for our interaction with the environment. The sensorimotor system displaces the eyes with a frequency of about 3 Hz with high speed in order to enable us to rapidly access information in our environment. These eye movements, called saccades, bring the point of highest resolution, the fovea, onto regions of interest. Gaze shifts come at a cost: The sensorimotor system must distinguish between motion on the retina, produced by the saccade or the head movement, and motion occurring in the external world (Angelaki & Hess, 2005; Cullen, 2019). To this end, the system must know about the amplitude of the performed movements. Electrophysiological studies found that neurons in the lateral intraparietal area (Duhamel et al., 1992) and visual area V3 (Nakamura & Colby, 2002), about 50 ms before the eye starts moving, receive information from spatial locations that would fall into the neuron's receptive field once the saccade has landed. This process has been termed remapping and is likely responsible for creating a transient supra-retinal reference frame at least for a few attended items (Goldberg & Bruce, 1990; Gottlieb et al., 1998; Kusunoki et al., 2000; Sommer & Wurtz, 2004; Szinte & Cavanagh, 2011).

Several behavioral experimental setups have been established to measure the integration of vision across the execution of saccades in the laboratory (Bridgeman, 1983; Honda, 1991; Szinte & Cavanagh, 2011). In a convenient method, observers are asked to compare the position of a stimulus presented before saccade performance against the location of a second stimulus presented after the saccade has been terminated (Szinte & Cavanagh, 2011). Localization across gaze shifts has mostly been studied in saccades, while the head was immobile. However, in larger gaze shifts, saccades are accompanied by head movements that expand the field of vision and align our visual field with the region of interest. Head movements start contributing to the gaze shift when targets are presented at an eccentricity larger than 20° (Freedman & Sparks, 1997). Studies that tested localization across eye-head gaze shifts found evidence for an accurate updating process (Kopinska & Harris, 2003; Mergner et al.,

2001).

In principle, internal knowledge about the saccade vector might be sufficient to bridge the pre- and the post-saccadic visual space. However, head movements could contribute to spatial localization across gaze shifts by potentially adding three signals to the remapping process. Rotational head movement signals are detected by the semicircular canals and processed in vestibular neurons in the cerebellum (Angelaki & Cullen, 2008). Active and passive head movements are distinguished in the response strength of the vestibular neurons. During active movements activation of these neurons is attenuated (Roy & Cullen, 2004). In addition to vestibular signals, neck proprioception information modulates the responses of vestibular neurons. However, this effect cannot be generated by passive head movements. Only if an intended and the actual head movement match the vestibular processing becomes attenuated. These findings suggest the existence of an internal model that predicts the sensory consequences of an upcoming movement and suppresses the actual sensorimotor contingencies if they match the expectation.

Impairing effects of head movements on visual localization seem unlikely given that visual processing is fully functional during their performance, unlike for saccades, during which gaze-induced retinal visual motion is omitted from conscious perception (Binda & Morrone, 2018; Castet et al., 2002). No such omission of the visual input is observed during the performance of head movements. Head movements last on average between 400 and 800 ms (Andres & Hartung, 1989; Hoffmann et al., 2017). Perception cannot afford to be refrained from visual input for such a long period. During head movements, eye movements, including saccades and eye movements that are triggered by the vestibulo-ocular reflex and the optokinetic nystagmus, are performed (Barnes, 1979; Zangemeister et al., 1981). The vestibulo-ocular reflex causes eye movements in the opposite direction of the performed head movement in order to keep the gaze aligned with the orientation of the head. The optokinetic nystagmus on the other hand is a sawtooth movement of the eye, which aims to stabilize a moving image on the retina. An involvement of the head movement in the context of gaze shifts and

localization is also highlighted by the fact that gaze shifts including both types of movement are initially coded as an integrated signal and are only later - downstream of the superior colliculus - separated into individual signals (Freedman et al., 1996; Freedman & Sparks, 1997; Walton et al., 2008).

In the present study, we investigated the accuracy and precision of spatial localization across saccades and combined eye-head movements. We used virtual reality to vary the availability of visual references and visual backgrounds across conditions. Additionally, participants were either free in their execution of the combined eye-head movement or were instructed to perform the gaze shift sequentially, i.e. the saccade first and then the head movement while keeping ocular fixation on the saccade target. This way we manipulated the occurrence of eye movements and the amount of motion perceived during the head movement. In a separate experiment, we further manipulated the background motion by either using a static background or rotating the background against the direction of the performed eye-head movement. This allowed us to explore the use of background motion as a source of information in the context of spatial remapping processes.

Methods

Participants

The sample of the first and second experiment consisted of 35 participants. Due to poor task performance 11 participants had to be excluded from the analyses. The final sample included 18 females and 6 males with a mean age of 24.29 years (24.29 years $\pm$ 0.84 years). The sample of the third experiment consisted of 37 participants. Due to poor task performance 11 participants had to be excluded from the analyses. The final sample included 24 females and 2 males with a mean age of 21.04 years (21.04 years $\pm$ 0.41 years). Every participant gave written informed consent prior to the experiment in accordance with the declaration of Helsinki, participated voluntarily and received either course credit or monetary compensation. This study was approved by the local ethics committee of the mathematical and natural science faculty at Heinrich-Heine-University Düsseldorf (Ethics-approval associated with ERC grant 757184).

Setup

Stimuli were presented by a custom program created with Unreal Engine (version 4.26), running on a Windows 10 desktop computer (Alienware Aurora R8, Intel Core i7 8700 @3.2 GHz, 16 GB RAM, NVIDIA GeForce RTX 2080 graphics card). The used head-mounted display was an HTC Vive Pro with two dual AMOLED 3.5" screens, a resolution of 1440 x 1600 pixels per eye (2880 x 1600 pixels combined), a refresh rate of 90 HZ and a field of view of 110 degrees. The built in 120 HZ eye tracker was used to record the eye position and a custom python app, working at a mean sample rate of 913.68 HZ ($913.68 \text{ HZ} \pm 0.74 \text{ HZ}$), to record the head rotation. The virtual environment was run using SteamVR (version 1.25.3) with the SteamVR 2.0 tracking system. Previous research has shown that the system provides suitable tracking of head and hand positions for research purposes if tracking loss is prevented (Niehorster et al., 2017; Verdelet et al., 2019). All positions of stimuli are given in rotational degree.

Procedure

Experiment 1: Localization across eye-head gaze shifts

The first experiment was divided into eight sessions. Each session started with a short exploration of the virtual environment followed by 10 training trials and 56 experimental trials. The virtual environment of this experiment consisted of a grey background void of any form of reference. All stimuli presented during the experiment were shown 250 cm in depth from the participants point of view. Throughout each trial two crosses ($3.44^\circ \times 3.44^\circ$) were presented, at an eccentricity of 10° to the left and right of the center view. The left cross was the fixation cross while the right was the gaze shift target. At the beginning of each trial participants were instructed to fixate the fixation cross (shown in green color) with their eyes and head. Ocular fixation was counted as successful if the eye remained within an invisible circular window of 1° around the fixation cross. After the fixation cross was fixated successfully for 55 ms, a black dot shaped probe stimulus with a radius of 10 cm was presented either 22° above or below the center view direction for 400 ms. During this period participants were required to maintain fixation at the fixation cross. If they failed to maintain fixation on

the fixation cross, the trial was aborted and restarted. After the stimulus disappeared the fixation cross turned red and the gaze shift target turned green indicating that participants should perform a gaze shift to the target and fixate it with their eyes and head. Participants were explicitly instructed to perform a combined eye-head movement to the target cross. After a successful fixation at the target cross, a black dot shaped reference stimulus, i.e. the visual comparison stimulus was presented on the opposite vertical position of the probe, again either 22° below or above the center view direction for 400 ms (see Figure 1). During this period participants were required to fixate the gaze shift target. The comparison stimulus could appear at 7 different positions ranging from 1.37° to the left to 1.37° to the right of the center view direction. The exact possible shift steps were -1.37° , -0.92° , -0.46° , 0° , 0.46° , 0.92° and 1.37° , negative values indicate a shift to the left. Each position of the comparison stimulus was presented 8 times within a session, resulting in a total of 56 trials. At the end of a trial, participants had to respond if they perceived the comparison stimulus to the left or right of the probe in a two-alternative forced choice task. Participants gave their response by pressing the touchpad of a HTC Vive Controller which they were holding with their dominant hand.

In 8 different experimental sessions, we varied the visual stimulation during the gaze shift. We either presented (i) a homogeneous grey background with stationary fixation and target crosses, (ii) a grey background with fixation and target crosses that were only briefly flashed, (iii) a background consisting of a whole-field grating and stationary fixation and target crosses or (iv) a whole-field grating and flashed fixation and target crosses. These four visual stimulation conditions were presented under two different gaze shift instructions. Participants had to move eye and head either in an unrestricted manner or sequentially. In the unrestricted condition, participants could move their eyes and head in a free, self-paced fashion. In the sequential condition, participants were required to first perform a saccade to the gaze shift target, maintain ocular fixation and then move their head. During the execution of the head movement, they had to keep ocular fixation on the gaze shift target.

Experiment 2: Localization across eye-only gaze shifts

In Experiment 2 subjects were instructed to perform a saccade to the gaze shift target while keeping their head immobile. For this purpose, a chin rest was used. We applied the same four manipulations of visual stimulation as in Experiment 1.

Experiment 3: Background motion in eye-head gaze shifts

In Experiment 3 participants had to perform eye-head gaze shifts. Participants were either unrestricted in their performance of the gaze shift or had to lead the gaze shift with an eye movement followed by a fixation of the gaze shift target with their eyes during the head movement. The fixation cross and gaze shift target were stationary, i.e. presented throughout the whole trial and a grating was presented during the movements from the fixation cross to the gaze shift target. The grating was moving against the direction of the head movement with three visual velocity gains, either as fast as the head movement (gain: 1) or faster (gains: 1.15, 1.3). The required gaze shift amplitude was 20 °. For instance, when a visual velocity gain of 1.15 was active and a head movement of 20 ° was performed, the grating shifted 23 ° to the left, against the direction of the head movement. The order of all conditions within all three experiments was randomized for each participant and counterbalanced across participants. With three visual velocity gains and two gaze shift instructions (unrestricted, sequential), Experiment 3 consisted of 6 conditions.

Head and eye movements

To analyze the head movement data, we first determined the start and end position of the individual movements. The data was analyzed with regard to the rotation around the vertical axis (yaw). Moving averages were used to smoothen the data which took ten consecutively following visual velocity values into account. Afterwards the time point of the peak velocity of the movement was determined. The start and end of the movement were determined by checking when the visual velocity values were below 3 °/s before and after the peak velocity time point. Trials in which the respective participant failed to exceed the visual velocity of 3 °/s or performed a head movement smaller than 10 ° were excluded. These criteria led to the exclusion of

2.72 % of trials across all participants.

To check if the performed head movements were in line with common head movement dynamics reported in the literature we analyzed the amplitude in relation to the achieved peak velocity of the individual head movements in Experiment 1 (see Figure 2a). Participants moved their head further to the sides the higher the achieved peak velocity was during the head movement (one-sample, paired t -test; $t(23) = 12.23$, $p < .001$). The same holds true for the third experiment, participants moved their head further to the sides the higher the achieved peak velocity was during the head movement (one-sample, paired t -test; $t(25) = 13.90$, $p < .001$).

To analyze the eye data, we determined all saccades performed within one respective trial. As the 120 HZ eye tracker did not allow to use standard velocity based detection algorithms, we used an algorithm that utilizes amplitude changes between individual data points irrespective of the time that has passed between the recording of the consecutive data points. We first calculated the differences between individual consecutive positional data points. Saccade onset was defined as the first data point prior to a difference of 1° . Saccade offset was defined as the data point which was recorded at least 30 ms later than the onset and which only deviated less than 0.1° from the previous data point. Based on visual inspection, this procedure yielded appropriate results with regard to the detection of individual saccades performed within a trial.

Results

Experiment 1: Localization across eye-head gaze shifts

Eye and head movements

In Experiment 1, participants had to move their eyes and head either in an unrestricted or in a sequential manner and had to localize the remembered position of a stimulus, seen before the gaze shift. We first analyzed head and eye movement parameters. Participants performed saccades with a mean amplitude of 17.97° ($17.97^\circ \pm 0.38^\circ$). Head movement dynamics followed the main sequence (Zangemeister et al., 1981) with an average peak velocity of $37.01^\circ/\text{s}$ ($37.01^\circ/\text{s} \pm 1.83^\circ/\text{s}$), a mean amplitude of 18.83° ($18.83^\circ \pm 0.46^\circ$) and a duration of 765.40 ms ($765.40 \text{ ms} \pm 28.55$

ms) on average (see Figure 2a). Figure 2b illustrates the sequence of movements performed during these gaze shifts. One can see that participants first performed a saccade and then a head movement to the gaze target.

To investigate the differences between unrestricted and sequential gaze shifts, we analyzed the fixation duration of the gaze shift target during the head movement component of the gaze shifts the participants performed. Figure 3a displays the fixation durations for the different conditions split by the type of gaze shifts. Participants fixated the gaze target for most of the duration of the head movement. Stationary targets led to longer fixations ($89.23 \% \pm 1.31 \%$) compared to flashed targets ($80.68 \% \pm 2.01 \%$; repeated measures ANOVA; $F(1,23) = 18.49$, $p < .001$, $\eta_p^2 = .45$, power = .98). The performance of sequential gaze shifts also increased the fixation duration ($89.64 \% \pm 1.68 \%$) compared to unrestricted gaze shifts ($80.28 \% \pm 1.69 \%$; $F(1,23) = 48.54$, $p < .001$, $\eta_p^2 = .68$, power = .99). There were no other significant effects (all $p \geq .471$).

We next checked for differences in saccade and head movement amplitudes in relation to the available visual references. Figure 3b displays the eye movement (triangle) and head movement amplitudes (circle) for the unrestricted (orange) and sequential gaze shifts (green). One can see that both eye and head movements had roughly the same amplitudes across conditions and that both movements generally undershot the gaze target. This undershoot became larger the fewer visual references were available. Participants performed smaller saccades when the targets were briefly flashed ($17.25^\circ \pm 0.29^\circ$) compared to when they were stationary (repeated measures ANOVA; $18.58^\circ \pm 0.22^\circ$; $F(1,23) = 26.51$, $p < .001$, $\eta_p^2 = .54$, power = .99). The same was the case for head movements, participants performed larger head movements when the targets were stationary ($19.16^\circ \pm 0.27^\circ$) compared to when they were flashed ($18.20^\circ \pm 0.38^\circ$; repeated measures ANOVA; $F(1,23) = 5.59$, $p = .027$, $\eta_p^2 = .20$, power = .62).

There was a significant interaction between the performance of the gaze shift and the target presentation for head movement amplitudes ($F(1,23) = 6.88$, $p = .015$, $\eta_p^2 = .23$, power = .71). Participants performed larger head movements in unrestricted gaze

shifts when the target was stationary ($19.31^\circ \pm 0.38^\circ$) compared to when the target was flashed ($17.69^\circ \pm 0.55^\circ$; $t(23) = 3.12$, $p = .005$). There were no other significant effects for saccade and head movement amplitudes in experiment 1 (all $p \geq .085$).

Psychometric data

We then estimated the horizontal position at which participants perceived the stimulus presented prior and after the gaze shift to be aligned vertically. To this end, we calculated psychometric functions, based on the average responses indicating whether the comparison stimulus was to the left or to the right of the probe stimulus. The horizontal position where the cumulative gaussian function reached 50% was chosen as the point of horizontal alignment (PHA). We calculated the PHA and the just-noticeable difference (JND) of every individual participant in each condition of the first experiment (see Figure 2d and e).

Figure 3c illustrates the PHAs for the individual conditions. We observed more overcompensation the more visual references were available to the participants. Participants overcompensated more for stationary ($0.24^\circ \pm 0.04^\circ$) compared to flashed targets ($0.00^\circ \pm 0.06^\circ$; repeated measures ANOVA; $F(1,23) = 11.00$, $p = .003$, $\eta_p^2 = .32$, power = .89). The presence of a grating as background also led to more overcompensation ($0.21^\circ \pm 0.06^\circ$) compared to a grey background ($0.03^\circ \pm 0.06^\circ$; $F(1,23) = 12.30$, $p = .002$, $\eta_p^2 = .35$, power = .92). There were no other significant results (all $p \geq .161$).

We then determined the discrimination performance by analyzing JNDs. Participants were more sensitive to positional differences when the targets were presented stationary ($0.76^\circ \pm 0.03^\circ$) compared to when they were flashed ($0.96^\circ \pm 0.04^\circ$; repeated measures ANOVA; $F(1,23) = 9.68$, $p = .005$, $\eta_p^2 = .30$, power = .85). There were no other significant results (all $p \geq .325$).

Experiment 2: Localization across eye-only gaze shifts

Eye and head movements

In Experiment 2, we let participants perform only saccades with a fixed head position. Participants performed saccades with a mean amplitude of 16.78° ($16.78^\circ \pm$

0.32 °). The mean head position was at 1.28 ° ($1.28^{\circ} \pm 0.48^{\circ}$) and differed significantly from a rotation of 0, $t(23) = 2.61$, $p = 0.015$. No head movements were detected across all trials performed.

Figure 4a illustrates the amplitudes of eye-only (blue) and eye-head gaze shifts (red). One can see that participants undershot the gaze target in both types of gaze shift and that this undershoot was more pronounced in conditions with fewer visual references. Participants performed smaller saccades when the targets were flashed ($15.96^{\circ} \pm 0.31^{\circ}$) compared to when they were stationary ($17.51^{\circ} \pm 0.29^{\circ}$; repeated measures ANOVA; $F(1,23) = 14.33$, $p = .001$, $\eta_p^2 = .38$, power = .99). There were no other significant results (all $p \geq .154$).

Psychometric data

Next we wanted to compare the localization accuracy between eye-head and eye-only gaze shifts. Figure 4b shows the PHAs for eye-only and unrestricted eye-head gaze shifts for the individual conditions. For both, eye-only and unrestricted eye-head gaze shifts, the compensation accuracy decreased with fewer visual references. Eye-only gaze shifts overall led to more undercompensation, while unrestricted eye-head gaze shifts tended to lead to overcompensation. Participants undercompensated more when they performed eye-only ($-0.26^{\circ} \pm 0.05^{\circ}$) compared to eye-head gaze shifts ($0.10^{\circ} \pm 0.06^{\circ}$; $F(1,23) = 34.73$, $p < .001$, $\eta_p^2 = .60$, power = .99). Flashed targets also led to more undercompensation ($-0.25^{\circ} \pm 0.06^{\circ}$) compared to when they were stationary ($0.09^{\circ} \pm 0.05^{\circ}$; $F(1,23) = 16.60$, $p < .001$, $\eta_p^2 = .42$, power = .99). When a grey background was presented during the gaze shifts, participants also undercompensated more ($-0.15^{\circ} \pm 0.06^{\circ}$) compared to when the background was a grating ($0.00^{\circ} \pm 0.06^{\circ}$; $F(1,23) = 5.09$, $p = .023$, $\eta_p^2 = .20$, power = .64). There were no other significant effects (all $p \geq .145$).

Participants were more precise in their spatial judgement when the targets were stationary ($0.70^{\circ} \pm 0.03^{\circ}$) compared to when they were flashed ($0.90^{\circ} \pm 0.04^{\circ}$; repeated measures ANOVA; $F(1,23) = 8.85$, $p = .007$, $\eta_p^2 = .28$, power = .81). There were no other significant differences (all $p \geq .088$).

Experiment 3: Background motion in eye-head gaze shifts

Eye and head movements

We then asked whether the sensorimotor system might take into account the visual motion that is contingently produced by a head movement. To this end, we artificially moved the background during the execution of head movements. In Experiment 3, participants performed saccades with a mean amplitude of 17.38° ($17.38^\circ \pm 0.26^\circ$). Head movement dynamics followed the main sequence for head movements (Zangemeister et al., 1981) with an average peak velocity of $48.14^\circ/\text{s}$ ($48.14^\circ/\text{s} \pm 2.63^\circ/\text{s}$), a mean amplitude of 20.25° ($20.25^\circ \pm 0.32^\circ$) and a duration of 940.73 ms ($940.73\text{ s} \pm 44.63\text{ ms}$) on average.

First, we analyzed the duration participants fixated the gaze shift target during their head movement. Figure 5a displays the fixation duration for the two types of gaze shift and the different visual velocity gains. Also in this experiment participants fixated the gaze target for most of the duration of the head movement. In sequential eye-head gaze shifts, participants almost fixated the gaze target for the whole head movement duration. Participants fixated the right fixation cross longer during their head movement in sequential ($91.72\% \pm 0.89\%$) compared to unrestricted gaze shifts ($75.12\% \pm 0.88\%$; repeated measures ANOVA; $F(1,25) = 73.76$, $p < .001$, $\eta_p^2 = .75$, power = .99). The application of a visual velocity gain also had an impact on the fixation duration ($F(2,50) = 5.19$, $p = .009$, $\eta_p^2 = .17$, power = .95). Bonferroni corrected post hoc dependent t -tests were able to reveal a significant difference between the unity visual velocity gain and a visual velocity gain of 1.3 ($t(25) = 2.88$, $p = .024$). Participants fixated the right fixation cross during the head movement for a longer duration while unity visual velocity gain was active ($84.63\% \pm 1.57\%$) compared to a visual velocity gain of 1.3 ($81.83\% \pm 1.56\%$). There were no other significant differences regarding the fixation duration (all $p \geq .080$).

Figure 5b shows the saccade and head movement amplitudes participants performed for the two types of gaze shift and the different visual velocity gains. We observed larger amplitudes in head movements compared to saccades. Unrestricted gaze

shifts produced the largest amplitudes across all movements. Little to no differences were observed for amplitudes in relation to the applied visual velocity gains. There were no significant differences between the amplitudes of the performed saccades across conditions (all $p \geq .061$) but participants performed larger head movements in unrestricted ($20.93^\circ \pm 0.26^\circ$) compared to sequential gaze shifts ($19.53^\circ \pm 0.23^\circ$; $F(1,25) = 13.02$, repeated measures ANOVA; $p = .001$, $\eta_p^2 = .34$, power = .99; see Figure 5b). There were no other significant effects for the performed head movements (all $p \geq .216$).

Psychometric data

To test if the performance of the gaze shift had an effect on the localization accuracy we performed regression analyses. We fitted linear functions to the visual velocity gain values and their respective PHAs on the single subject level. This procedure was performed separately for unrestricted and sequential gaze shifts. Figure 5c illustrates the mean PHAs for the two types of gaze shift and the individual conditions. One can see that for both gaze shifts and all visual velocity gains participants overcompensated for their movements. This overcompensation was even more pronounced in sequential eye-head gaze shifts. Participants overcompensated more when they performed sequential ($0.26^\circ \pm 0.08^\circ$) compared to unrestricted gaze shifts ($0.08^\circ \pm 0.08^\circ$; paired t -test; $t(24) = -2.90$, $p = .008$). Additionally, we compared the slopes of the unrestricted and restricted movement condition against 0 in order to test how the application of visual velocity gains influenced localization accuracy. Participants' localization accuracy was influenced by the applied visual velocity gain when they performed unrestricted gaze shifts ($t(24) = 2.03$, $p = .027$) but not when they performed sequential gaze shifts ($t(24) = 0.63$, $p = .268$). We performed the same regression analyses for the localization sensitivity but did not find significant differences (all $p \geq .203$).

Since we found a difference in localization accuracy between unrestricted and sequential gaze shifts in Experiment 3 but not in Experiment 1, we calculated the difference in the fixation duration of the saccade target during the head movement

between unrestricted and sequential gaze shifts for each participant for both experiments. There was a bigger difference in the fixation duration during the head movement between unrestricted and sequential gaze shifts in Experiment 3 ($16.60\% \pm 1.90$) compared to Experiment 1 ($9.36\% \pm 1.32\%$; unpaired t -test; $t(24) = -3.02$, $p < .001$).

Discussion

In this study, we found that performing head movements contributes to the accuracy in spatial localization across gaze shifts. In natural vision, head movements are an integral part of gaze shifts. On the one hand, the head-movement component complicates the updating of visual space across gaze shifts. In an eye-head gaze shift two reference frames, that of the eye and that of the head, move and thus must be compensated for. On the other hand, the position of the head can be used by the sensorimotor system to measure the size of the gaze shift. Four signals provide information about the amplitude of a head-movement and can contribute to spatial updating. Three of these signals are exclusively internal, i.e. changes in vestibular and proprioceptive states and the efference copy which is a copy of the motor command that drives the head movement. The fourth signal consists in the self-produced visual motion on the retina that originates from the relative displacement between the head-movement and the external visual scene.

Vestibular signals inform about head position extremely fast with a latency of only 14 ms (Lisberger, 1984). Vestibular and neck proprioception input is likely combined to decode head position, as the convergence of both signals is required for posture, balance and vestibular spinal reflexes (Crowell et al., 1998; Harris, 1994; Mergner et al., 1992). Like for saccades, an efference copy of the head movement amplitude could also be involved in spatial localization of eye-head gaze shifts. The efference copy is likely involved in the build up of an internal model predicting the sensory consequences, i.e. vestibular and proprioceptive changes, following the performance of head movements. The interaction of these three signals could provide predictively available, precise information of the head movement size. Combining this

information with the internal estimate of the upcoming saccade vector might improve the accuracy of spatial updating across a gaze shift. In addition, even the visual motion that is produced by the head movement entails information about the head movement amplitude. The sensed visual motion velocity could be read out to determine the actual head movement speed, allowing to detect deviations from the predicted head movement amplitude. Instead of suppressing this latter signal, as suggested by traditional theories (Binda & Morrone, 2018; Campbell & Wurtz, 1978), the sensorimotor system uses the information to monitor the spatial extent of the head movement.

We first directly compared spatial updating in eye and in combined eye-head gaze shifts. When participants performed saccades without moving their head, localization was accurate when a background was shown and the fixation and the saccade target were permanently visible. When we decreased the availability of visual references by showing no background or in addition by only flashing the fixation cross and the saccade target in the pure saccade condition, we found that participants underestimated the location of the pre-saccadic stimulus. They undershot the target with their saccades and they also indicated the position of the localization stimulus closer to the fixation point. The underestimation in trans-saccadic localization might be a consequence of the saccadic undershoot. If a saccade fails to reach the target and the sensorimotor system updates visual space according to the executed saccade vector, underestimation in trans-saccadic localization would be the consequence. However, studies have shown that the sensorimotor system is aware of the saccade errors even before the saccade is initiated (Murthy et al., 2007; Tian et al., 2013). The updating process should therefore take into account the saccade landing. Furthermore, results from our lab have shown that visual localization also drifts toward the fovea even during ocular fixation when visual references are absent (Cont & Zimmermann, 2021). The underestimation of the pre-saccadic stimulus location in the present study and the accompanied saccade undershoot in conditions in which visual references were absent might thus have the same origin, that is uncertain information about external space. Indeed, previous findings have demonstrated that uncertain target information leads to

saccadic undershoot (Lisi et al., 2019).

In Experiment 3, we measured spatial localization across an eye-head gaze shift when the background either was stationary or when it moved against the direction of the head movement. The latter condition served to increase the experience of background motion. We compared localization performance when observers could freely execute an eye-head gaze shift or when they were required to keep their eye direction fixated on the target cross while executing the head movement. We found overestimation of the target as a function of background displacement velocity in both conditions. However, the overestimation was significantly stronger when subjects were required to keep their eye direction fixated on the target cross. A moving grating induces an optokinetic nystagmus if observers do not maintain fixation. The nystagmus stabilizes the grating on the retina and thus partly cancels out the motion experience. Under natural conditions, i.e. when the background moved with unity gain and when they performed unrestricted head movements, participants localized the probe accurately. These results reveal that it is the visual motion registered during the execution of a head movement that contributes to updating visual space in eye-head gaze shifts. Vestibular and proprioceptive signals that code the location of the head after movement termination are unlikely to be in charge of spatial updating. Neither did head movement amplitudes show any modulation by background motion nor was their distribution suited to explain where subjects localized the probe stimulus.

In a previous study we showed that subjects are sensitive to deviations between the expected and the actual head movement - contingent visual motion velocity (Bayer & Zimmermann, 2023). We found that even after a single experience of such a deviation, subjects shifted their expectation about the motion velocity occurring during head movement execution. The need to maintain the expectations about movement-contingent motion accurate could lie in their contribution to spatial updating. Head movements arise in a gaze shift when desired objects have an eccentricity of 20° or more (Freedman & Sparks, 1997). Targets at that distance will necessarily provide an uncertain visual signal as they fall on peripheral retinal locations.

We found in the present study that uncertain visual signals generate saccades that undershoot their targets combined with undercompensation of visual space. A gaze shift consisting of an eye and a head movement component solves the problem since first, the combined gaze movement is more accurate and second, the visual motion produced by the head movement is used by the sensorimotor system to update visual space. We conclude that the self-produced movement-contingent visual motion during a head rotation contributes to updating visual space across gaze shifts.

References

- Andres, R. O., & Hartung, K. J. (1989). Prediction of head movement time using fitts' law. *Human Factors*, 31(6), 703–713.
- Angelaki, D. E., & Cullen, K. E. (2008). Vestibular system: The many facets of a multimodal sense [PMID: 18338968]. *Annual Review of Neuroscience*, 31(1), 125–150. <https://doi.org/10.1146/annurev.neuro.31.060407.125555>
- Angelaki, D. E., & Hess, B. J. M. (2005). Self-motion-induced eye movements: Effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6(12), 966–976. <https://doi.org/10.1038/nrn1804>
- Barnes, G. R. (1979). Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets. *The Journal of Physiology*, 287(1), 127–147. <https://doi.org/10.1113/jphysiol.1979.sp012650>
- Bayer, M., & Zimmermann, E. (2023). Serial dependencies in visual stability during self-motion [PMID: 37465870]. *Journal of Neurophysiology*, 130(2), 447–457. <https://doi.org/10.1152/jn.00157.2023>
- Binda, P., & Morrone, M. C. (2018). Vision during saccadic eye movements. *Annual Review of Vision Science*, 4(1), 193–213. <https://doi.org/10.1146/annurev-vision-091517-034317>
- Bridgeman, B. (1983). Mechanisms of space constancy. In A. Hein & M. Jeannerod (Eds.), *Spatially oriented behavior* (pp. 263–279). Springer New York. https://doi.org/10.1007/978-1-4612-5488-1_15
- Campbell, F. W., & Wurtz, R. H. (1978). Saccadic omission: Why we do not see a grey-out during a saccadic eye movement. *Vision Research*, 18(10), 1297–1303. [https://doi.org/https://doi.org/10.1016/0042-6989\(78\)90219-5](https://doi.org/https://doi.org/10.1016/0042-6989(78)90219-5)
- Castet, E., Jeanjean, S., & Masson, G. S. (2002). Motion perception of saccade-induced retinal translation. *Proceedings of the National Academy of Sciences*, 99(23), 15159–15163. <https://doi.org/10.1073/pnas.232377199>
- Cont, C., & Zimmermann, E. (2021). The motor representation of sensory experience. *Current Biology*, 31(5), 1029–1036.e2. <https://doi.org/10.1016/j.cub.2020.11.032>

- Crowell, J. A., Banks, M. S., Shenoy, K. V., & Andersen, R. A. (1998). Visual self-motion perception during head turns. *Nature neuroscience*, 1(8), 732–737. <https://doi.org/10.1038/3732>
- Cullen, K. E. (2019). Vestibular processing during natural self-motion: Implications for perception and action. *Nature Reviews Neuroscience*, 20(6), 346–363. <https://doi.org/10.1038/s41583-019-0153-1>
- 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(5040), 90–92. <https://doi.org/10.1126/science.1553535>
- Freedman, E. G., Stanford, T. R., & Sparks, D. L. (1996). Combined eye-head gaze shifts produced by electrical stimulation of the superior colliculus in rhesus monkeys. *Journal of Neurophysiology*, 76(2), 927–952. <https://doi.org/10.1152/jn.1996.76.2.927>
- Freedman, E. G., & Sparks, D. L. (1997). Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys. *Journal of Neurophysiology*, 77(5), 2328–2348. <https://doi.org/10.1152/jn.1997.77.5.2328>
- Goldberg, M. E., & Bruce, C. J. (1990). Primate frontal eye fields. iii. maintenance of a spatially accurate saccade signal. *Journal of neurophysiology*, 64(2), 489–508.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391(6666), 481–484.
- Harris, L. R. (1994). Visual motion caused by movements of the eye, head and body. *Visual detection of motion*, 1, 397–435.
- Hoffmann, E. R., Chan, A. H., & Heung, P. (2017). Head rotation movement times. *Human factors*, 59(6), 986–994.
- Honda, H. (1991). The time courses of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. *Vision Research*, 31(11), 1915–1921. [https://doi.org/https://doi.org/10.1016/0042-6989\(91\)90186-9](https://doi.org/https://doi.org/10.1016/0042-6989(91)90186-9)
- Kopinska, A., & Harris, L. (2003). Spatial representation in body coordinates: Evidence from errors in remembering positions of visual and auditory targets after active

- eye, head, and body movements. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 57(1), 23.
<https://doi.org/10.1037/h0087410>
- Kusunoki, M., Gottlieb, J., & Goldberg, M. E. (2000). The lateral intraparietal area as a salience map: The representation of abrupt onset, stimulus motion, and task relevance. *Vision research*, 40(10-12), 1459–1468.
- Lisberger, S. (1984). The latency of pathways containing the site of motor learning in the monkey vestibulo-ocular reflex. *Science*, 225(4657), 74–76.
- Lisi, M., Solomon, J. A., & Morgan, M. J. (2019). Gain control of saccadic eye movements is probabilistic. *Proceedings of the National Academy of Sciences*, 116(32), 16137–16142. <https://doi.org/10.1073/pnas.1901963116>
- Mergner, T., Rottler, G., Kimmig, H., & Becker, W. (1992). Role of vestibular and neck inputs for the perception of object motion in space. *Experimental Brain Research*, 89(3), 655–668. <https://doi.org/10.1007/BF00229890>
- Mergner, T., Nasios, G., Maurer, C., & Becker, W. (2001). Visual object localisation in space. interaction of retinal, eye position, vestibular and neck proprioceptive information. *Experimental brain research*, 141 1, 33–51.
- Murthy, A., Ray, S., Shorter, S. M., Priddy, E. G., Schall, J. D., & Thompson, K. G. (2007). Frontal eye field contributions to rapid corrective saccades [PMID: 17135479]. *Journal of Neurophysiology*, 97(2), 1457–1469.
<https://doi.org/10.1152/jn.00433.2006>
- Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences*, 99(6), 4026–4031. <https://doi.org/10.1073/pnas.052379899>
- Niehorster, D. C., Li, L., & Lappe, M. (2017). The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research. *i-Perception*, 8. <https://doi.org/10.1177/2041669517708205>
- Roy, J. E., & Cullen, K. E. (2004). Dissociating self-generated from passively applied head motion: Neural mechanisms in the vestibular nuclei. *Journal of*

Neuroscience, 24(9), 2102–2111.

<https://doi.org/10.1523/JNEUROSCI.3988-03.2004>

Sommer, M. A., & Wurtz, R. H. (2004). What the brain stem tells the frontal cortex. ii. role of the sc-md-fef pathway in corollary discharge. *Journal of neurophysiology*, 91(3), 1403–1423.

Szinte, M., & Cavanagh, P. (2011). Spatiotopic apparent motion reveals local variations in space constancy. *Journal of Vision*, 11(2), 4–4.

Tian, J., Ying, H. S., & Zee, D. S. (2013). Revisiting corrective saccades: Role of visual feedback. *Vision Research*, 89, 54–64.

<https://doi.org/10.1016/j.visres.2013.07.012>

Verdelet, G., Desoche, C., Volland, F., Farnè, A., Coudert, A., Hermann, R., Truy, E., Gaveau, V., Pavani, F., & Salemme, R. (2019). Assessing spatial and temporal reliability of the vive system as a tool for naturalistic behavioural research. *2019 International Conference on 3D Immersion (IC3D)*, 1–8.

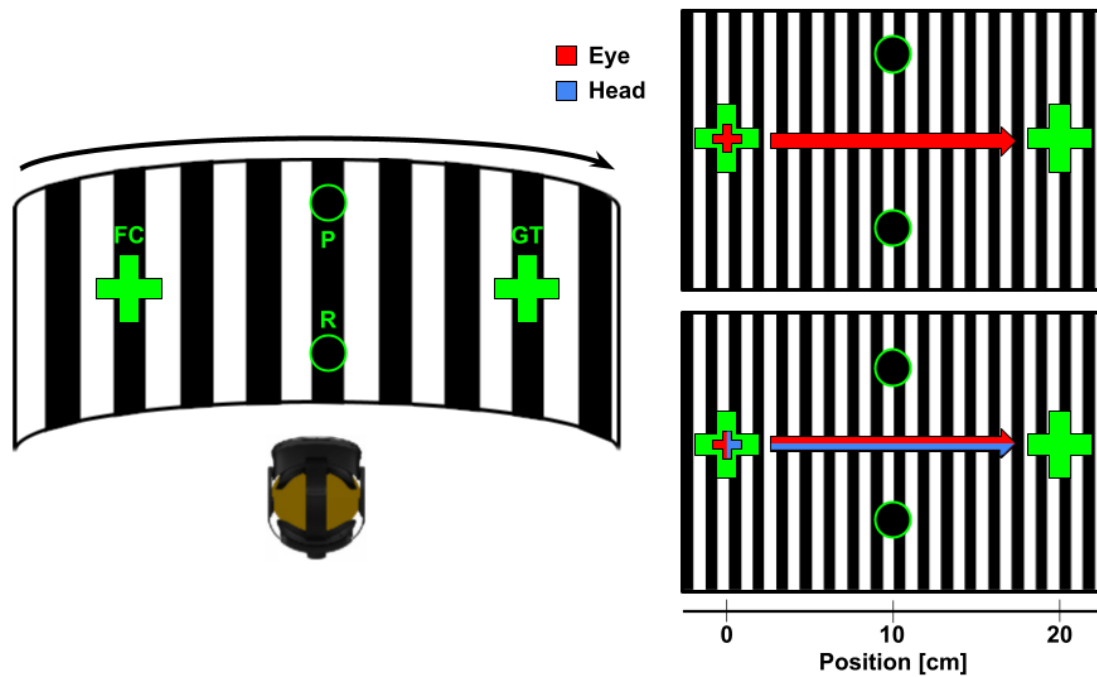
<https://doi.org/10.1109/IC3D48390.2019.8975994>

Walton, M. M. G., Bechara, B., & Gandhi, N. J. (2008). Effect of reversible inactivation of superior colliculus on head movements. *Journal of Neurophysiology*, 99(5), 2479–2495. <https://doi.org/10.1152/jn.01112.2007>

Zangemeister, W. H., Lehman, S., & Stark, L. (1981). Simulation of head movement trajectories: Model and fit to main sequence. *Biological Cybernetics*, 41(1), 19–32. <https://doi.org/10.1007/BF01836124>

Figure 1

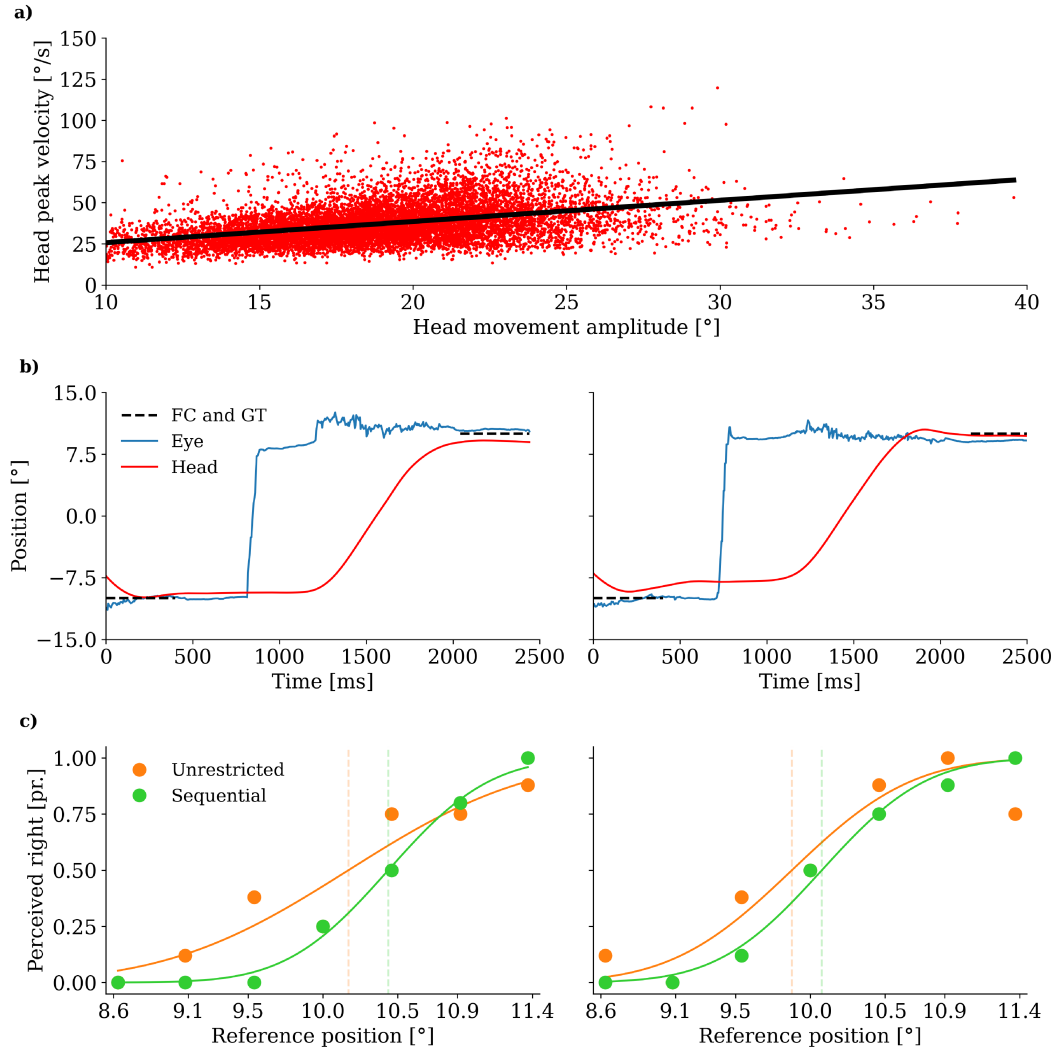
Experimental setup



Note. Illustration of the experimental setup. Participants performed either eye-only or eye-head gaze shifts from the fixation cross (FC) to the gaze shift target (GT). Participants had to judge the position of the comparison stimulus (C), presented after the gaze shift, against the probe stimulus (P), presented prior to the gaze shift. Please note, the displayed grating was only visible during the gaze shift and not during the presentation of the stimuli.

Figure 2

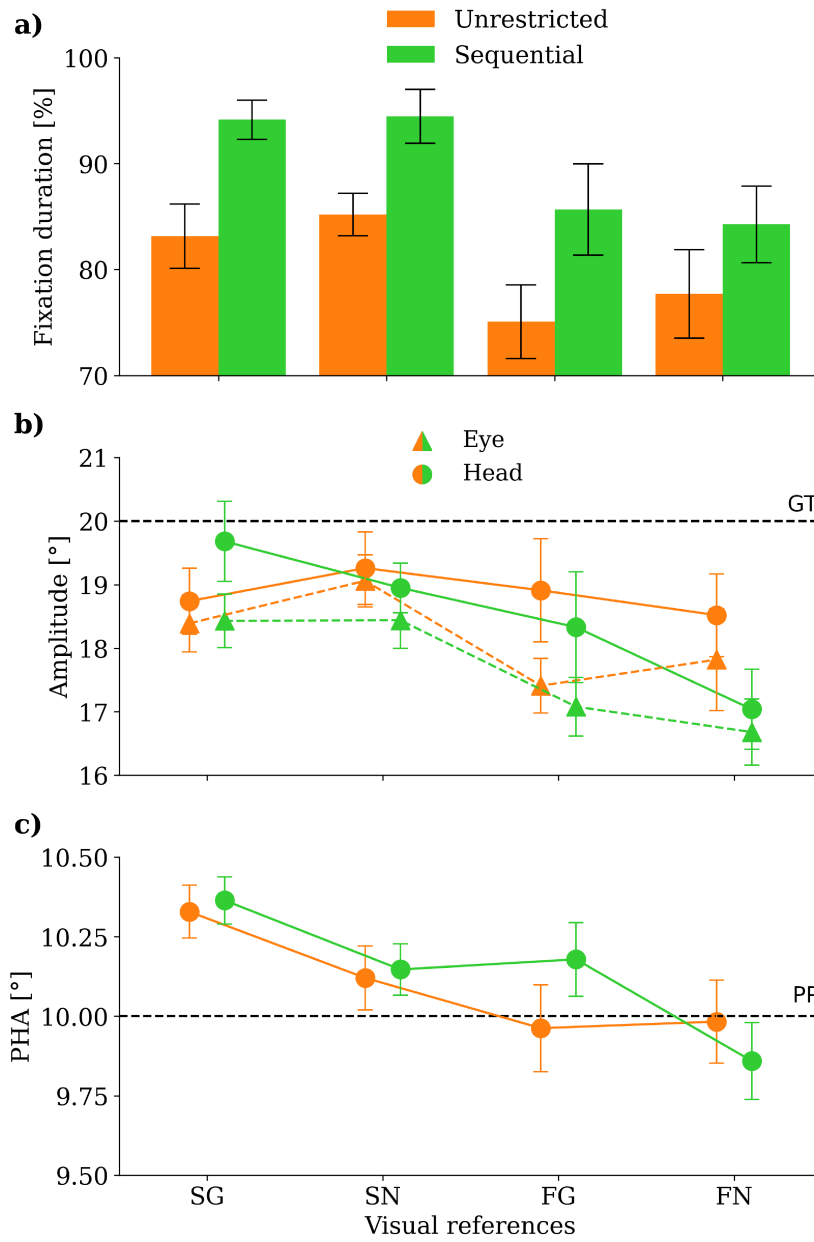
Head movement main sequence and examples for the movement sequence and psychometric functions



Note. **a)** Average head movement main sequence, illustrating the linear relationship ($y = 13.22 + 1.27x$) between head movement amplitude and peak velocity. Data were collapsed across trials and participants in the first experiment. **b)** Gaze shift traces for the head movement (shown in red) and the eye movement component (shown in blue) from an example trial in which the participant performed an unrestricted (left) and sequential gaze shift (right). The dashed lines represent the position of the fixation cross and gaze shift target on the horizontal axis. The eye moves to the target first, followed by the rotation of the head. **c)** Psychometric functions of two representative participants for the stationary target - grating condition split by the eye movement restriction during the head movement.

Figure 3

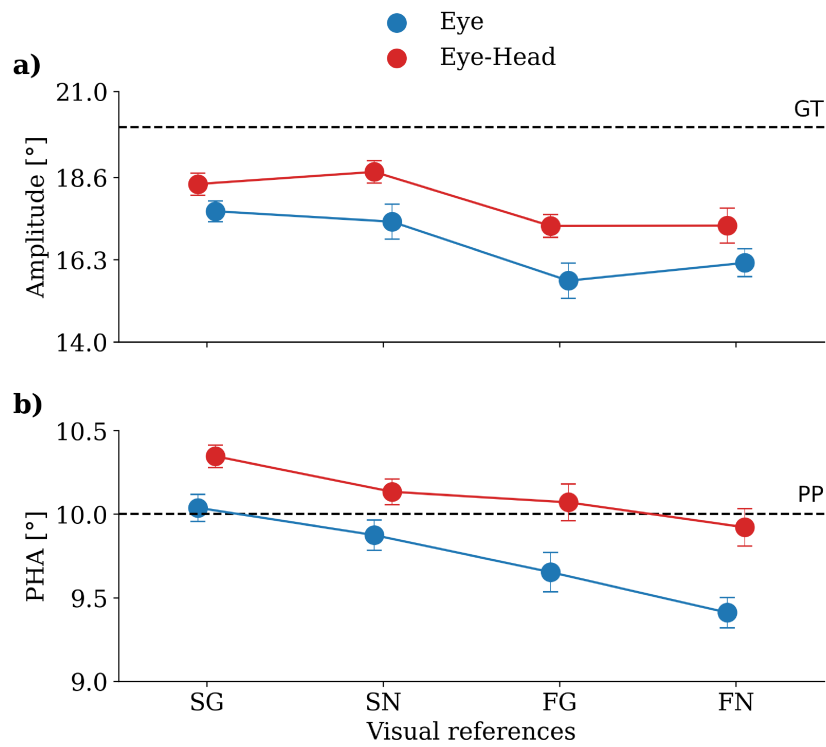
Unrestricted and sequential eye-head gaze shifts



Note. **a)** Duration of the fixation of the gaze shift target during the execution of the head movement component of the gaze shift for the stationary target - grating (SG), stationary target - grey background (SN), flashed - grating (FG) and flashed - grey background condition (FN) split by the eye movement restriction during the head movement. **b)** Saccade and head movement amplitudes for the individual conditions. The dashed line represents the position of the gaze shift target (GT). Same conventions as in a). **d)** PHAs for the individual conditions. The dashed line represents the position of the probe (PP). Same conventions as in a).

Figure 4

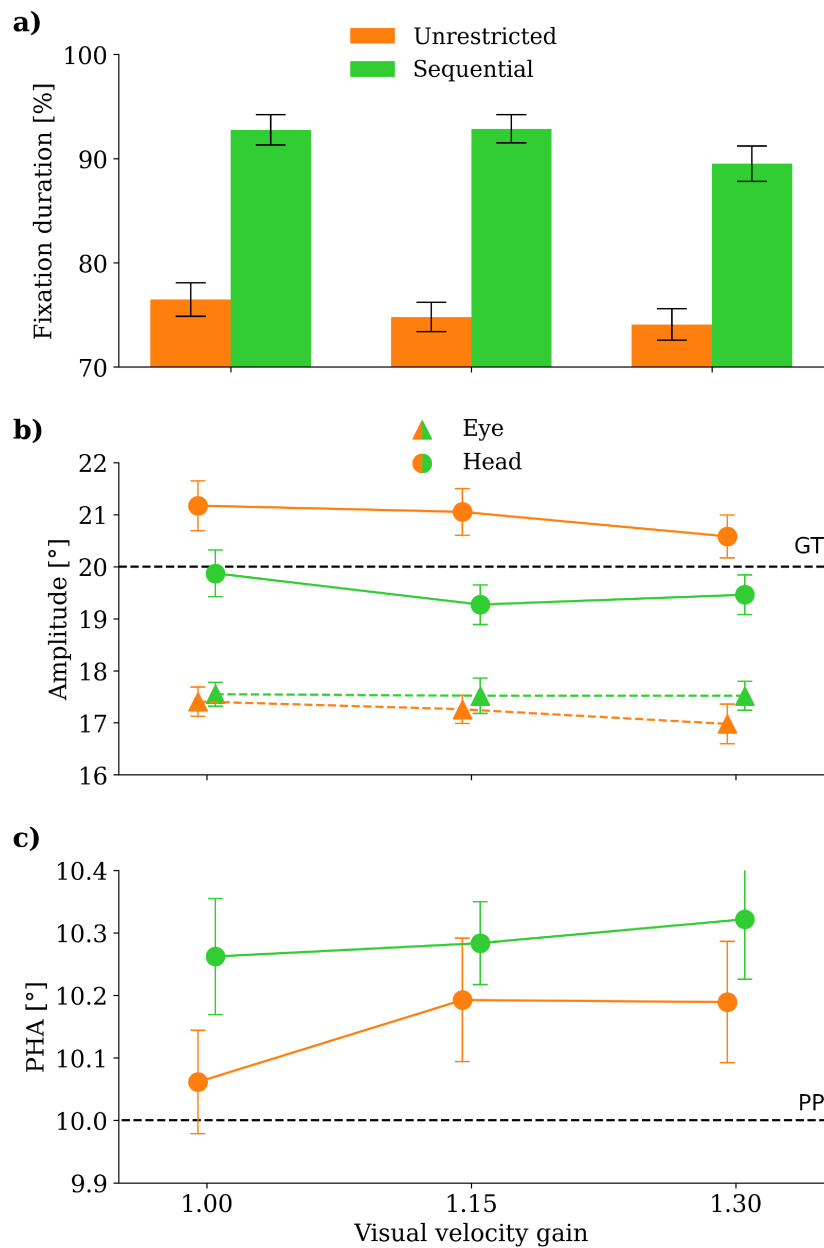
Eye versus eye-head gaze shifts



Note. **a)** Saccade amplitudes for the stationary target - grating (SG), stationary target - grey background (SN), flashed - grating (FG) and flashed - grey background condition (FN) of the first and second experiment split by color. The dashed line represents the position of the gaze shift target (GT). **b)** PHAs for the individual conditions of the first and second experiment. The dashed line represents the position of the probe (PP). Same conventions as in a).

Figure 5

Eye-head gaze shifts and the role of background motion



Note. **a)** Duration of the fixation of the gaze shift target during the execution of the head movement component of the gaze shift for the individual visual velocity gains split by the eye movement restriction during the head movement. **b)** Saccade and head movement amplitudes for the individual visual velocity gains. The dashed line represents the position of the gaze shift target (GT). Same conventions as in a). **c)** PHAs for the individual visual velocity gains. The dashed line represents the position of the probe (PP). Same conventions as in a).