# From retinotopic processing to nonretinotopic representation

THÈSE N$^O$ 8126 (2017)

PRÉSENTÉE LE 13 DÉCEMBRE 2017
À LA FACULTÉ DES SCIENCES DE LA VIE
LABORATOIRE DE PSYCHOPHYSIQUE
PROGRAMME DOCTORAL EN NEUROSCIENCES

## ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

## Marc Michael LAUFFS

acceptée sur proposition du jury:

Prof. B. D. McCabe, président du jury
Prof. M. Herzog, directeur de thèse
Prof. D. Kerzel, rapporteur
Dr K. Hamburger, rapporteur
Prof. F. Hummel, rapporteur

*(EPFL*
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

# Acknowledgements

# Abstract

Encoding of visual information in the brain is *retinotopic:* Neighboring points in the visual field are mapped onto neighboring photoreceptors in the retina, and these neighborhood relations are maintained in the early stages of cortical processing. However, perception is *nonretinotopic*. First, vision is stable and continuous although the retinal image is in constant flux and frequently interrupted by eye blinks. Second, object parts are perceived relative to the object, rather than in retinal coordinates. For example, the reflector on the wheel of a moving bicycle is perceived to circle, although it follows a cycloidal trajectory across the retina. The precise mechanisms of how, where, and when visual information is transformed from a retinotopic into a nonretinotopic representation, remain largely elusive. While nonretinotopic perception across eye movements is relatively well understood, almost nothing is known about nonretinotopic processing without eye movements. Here, I used psychophysical experiments to uncover the perceptual mechanisms that are at play. My results show that A) the computation of nonretinotopic perception does not require predictability of the stimulus, suggesting that the computations are very fast; B) unconscious processing of retinotopic motion can interfere with the conscious nonretinotopic motion percept; C) the computation of nonretinotopic motion is likely a two-stage process, in which first a reference motion is computed, that is then subtracted from the retinotopic motion; D) nonretinotopic motion perception is intact in schizophrenia patients; E) endogenous attention affects nonretinotopic feature integration primarily by increasing the resolution of the initial stimulus encoding.

1743/3499 characters incl. spaces

# Keywords

Vision, perception, nonretinotopic processing, Ternus-Pikler display, apparent motion, illusions, consciousness, attention.

# Zusammenfassung

Die Verarbeitung von visuellen Informationen im Gehirn ist zunächst *retinotop*: Benachbarte Punkte im Gesichtsfeld werden auf benachbarte Fotorezeptoren in der Netzhaut projiziert, deren Nachbarschaftsverhältnisse in frühen Stadien der kortikalen Verarbeitung erhalten bleiben. Wahrnehmung ist jedoch *nichtretinotop*: Zum einen ist das Gesehene stabil und kontinuierlich, obwohl das retinale Abbild in ständiger Bewegung ist und durch das Blinzeln der Augen unterbrochen wird. Zum anderen werden Teilobjekte relativ zum Objekt wahrgenommen, anstatt in retinalen Koordinaten. Zum Beispiel wird die Bewegung des Speichenreflektors eines Fahrrads als kreisförmig wahrgenommen, obwohl er sich entlang einer zykloiden Trajektorie über die Netzhaut bewegt. Die genauen Mechanismen wie, wo, und wann visuelle Information von einer retinotopen in eine nichtretinotope Darstellung umgewandelt werden ist weitestgehend unbekannt. Ich habe psychophysikalische Experimente genutzt um die Wahrnehmungsmechanismen zu untersuchen die hierbei eine Rolle spielen. In dieser Dissertation beschreibe ich einige der Determinanten und Einschränkungen dieser Verarbeitung. Meine Ergebnisse zeigen, A) dass die Berechnung nichtretinotoper Wahrnehmungen nicht auf Vorhersehbarkeit des Reizes angewiesen ist, was darauf hindeutet, dass sie extrem schnell erfolgen kann; B) dass die unbewusste Verarbeitung von retinotoper Bewegung mit der bewussten Wahrnehmung nichtretinotoper Wahrnehmung interferieren kann; C) dass die Berechnung nichtretinotoper Bewegung wahrscheinlich ein zweistufiger Prozess ist, in dem zunächst eine Referenzbewegung berechnet und diese dann von der retinalen Bewegung subtrahiert wird; D) dass nichtretinotope Bewegungswahrnehmung in Schizophreniepatienten intakt ist; E) dass endogene Aufmerksamkeit nichtretinotope Merkmalsfusion primär über die Auflösung der ursprünglichen Reizwahrnehmung beeinflusst.

1904/3499 Zeichen inkl. Leerzeichen

## Schlüsselwörter

# Contents

# Preface

In this thesis I present eight research projects that I worked on as doctoral candidate at EPFL. At the time of writing, two of these projects are published, two are submitted for publication, two are in preparation for submission, and two are ongoing.

A list of these projects is provided below. Published articles and pre-prints of the submitted manuscripts are provided in the appendix with permission of the copyright holders.

<u>Nonretinotopic processing</u>

Lauffs, M. M., Öğmen, H., & Herzog, M. H. (2017). Uncertainty does not hamper nonretinotopic motion perception. *Journal of Vision*, 19(7):6, 1-10.

Lauffs, M. M., Shaqiri, A., Brand, A., Roinishvili, M., Chkonia, E., Öğmen, H., & Herzog, M. H. (2016). Local versus global and retinotopic versus nonretinotopic motion processing in schizophrenia patients. *Psychiatry Research, 246*, 461-465.

Lauffs, M. M., Choung, O. H., Öğmen, H., & Herzog, M. H. (*submitted*). Unconscious retinotopic motion processing affects nonretinotopic motion perception.

Hochmitz, I., Lauffs, M. M., Herzog, M. H., & Yeshurun, Y. (*in preparation*). Sustained spatial attention affects feature fusion through enhanced signal encoding.

Lauffs, M. M., Öğmen, H., & Herzog, M. H. (*ongoing*). Retinotopic perception despite group motion at intermediate ISI durations.

<u>Other fields</u>

Ballhausen*, N., Lauffs*, M. M., Herzog, M. H., & Kliegel, M. (*submitted*). Investigating prospective memory via the eyes: No evidence for a monitoring deficit in older adults. (*equal contributions)

Lauffs, M. M., Geoghan, S. A., Favrod, O., Herzog, M. H., & Preuschoff, K. (*in preparation*). EEG correlates of risk prediction error magnitude.

Lauffs M. M., Karacsony, D., & Herzog, M. H. H. (*ongoing*). What makes a masterpiece? Surprisingly low inter-observer gaze-path similarity when viewing Renaissance art.

# Chapter 1    Introduction

## 1.1    Visual information processing in the brain

Light is reflected by an object, transverses the optical apparatus of the eye and falls on the photoreceptors in the retina. Photoreceptors hyperpolarize according to the amount of light they receive, which changes the amount of neurotransmitters they release. Bipolar and ganglion cells pool this information and encode it in the firing of action potentials. This information leaves the eye via long axons, that interface with cells in the lateral geniculate nucleus (LGN) in the thalamus. It ultimately reaches the primary visual cortex (striate cortex, V1) in a largely unprocessed form. In cortex, processing then proceeds hierarchically in functionally specialized areas along two primary routes, the dorsal and the ventral stream (Figure 1; e.g., Ungerleider & Mishkin, 1982; Goodale & Milner, 1992; Grill-Spector & Malach, 2004). The dorsal stream is fast and processes information related to the location of objects, allowing for rapid response action. For example, neurons in the middle temporal cortex are specialized in the perception of motion. The ventral stream is slower and performs a more detailed analysis of the objects in the visual field. For example, cells in infero-temporal cortex respond selectively to faces and cells in the lateral occipital complex to objects.

## 1.2    Retinotopic processing

The early stages of visual processing are organized in a *retinotopic* fashion: Similar to the workings of a photo camera, the optics of the eye map an image of the 3-dimensional world onto the 2-dimensional retina (Figure 2). Neighboring points in the visual field map onto neighboring photoreceptors of the retina. These neighborhood relations are preserved in the projections of retinal cells to the LGN and the early visual cortices. Hence, the early visual cortices contain a retinotopic representation of the visual field, i.e., a representation of the visual scene in the layout of the retinal image (Figure 3; e.g., Wandell, Dumoulin, & Brewer, 2008; Engel, Glover, & Wandell, 1997). Retinotopic maps have also been found in higher visual areas (Arcaro, McMains, Singer, & Kastner, 2009; and Saygin & Sereno, 2008), but are more difficult to investigate, because receptive fields increase in size and overlap as information travels up the visual hierarchy.
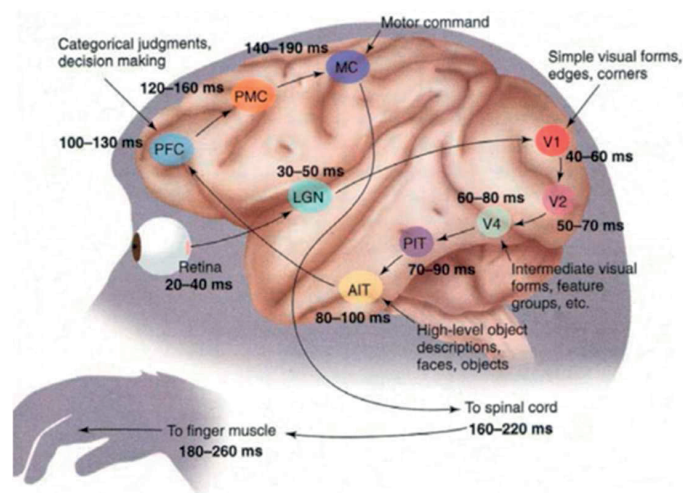
*Figure 1 An example for the functional specializations of human visual cortex: A plausible processing circuit in a visual categorization task. Visual information from the retina arrives via the lateral geniculate nucleus in cortical area V1, where luminance steps (edges) are detected. Higher visual areas in the ventral visual stream (e.g., V4, inferotemporal cortex) respond selectively to intermediate-level properties, such as shapes, and high-level properties, such as faces. Adapted from Thorpe, S.J. & Fabre-Thorpe, M. (2001) Seeking categories in the Brain. Science 291(5502) 260-263. Reprinted with permission from AAAS.*

## 1.3    Nonretinotopic perception: The problem of visual stability

Despite the retinotopic organization of large parts of visual cortex, perception is usually *nonretinotopic*. For example, as we interact with visual information around us, our eyes move rapidly to focus the fovea, the part of the retina where vision is sharpest, on different parts of the visual field. Most of these eye movements are *saccades*, ballistic eye movements of very high velocity (with maximum velocities of 750-1000 °/sec among the fastest movements the human body can perform). During saccades, the retinal image shifts rapidly and abruptly. This does not leave enough time to fully process the image, which should lead to a "smeared" percept (Öğmen, 2007; Öğmen & Herzog, 2010). But perception is sharp and crisp. Retinal input is interrupted during eye blinks and we are virtually blind during saccades due to saccadic suppression (e.g., Matin, 1974; Riggs, Merton, & Morton, 1974; Bridgeman, Hendry, & Stark, 1975; Burr, Morrone, & Ross, 1994; Ross, Burr, & Morrone, 1996). However, perception is continuous, uninterrupted. Every saccade entails a drastic displacement of the retinal image. Even during fixations, micro-saccades keep the eye in constant motion. However, perception is stable: Although the retinal image moves, we correctly perceive the world as static. In addition to the retinotopic representation that characterizes the early-stage processing in the brain, there must hence exist a more stable nonretinotopic representation of the world. How is this nonretinotopic representation computed? How does the brain extract a stable representation of the world from the ever changing retinal image? And how is this information stored and represented in the brain?
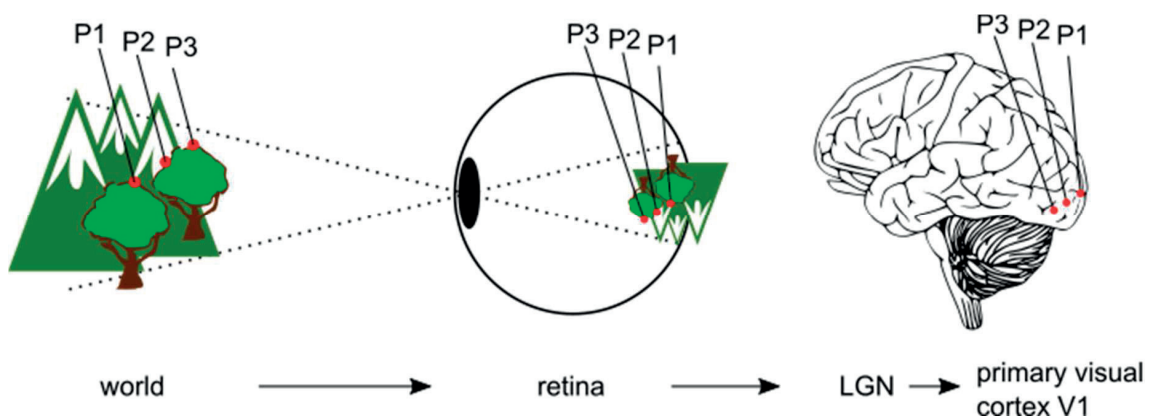


*Figure 2 Schematic explanation of retinotopic coding. The optics of the eye map neighboring points in the field of view onto neighboring photoreceptors in the retina. These neighborhood relations are maintained in the cells of the lateral geniculate nucleus (LGN) of the thalamus and early visual areas in cortex. P1, P2, P3, are three arbitrarily chosen points in the visual field, whose spatial inter-relations are preserved up to the cells in early visual cortex.*
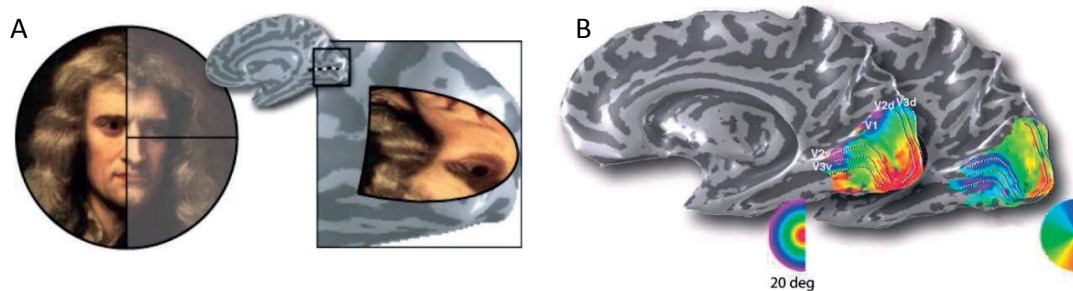
*Figure 3 Retinotopic mapping in the human brain. A) An illustration of how the visual field is transformed and mapped to primary visual cortex V1. In the cortical representation, the image is inverted and the central part of the visual field is overrepresented (cortical magnification), but the spatial relations of the original image are maintained (retinotopy). B) Example of retinotopic V1, V2, and V3 visual field maps in the right hemisphere of a human observer, localized with functional magnetic resonance imaging (fMRI). Stimulation at different eccentricities and polar positions in the visual field activates corresponding parts of visual cortex. Reprinted from Wandell, Dumoulin, & Brewer (2007) Visual field maps in human cortex. Neuron 56(2), 366-383, with permission from Elsevier.*

*Visual stability* across saccades is probably the best-investigated nonretinotopic phenomenon. The visual system seems capable of telling apart whether a displacement of the retinal image is caused by an eye movement or a change in the outer world, because only the latter is perceived as motion. Why are eye movements not perceived as motion, although the image on the retina shifts? Other than motion in the world, eye movements are intentional, i.e., they are controlled by the observer. The visual system hence "knows" in advance that an eye movement will occur. This makes it possible to correct the percept for the portion of retinal motion that is caused by the eye movement, so that only motion in the outer world is perceived. The special role of volition was already recognized by Descartes (1664) and Von Helmholtz (1867). They described perceiving motion when moving the eyeball by applying external pressure with a finger, but not when making a normal, voluntary eye movement. This suggests, that there is a mechanism that corrects perception for shifts of the retinal image that are cause by the own intentional eye movements, but not other factors.

Such a mechanism was independently conceptualized by von Holst and Mittelstädt (1950) and Sperry (1950). Von Holst and Mittelstädt (1950) observed that flies continue turning in circles after their heads had been rotated upside-down. Normally, turning to the left causes the retinal image to displace to the right, and vice versa. Von Holst and Mittelstädt hypothesized, that this expectation is represented in the action-and-perception circuits of the brain in the form of *efference copies.* They conceived of efference copies as analogs of the motor signals that cause the movement (hence the term *efference* copies). Perceptual areas receive this signal and use it to predict the change of the retinal image the movement will cause. This prediction is then discounted from the shift of the retinal image, allowing for stable visual perception, although the eye movement drastically changes the retinal image.

Since in the experiment the flies' heads were rotated upside-down, the animal experienced the retinal image to shift in the direction opposite of what it intended, and hence of what the efference copy predicted. To compensate, the animal turned more in the same direction, causing the retinal image to shift more in the opposite direction, and so forth, leading to the observed circling pattern. Similar observations were made by Sperry (1950) who rotated the eyeballs of fish and called these signals *corollary discharges*. The terms 'efference copy' and 'corollary discharge' are functionally equivalent and used interchangeably in the literature. In humans, evidence for the existence of efference copy-like signals was found in self-experiments, in which the experimenter paralyzed the own eye-muscles, or even underwent complete paralysis (Stevens et al., 1976). Although the eye could not move due to the paralysis, attempted saccades resulted in a "non-visual" perception of displacement in the direction of the attempted saccade. This was taken as evidence for an influence of the motor intent on perception, that is compatible with the theory of efference copies.

Through neurophysiological recordings in monkeys, an exact neural pathway has been identified, that is compatible with the concept of efference copies and their effect on perception (Cavanaugh, Berman, Joiner, & Wurtz, 2016; Sommer & Wurtz, 2008; Wurtz, 2008). This pathway runs from the superior colliculus in the midbrain, via the medial-dorsal part of thalamus, to the Frontal Eye Fields in cortex (Sommer & Wurtz, 2008). Importantly, signaling along this pathway is not only linked to the occurrence of eye movements, but *precedes* them. This is in line with the idea that efference copies *predict* an eye movement before it occurs.

Signaling along this pathway is thought to drive the shifting of receptive fields during saccade preparation, known as *pre-saccadic remapping*. Remapping of receptive fields has been observed in the lateral intraparietal area (Duhamel, Colby, & Goldberg, 1992), superior colliculus (Walker, Fitzgibbon, & Goldberg, 1995), frontal eye fields (Umeno & Goldberg, 1997; Zirnsak, Steinmetz, Noudoost, Xu, & Moore, 2014), and other areas (e.g., Nakamura & Colby, 2002; Tolias, Moore, Smirnakis, Tehovnik, Siapas, & Schiller, 2001; but see Cavanagh, Hunt, Afraz, & Rolfs, 2010; Rolfs, Jonikaitis, Deubel, & Cavanagh, 2011). The activity of a single brain-cell is recorded in an awake, behaving monkey. The receptive field of the cell is found by consecutively probing all parts of the visual field with light-flashes. The cell only responds to stimuli in its receptive field, but not to stimuli elsewhere (Figure 4a). Then, the monkey saccades to a second fixation point. Usually, the receptive fields have fixed retinotopic positions, so the receptive field displaces in accordance with the point of gaze (Figure 4b). Surprisingly, shortly before the saccade takes place, the cell already responds to stimuli in the *future* position of its receptive field, although the eye has not moved, yet (Figure 4c). Hence, processing of visual information at the future location begins before the eye has started to move, allowing for a smooth transition from the pre-saccadic to the post-saccadic image that might explain visual stability. More recent data even suggest that receptive fields are not immediately remapped to their future location, but converge on the saccade target first (Zirnsak et al., 2014; Zirnsak & Moore, 2014; but see Rolfs & Szinte, 2016). Concurrent

theories posit that not the receptive field is remapped, but attention (Cavanagh et al, 2010; Rolfs et al., 2011, 2016). In these theories, efference copies remain crucial, because they provide the prediction of the upcoming saccade.
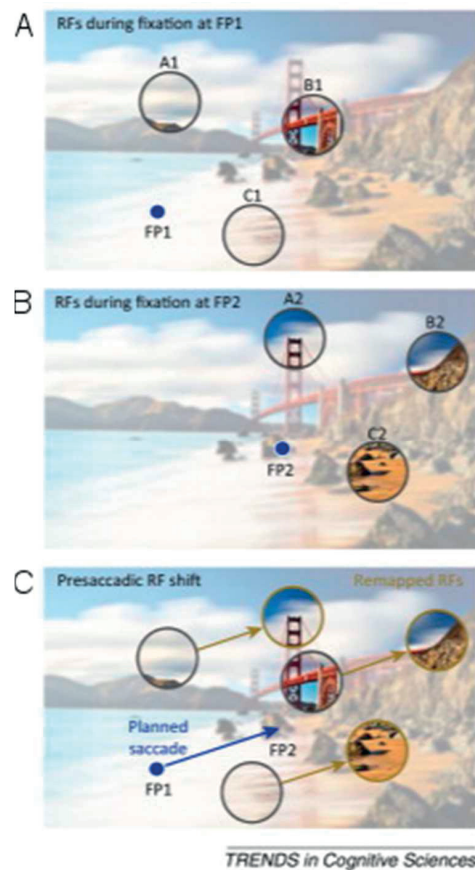


*Figure 4 The concept of shifting receptive fields. A) A visual neuron is typically activated by stimulation in only a subpart of the visual field, its receptive field. The position of the receptive field is defined in retinotopic coordinates and hence dependent on the current fixation position. A1, B1, and C1 are the receptive fields of cells A, B, and C, when the gaze is directed at fixation position 1 (FP1). B) The receptive fields stay in the same retinotopic position relative to fixation when the gaze is directed at a different part of the image, i.e. a second fixation position (FP2). C) Immediately preceding a saccadic eye movement, cells in many visual areas respond to stimulation in parts of the retina that will only be covered by their receptive fields after the saccade. This suggests that the receptive fields are remapped pre-saccadically, i.e. before the eye movement occurs. More recent evidence suggests that receptive fields initially converge on the future fixation position. Reprinted from Zirnsak & Moore (2014) Saccades and shifting receptive fields: anticipating consequences or selecting targets? Trends in Cognitive Sciences, 18(12), 621-628, with permission from Elsevier*

## 1.4 Nonretinotopic perception without eye movements

Efference copies have been found to be the single crucial explanation of visual stability across saccades (Wurtz, 2008). Other factors, such as information about the current orientation of the eye (proprioception), and the fact that the *entire* visual field moves (*optic flow*), are thought to play secondary roles at best (review: Wurtz, 2008; Cavanaugh et al., 2016).

However, nonretinotopic perception also occurs in the absence of eye movements, that is, when there are no motor signals and hence no efference copies. For example, moving objects are generally perceived relative to other objects, rather than in retinal coordinates (Johansson, 1950, 1974, 1976; Ağaoğlu, Herzog, & Öğmen, 2015; Ağaoğlu, Clarke, Herzog, & Öğmen, 2016). In the same way, the parts of an object are perceived relative to the object itself. For example, when a bicycle passes through the visual field, the reflector on its wheel appears to circle, although it follows a cycloidal trajectory across the retina (Figure 5; Johansson, 1950; Duncker, 1929). The reflector motion is hence perceived nonretinotopically. The retinotopic, cycloidal trajectory cannot be perceived, because the nonretinotopic, circular motion percept dominates. This percept arises because the reflector is perceived relative to the bicycle, rather than in retinal coordinates. The bicycle serves as a moving reference system, whose linear motion across the retina is subtracted from the cycloidal reflector motion. Mathematically, this is similar to the discounting of an efference copy from a retinal shift. The percept is fundamentally the same, whether the bicycle is tracked with the eyes or not. Hence, it cannot be explained by efference copies.



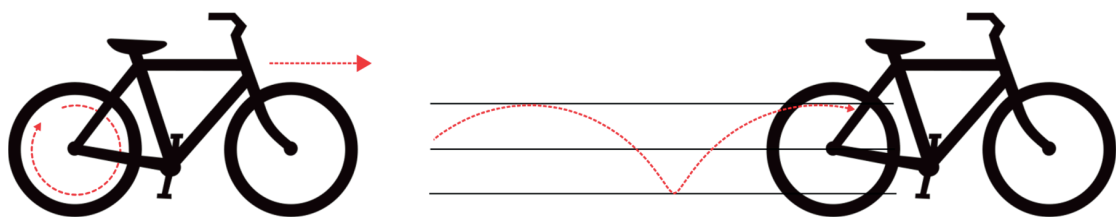*Figure 5 Nonretinotopic perception of a bicycle reflector. Moving parts of an object are perceived relative to the object. The reflector of a moving bicycle is perceived as circling (left) even though the motion is cycloidal in retinotopic coordinates (right). The motion of the bicycle is perceptually subtracted from the cycloidal motion, yielding circular motion. Reprinted from Lauffs, Öğmen, & Herzog (2017).*

## 1.5   Nonretinotopic processing in the Ternus-Pikler display

Based on this observation, Boi and colleagues (2009) have developed a paradigm that allows the study of nonretinotopic processing in the absence of eye movements. Their version of the *Ternus-Pikler display* is based on a perceptual phenomenon first described by Pikler (1917) and studied in detail by his student Ternus (1926). Two horizontally aligned elements (e.g., black disks) are briefly presented in the center of the screen. They reappear in the same positions after an inter-stimulus interval (ISI). After another ISI, the sequence starts over (Figure 6a). Two stationary, flickering disks are perceived (*no motion* percept). When a third disk is added alternately to the left and to the right, two concurrent percepts are possible: At very short ISIs (e.g., 0-50 ms), the two original disks are perceived to flicker on and off stationarily, and the third disk appears to jump back and forth between their left and right (*element motion* percept; Figure 6b). However, at longer ISIs (e.g., 200 ms or more) the percept is very different: The three disks form a perceptual group that appears to move left and right in concert (*group motion* percept; Figure 6c). When the observer's gaze is stable, the screen-based position of each element matches its position in *retinotopic* coordinates. However, the group motion percept is *nonretinotopic*, because the perceived object correspondence is not in line with the screen-based and hence retinotopic positions of the elements.
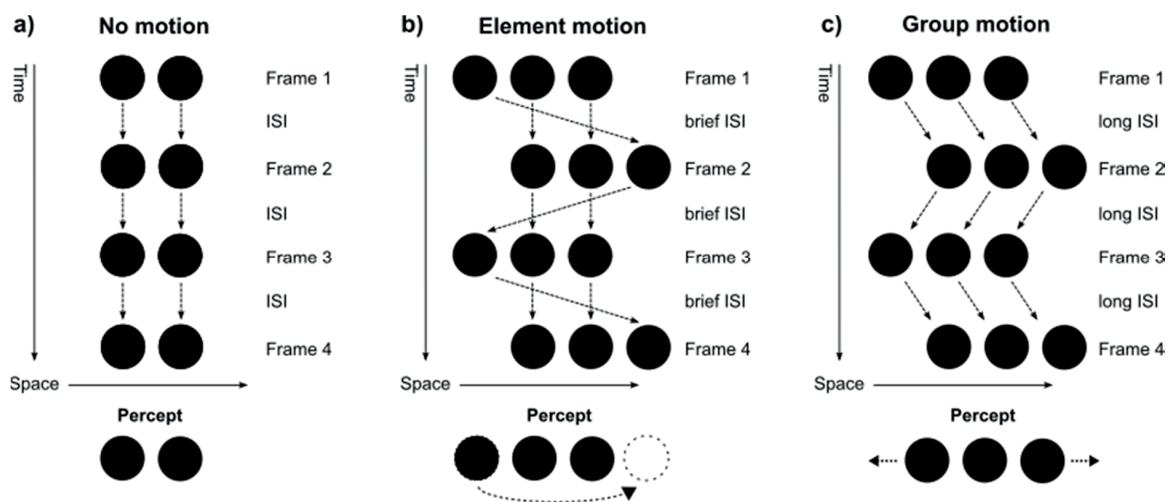


*Figure 6 The Ternus-Pikler display. A) Two disks are briefly presented (e.g., 200 ms) and reappear in the same positions after an inter-stimulus interval (ISI). Two stationary flickering disks are perceived. B) When a third disk is added alternately left and right in every frame, two concurrent percept are possible. At very brief ISI durations (e.g., 0 ms) two stationary flickering disks are perceived, with the third element jumping from the left to the right (element motion). C) At longer ISIs (e.g., 200 ms) three disks are perceived to move left-and-right in concert (group motion). The percept does not correspond to the retinotopic position of the individual disks and is, hence, nonretinotopic. Dashed lines indicate the perceived object correspondence and were not presented during the experiment. © CC-BY 4.0.*

When a feature is added to the objects, it can easily be determined whether this feature is processed in retinotopic or nonretinotopic coordinates, by comparing its perception when the objects are perceived to move (group motion -> nonretinotopic percept) and when they appear static (no or element motion -> retinotopic percept). Since almost any visual feature can be added, the Ternus-Pikler display is an extremely versatile tool to directly pit retinotopic against nonretinotopic processing in the absence of eye movements. Most processes investigated with the Ternus-Pikler display were found to be nonretinotopic. These include the perception of motion (Boi et al., 2009), orientation (Wutz, Drewes, & Melcher, 2016), feature search (Boi et al., 2009), and even exogenous attention (Boi, Vergeer, Öğmen, & Herzog, 2011). On the other hand, motion after-effects (Boi et al., 2009; Boi, Öğmen, & Herzog, 2011) and tilt after-effects (Boi, Öğmen, & Herzog, 2011) were found to be retinotopic.

Of particular importance to my own work is the dot-motion Ternus-Pikler display introduced by Boi and colleagues (2009). White dots were added to the black disks of a Ternus-Pikler display. From one stimulus presentation to another, the white dots were displaced, so as to create an apparent motion percept of up-and-down motion in one, and left-and-right motion in the other disk (no motion; Figure 7a). Then a third disk with central dot was added alternately to the left and right. When the ISI was very brief (e.g., 0 ms) the third disk appeared to jump back and forth and the up-down, left-right dot motions were seen in the two stationary disks (element motion; Figure 7b).
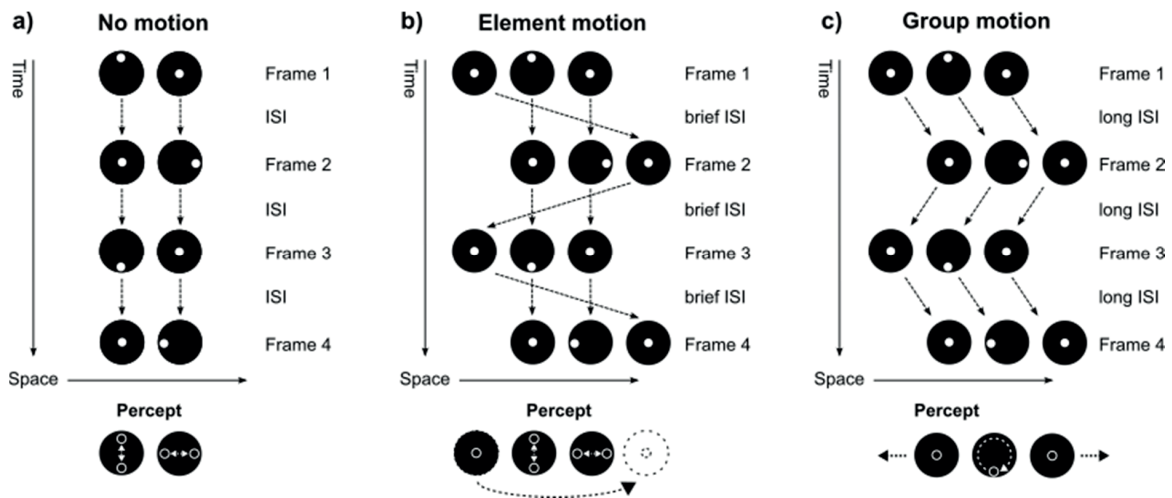


*Figure 7 Dot motion Ternus-Pikler display after Boi et al. (2009). A) Each disk carries a white dot that is displaced from frame to frame. The dot positions are chosen to create an apparent up-down motion in one, and left-right motion in the other disk (retinotopic dot motion). B) A third disk with central dot is added alternately to the left and right. When element motion is perceived (brief ISIs), the retinotopic motion is perceived in the two central disks. C) When group motion is perceived (long ISIs), the retinotopic up-down and left-right motion is perceptually discounted for the left-right motion of the group. As a result, the dot in the middle disk is perceived to rotate (nonretinotopic motion percept). The rotation percept is owed to the different integration of the left-right and up-down dot motions. © CC-BY 4.0.*

However, at longer ISIs (e.g., 200 ms) the three disks appeared to move left-and-right in concert and the dot in the middle disk appeared to rotate (group motion; Figure 7c). The dots in the outermost disks appeared to be always in the center. The dot rotation is due to a nonretinotopic integration of the retinotopic up-down and left-right dot motions, that are discounted for the left-right motion of the disks. The dot rotation is nonretinotopic, because the percept is relative to the perceived motion of "its" disk. Hence, motion, or at least apparent motion, is processed in nonretinotopic coordinates. Importantly, the image on the retina is the same in all cases (cf. Figure 7a-c). All that changes is the *interpretation* of the retinal image, depending on the context of the third disk and the timing.

## 1.6   A two-stage model of nonretinotopic motion computation

How the visual system computes nonretinotopic percepts is not known. It has been proposed, that a strictly retinotopic one-stage model based on spatio-temporal filters can account for nonretinotopic motion perception in the Ternus-Pikler display (Pooresmaeili, Cicchini, Morrone, & Burr, 2012). However, there are serious problems with this model (Clarke, Repnow, Öğmen, & Herzog, 2013; reply: Pooresmaeili, Cicchini, Morrone, & Burr, 2013). A two-step approach involving a nonretinotopic processing stage seems more appropriate. Such a framework has been proposed by Öğmen and Herzog (2010) and implemented in computer models by Clarke, Öğmen, and Herzog (2016), and Ağaoğlu, Clarke, Herzog, & Öğmen (2016).

In the first stage, objects in the visual field are detected and their retinotopic motion vectors are computed. In the case of the Ternus-Pikler display, this requires solving the *object correspondence problem*: When the disks reappear after the ISI, the visual system has to establish which element in the previous frame corresponds to which element in the current frame. The difference between the previous and the current position of each element gives its motion vector in retinotopic coordinates. A computational solution to this problem was proposed by Dawson (1991).

In the second stage, nonretinotopic motion is computed from the retinotopic motion vectors (Figure 8). First, objects are grouped based on their proximity, the similarity of their retinotopic motion, and other factors. For each group, the average motion of its members is computed (global motion). Finally, this motion is subtracted from the retinotopic motion of each element, yielding its motion relative to the group (local motion). Hence, the global motion provides the reference frame for the perception of the nonretinotopic local motion. Only the local, relative, nonretinotopic motion is perceived, and the absolute, retinotopic motion is invisible. For example, in the Ternus-Pikler display, the retinotopic dot motion is discounted for the left-and-right group motion. Only the nonretinotopic circular dot motion is perceived, while the retinotopic up-down and left-right motions are invisible. A similar vector based approach to global and local motion perception was proposed by Johansson (1973) for biological motion perception, in which for example the arm is perceived relative to the torso.
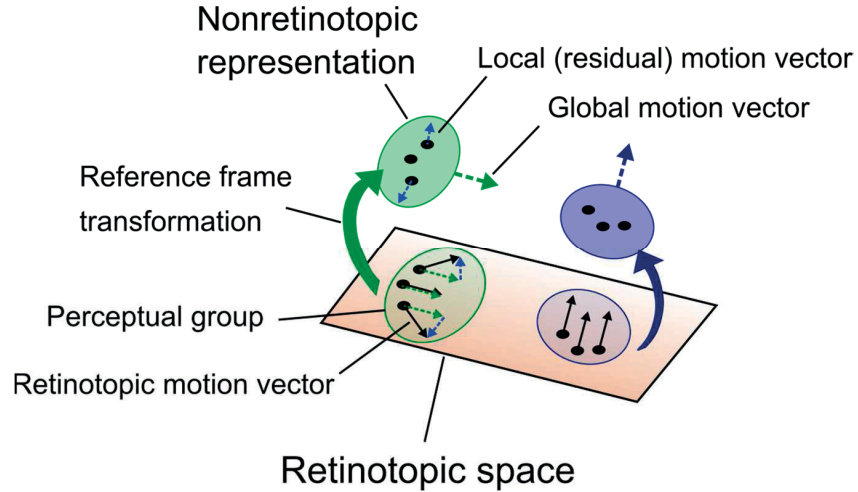
*Figure 8 Vector field approach to nonretinotopic motion perception. Green: Objects that move across the retina are grouped based on proximity, motion similarity, and other factors. Their retinotopic motion vectors are decomposed into their shared (global) and residual (local) motion components. A nonretinotopic representation of the group is created by subtracting the global component from the retinotopic motion vectors. For each object, only the local motion is perceived within the moving group. The global motion of the group serves as the reference frame for the computation of the nonretinotopic motion of its elements. Only the nonretinotopic representation is perceived, and the retinotopic motion vectors are invisible. Purple: The three objects' motion vectors are identical. After subtracting the average global motion, there is no residual local motion. The objects are perceived as stationary within the moving group. Idea by Öğmen & Herzog (2010) and Clarke et al. (2016). © CC-BY 4.0.*

## 1.7   Neural correlates of nonretinotopic processing

Very little is known about nonretinotopic computations at the neural level. Thunell and colleagues (Thunell, Plomp, Öğmen, & Herzog, 2016a) investigated the time course of nonretinotopic processing in the Ternus-Pikler display using electroencephalography (EEG). Three disks were always presented in the same position (retinotopic condition; no motion), or displaced to the left and right with every frame, creating the perception of group motion (nonretinotopic condition; group motion). In both conditions, a notch in the middle disk started and stopped rotating at random time points of a prolonged trial. In the condition with static disks the perception of rotation requires *retinotopic* integration. In the group motion condition, the rotation percept requires *nonretinotopic* integration across retinal positions.

The EEG data were analyzed in epochs of a single stimulus frame, i.e., a single *static* image. The EEG traces looked very similar in the group motion and no motion conditions. This is surprising, because the disks occupy different retinal positions in the two conditions. With

both the no motion and group motion stimulus, there were strong EEG differences between the static and rotating notch conditions. In the first frame after the notch transitioned from static to moving or vice versa, activity was stronger for the moving than the static notch around 220 ms post stimulus onset. Previous studies suggested that event-related potentials (ERPs) in this time-range signal motion onsets (Kuba, Kubova, Kremlacek, & Langrova, 2007). This interpretation is also compatible with source reconstructions, that pointed to a neural origin in the motion sensitive middle temporal cortex (MT). In the last frame *before* the rotation stopped or started, that is, when the static or rotation percept was already established, stronger activity for the rotating notch conditions was found as early as 120 ms after stimulus onset. Again there were no significant differences between the no motion and group motion conditions. This suggests that the computation of nonretinotopic notch rotation was completed as early as 120 ms after stimulus onset, allowing for a processing of the notch motion that is largely independent of whether the reference system is static or moving.

Using functional magnetic resonance imaging (fMRI), Thunell and colleagues (Thunell, van der Zwaag, Öğmen, Plomp, & Herzog, 2016b) factored in on *where* in the brain these computations take place. Instead of rotating dots or notches, they used squares that were half black and half white. From frame to frame, the squares inversed their contrast polarity, creating the illusion that the black half moved on a white background. The contrast polarity reversal happened either in retinotopic or nonretinotopic coordinates. When the reversal took place in retinotopic coordinates, the reversal was perceived when the squares were perceived as stationary (no motion), but not when they were perceived as moving (group motion). When the reversal took place in nonretinotopic coordinates, it was perceived in the group motion, but not the no motion trials. This factorial design allowed the authors to disentangle the retinotopic from the nonretinotopic processing of the stimulus. They found that the blood oxygen level dependent (BOLD) response in early visual areas V1/V2/V3 reflected the retinotopic encoding of the stimulus. However, activity in the motion sensitive middle temporal cortex reflected whether the stimulus was *perceived* as inversing, irrespective of whether the inversion was taking place in retinotopic or nonretinotopic coordinates. Hence, the nonretinotopic percept must have been computed in area MT or before. Both the EEG and fMRI evidence therefore indicates that the nonretinotopic percept is computed at an early stage of visual processing, both in terms of timing and hierarchical organization of the brain.

## 1.8   Summary and research goals

Although the early visual areas of the brain are organized in retinotopic coordinates, vision is typically nonretinotopic: We perceive the world as stable, although the retinal image changes drastically several times per second, with every eye and body movement. The brain computes this stable, nonretinotopic, percept of the world, by discounting the retinal shift for an efference copy that predicts the eye movement. However, nonretinotopic perception is at least as prevalent in situations without eye movements: The motion of object parts is typically perceived relative to the entire object. For example, the reflector on the wheel of a moving bicycle is perceived to circle, although its motion is cycloidal on the retina.

Surprisingly, the topic has received very little attention by the scientific community, although nonretinotopic perception is clearly a fundamental part of vision and a ubiquitous phenomenon. Consequently, almost nothing is known about how the brain computes nonretinotopic perception from its retinotopic input. It appears that elements in the visual scene are grouped based on their motion similarity. The motion of each element across the retina is then discounted for the common motion of the group, which is similar to the discounting of efference copies. As a result, only the nonretinotopic motion of the element is perceived, that is, its local, relative motion within the group.

This kind of processing can explain nonretinotopic perception in theory, but the exact mechanisms remain elusive: What determines which elements are grouped? What is the timeframe of these computations? Is predictability of the object required, as in visual stability? And what is the fate of the retinotopic motion percept? It may be clear that we're only beginning to understand how the brain computes nonretinotopic motion.

In the following chapter, I discuss my contributions to the continued investigation of nonretinotopic perception. The primary aim of my work was to achieve a better understanding of the mechanisms and determinants of nonretinotopic processing. With a focus on the example of nonretinotopic motion perception, I explored conditions under which nonretinotopic perception fails and compared these to the predictions of current models. In particular, I tested whether predictability of the reference object is crucial for the successful computation of nonretinotopic motion (Chapter 2.1), whether unconscious processing of the invisible retinotopic motion can affect the conscious nonretinotopic motion percept (Chapter 2.2), what the temporal limits of this computation are (Chapter 2.3), and whether nonretinotopic processing is intact in schizophrenia patients (Chapter 2.4). In addition, I explored the influence of attention on nonretinotopic feature integration (Chapter 2.5).

# Chapter 2  Results

## 2.1  Uncertainty does not hamper nonretinotopic motion perception

*This subchapter summarizes the publication*
*Lauffs, M. M., Öğmen, H., & Herzog, M. H. (2017)*
*Uncertainty does not hamper nonretinotopic motion perception. Journal of Vision, 19(7):6, 1-10.*
*(Appendix 1)*

*Contributions: I designed the experiment, programmed the stimulus, collected, analyzed, and interpreted the data, and wrote the manuscript.*

The motion of real-world objects is fairly predictable, at least over short periods of time. For example, the motion of a bicycle is rather consistent. It is therefore easy to rapidly subtract it from the motion of its parts, to compute the nonretinotopic part motion. Likewise, the Ternus-Pikler display is highly predictable: Three black disks move left-and-right in concert. Predictability also plays an important role in visual perception more generally. For example, the brain actively predicts upcoming sensory input instead of awaiting it passively (Muckli, Kohler, Kriegeskorte, & Singer, 2005; Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010). Predicted stimuli are processed more efficiently and are detected more readily (Alink et al., 2010; Vetter, Grosbras, & Muckli, 2015; Posner, 1980; Posner, Snyder, & Davidson, 1980).

In nonretinotopic perception across saccadic eye movements, visual stability is achieved through efference copies, that *predict* the shift of the retinal image before the eye moves (Wurtz, 2008; von Holst & Mittelstädt, 1950; Sperry, 1950). Visual stability breaks down when the retinal shift is rendered unpredictable by inactivating the neural pathway underlying the efference copy signalling (Sommer & Wurtz, 2006). Hence, predictability is crucial for nonretinotopic perception across eye movements. Here, we raised the question to what extent predictability is crucial for nonretinotopic motion perception *without* eye movements. In Experiment 1, we tested three factors: Unpredictable contrast polarity, unpredictable contours, and unpredictable motion trajectories. We were most interested whether unpredictable motion of the disks would impair nonretinotopic perception, as this motion is perceptually subtracted from the dot motion. We included conditions with unpredictable contours and contrast polarities to test the implementation of edge detection in computational models of the Ternus-Pikler display (Clarke et al., 2016; Ağaoğlu et al., 2016). In the models, this edge detection is contrast polarity invariant, but so far there were no

psychophysical data to confirm this implementation. If human observers' nonretinotopic perception were impaired in these conditions, this would call for a revision of the model. The stimuli are schematically depicted in Figure 9.



*Figure 9 Uncertainty Experiment 1: Stimuli and results. Top: Performance in terms of sensitivity (d'). Bottom: Mean change in sensitivity (d') compared to performance in Condition 1. Note, that in the bottom plot for Condition 1, the actual d' values (as opposed to change in d') are plotted for comparison. P values were computed in one-sample t tests against 0 and are reported Bonferroni corrected in red and uncorrected in black. In Condition 1, three black disks moved predictably left and right. In Condition 2, the contrast polarity of the disks reversed from frame to frame but was the same for all three disks. In Condition 3, the contrast polarity was randomly chosen for each individual disk from frame to frame. In Condition 4, the shape changed from frame to frame. Shapes were identical in each frame. In Condition 5, the three disks were as in Condition 1, but the position of the three disks changed randomly from frame to frame. In Condition 6, the disks moved as in Condition 5 and changed contrast polarity and shape as in Conditions 3 and 4. Only four stimulus frames with dot rotation were presented per trial, preceded by two and followed by one frame without dots. Reprinted from Lauffs, Öğmen, & Herzog (2017).*

To establish the individual baseline performance level, we used the same predictable stimulus as in earlier studies (cf. Boi et al., 2009, and Figure 7). Observers were asked to indicate the direction of the nonretinotopic dot-rotation in the middle disk. We then decreased the predictability of the stimulus and tracked the impact on performance. In Condition 2, the contrast polarity of the entire group inverted (reliably) from frame to frame. In Condition 3, the contrast polarity of each individual disk was chosen randomly from frame to frame. In Condition 4, the three elements had a different, randomly chosen shape in each frame. Most importantly, in Condition 5 the group of disks moved randomly from frame to frame, instead of reliably moving left-and-right. Its motion trajectory was, hence, largely unpredictable. In Condition 6, we combined all previous manipulations: The disks changed shape and contrast polarity unpredictably, and moved randomly over the screen. As control conditions, each condition was run a second time with only two disks. The instruction for the observer to indicate the rotation direction was maintained, although the rotation could not be perceived. We thereby controlled that the correct response could not be cognitively inferred, without perceiving the rotation.

Performance was good in the predictable baseline condition. When we decreased the predictability of the stimulus by randomly choosing the stimulus polarity, shape, and motion paths, performance did *not* break down, but was at roughly the same level as in the baseline condition. This indicates that the nonretinotopic dot rotation was well perceived. Only in Condition 6, where we combined all previous manipulations, there was a trend towards slightly lower performance than in the baseline condition, but nonretinotopic perception did not break down completely. Importantly, performance was at chance-level for all control conditions. Performance in the experimental conditions with three disks must hence have indeed been due to the *perception* of rotation.

These results indicate that unpredictability of contrast polarity, shape, and motion paths, did not significantly impair the nonretinotopic dot rotation percept. Although we used an apparent motion stimulus, we believe that the results can be generalized to "normal", ecological viewing conditions. If this is the case, our results suggest that nonretinotopic motion can be computed very rapidly, at least within the 333 ms before the next stimulus frame was presented, because predictability is not necessary. The finding of unimpaired performance when the shape and contrast polarity of the disks changed unpredictably confirms the contrast invariant implementation of edge detection in computational models.

In Experiment 2, we decreased the predictability of the stimulus by randomly jittering the position of the disks *within* the group. Jittering the positions within the group has the potential of impairing the grouping between the disks. Since group motion perception is crucial for the nonretinotopic motion computation, we expected performance to be impaired. The stimuli are schematically depicted in Figure 10.
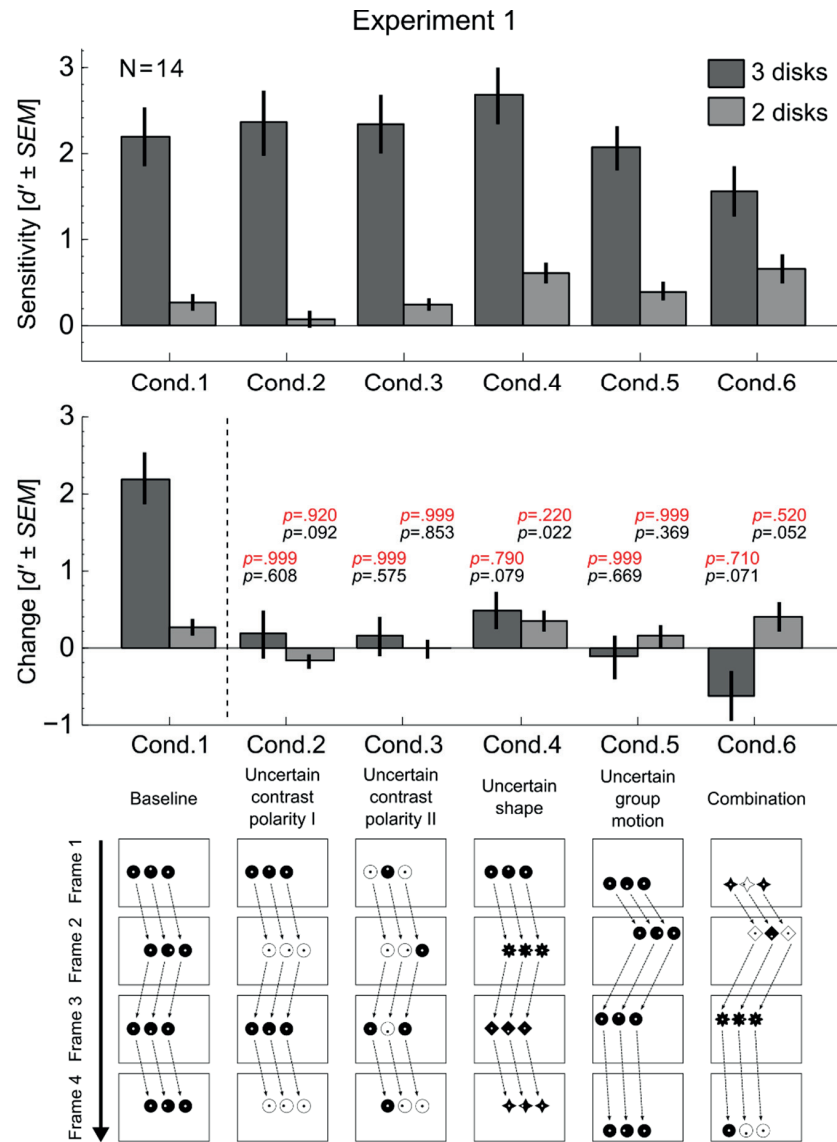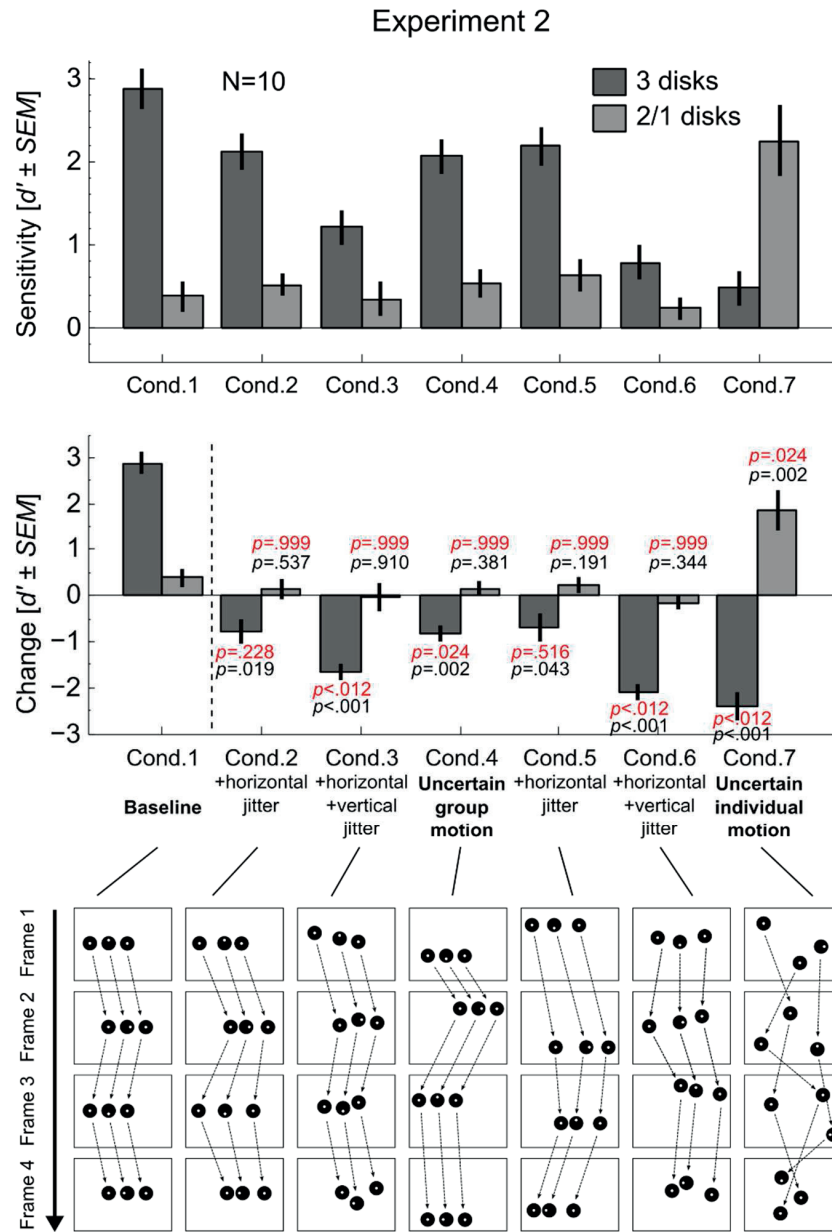
*Figure 10 Uncertainty Experiment 2: Stimuli and results. Top: Performance in terms of sensitivity (d'). Bottom: Mean change in sensitivity (d') compared to performance in Condition 1. Note that in the bottom plot for Condition 1, the actual d' values (as opposed to Change in d') are plotted for comparison. P values were computed in one-sample t tests against 0 and are reported Bonferroni corrected in red and uncorrected in black. In Conditions 1-3, three black disks moved predictably left and right and in Conditions 4-6, the three disks moved randomly across the screen as a group. In Conditions 2 and 5, we randomly jittered the inter-stimulus distance of the disks. In Conditions 3 and 6, we additionally jittered each disk's vertical position. Performance in all conditions was lower than in Condition 1 (all p≤.043). Performance in the control conditions with two disks did not differ significantly from Condition 1. In Condition 7, the disks did not move as a group, but each disk followed an individual, randomly determined motion trajectory. Performance was near chance level. Unlike in the other control conditions, only a single disk was presented in the control condition of Condition 7, to show that observers can track the disk well. Only four stimulus frames with dot rotation were presented per trial, preceded by two and followed by one frame without dots. Reprinted from Lauffs, Öğmen, & Herzog (2017).*

In Condition 1, we again established individual baseline performance levels with the predictable left-and-right moving Ternus-Pikler display. In Conditions 2-3, the group of disks moved left-and-right, but we individually added random jitter to each disk's position. In Condition 2, we added random jitter to only the horizontal positions. That is, the three disks shared the same vertical position, but their horizontal positions relative to the group center contained some randomness (cf. stimuli in Figure 10). In Condition 3, we additionally jittered each disk's vertical position within the group. In Condition 4-6, the disks moved randomly over the screen as a group. In Condition 4, no jitter was added to the disks' positions within the group. In Condition 5, we added only horizontal jitter to each disk's position within the randomly moving group. In Condition 6, we added horizontal *and* vertical jitter. In all conditions discussed so far, the order of the disks was always maintained throughout the trial. That is, the left-most disk in one frame corresponded to the left-most disk in any other frame, and so forth. In Condition 7, we abolished this group structure by letting each disk move randomly, fully independent of the other disks. As control conditions we again presented the same conditions with only two disks.

We found that all manipulations had a significant negative effect on performance compared to baseline (Figure 10). Performance was worse when the stimulus contained horizontal *and* vertical, rather than only horizontal jitter. This pattern was the same when the group of disks moved reliably left-and-right (Condition 2-3) and when it moved randomly over the screen (Condition 4-6). When the disks moved fully independently from each other, performance was close to chance level (Condition 7). Hence, the more unpredictable the individual disk positions were with respect to each other, the more performance was affected. In the control conditions with two disks, performance was again close to chance level in all conditions, showing that the correct answer could not be cognitively derived.

In the Ternus-Pikler display, the perceptual grouping of the disks is crucial, because the group motion is subtracted from the retinotopic dot motion to compute the nonretinotopic percept (Öğmen & Herzog, 2010; Boi et al., 2009; Clarke et al., 2016). When the grouping of the disks fails, the group motion is not clearly perceived. Our interpretation of the results is that by introducing unpredictable motion of the disks relative to each other, we impaired their perception as a group, causing the computation of nonretinotopic motion to fail. This interpretation is supported by a strong correlation between performance and subjective ratings of both grouping strength and predictability. Without the group motion percept, the rotation task can only be solved by tracking the rotating disk attentively. As the control conditions with two disks show, this is not possible with the current stimulus. In sum, our experiments indicate that nonretinotopic motion perception does not require predictability of the shape, contrast polarity, and motion path of the reference object (Experiment 1). However, when unpredictability of the individual elements impairs their grouping, group motion perception is impaired, and nonretinotopic motion perception breaks down (Experiment 2). The strongest grouping factor in our experiments was common motion, as predicted by the model.

## 2.2 Unconscious retinotopic motion processing affects nonretinotopic motion perception

Invisible stimuli can affect behavior and the perception of subsequent stimuli. For example, in masked priming a briefly presented prime stimulus is rendered invisible by the presentation of a trailing mask stimulus. Observers are unable to indicate whether the prime pointed to the left or right. Nonetheless, subsequent target stimuli on the primed side are detected faster than on the unprimed side. This is usually explained by pre-activation of the motor system during the first unconscious, feed-forward sweep of cortical stimulus processing, facilitating responses in the primed and inhibiting responses in other locations (Breitmeyer & Öğmen, 2006). While the effects of masked priming are short-lived, invisible stimuli can also have long lasting effects on perception. For example, in sequential metacontrast masking a centrally presented Vernier stimulus (i.e., two abutting vertical lines with a horizontal offset) is rendered invisible by subsequently presented flanking lines (Otto, Öğmen, & Herzog, 2006). Although the central stimulus itself is invisible, its offset is perceived in the trailing lines up to 450 ms after it was presented (Drissi Daoudi, Öğmen, & Herzog, 2016; also see Scharnowski et al., 2009; Otto, Öğmen, & Herzog, 2009).

In ambiguous figures, two interpretations of the stimulus compete for conscious perception. Unlike in masking, the stimulus itself is fully visible, but only one *interpretation* is consciously perceived at a time. The non-dominant interpretation is suppressed. Interestingly, this suppression transfers to later, similar stimuli. For example, Peterson and colleagues (Peterson & Kim, 2001; Peterson & Skow, 2008) performed experiments in which the background was shaped like an object, for example half a line drawing of a house. A meaningless figure in the foreground was perceived, suppressing the interpretation of the background. Although the object in the background was not consciously perceived, responses to a subsequently presented drawing were slower when this drawing depicted an object from the same category, in this example a building, rather than an animal. Hence, although the non-dominant interpretation were invisible, its unconscious processing interfered with the conscious perception of visible objects later on.

Similar to ambiguous figures, the dot motion in a Ternus-Pikler display can be interpreted in two ways: In retinotopic or nonretinotopic coordinates. When three disks in group motion are presented, the retinotopic dot motion interpretation is invisible, because the nonretinotopic dot motion dominates the percept. It is already known from previous studies that retinotopic motion can produce strong retinotopic motion after-effects, even

when nonretinotopic motion is perceived (Boi et al., 2009; 2011). However, it is uncertain if in these experiments the retinotopic motion was really invisible, because drifting gratings were used that also drifted within each frame.

Here, we tested whether the unconscious processing of retinotopic motion can affect the conscious perception of nonretinotopic motion. To this end, we adapted the dot motion Ternus-Pikler display to contain a *retinotopic* rotation, in addition to the usual nonretinotopic rotation (Figure 11b). The dots in the display change position from frame to frame, but stay in the same position within each frame. Hence, motion can only be perceived as apparent motion across frames, but never within one frame. Whether the disks are perceived as moving or stationary determines whether the retinotopic or nonretinotopic dot motion is perceived. The retinotopic and nonretinotopic rotation directions were combined in a factorial design and could hence be in the same or in opposite directions (Figure 11c). As expected from earlier experiments, only the nonretinotopic rotation was perceived when three disks were presented. On the other hand, only the retinotopic rotation was perceived when only two disks were presented. As in previous experiments, the retinal image is identical in both cases, except for a third contextual disk - the addition of the third disk changes the interpretation of the stimulus, which leads to the retinotopic dot motion being invisible.

To obtain a baseline performance measure, we first presented a Ternus-Pikler with nonretinotopic rotation in the middle disk and omitted the dots from the outer disks. Observers reported the direction of the nonretinotopic rotation. Performance was very good (Figure 12; C1: 84.0%). Next, we presented the adapted Ternus-Pikler display with both retinotopic *and* nonretinotopic rotations (Figure 11b, c). Overall, performance was lower than in the baseline condition ($C2_{all}$: 73.3%; $C2_{all}$ vs. C1: $p$ = .002). We then split up the data for this condition into trials in which the retinotopic and nonretinotopic rotation directions were in the same, and when they were in opposite directions. If the retinotopic rotation had no effect on the nonretinotopic percept, we would expect similar performance levels in both subsets. However, this was not at all the case: Performance was severely impaired and hardly above chance-level when the rotation directions were incongruent ($C2_{opposite}$: 59.3%; $C2_{opposite}$ vs. 50%: $p$ = .090). On the contrary, when the retinotopic and nonretinotopic rotations were in the same direction, the nonretinotopic rotation could be reported equally well as in the baseline condition ($C2_{same}$: 87.3%; $C2_{same}$ vs. C1: $p$ = .284).

This strong influence of retinotopic rotation on the conscious percept is surprising, because the retinotopic rotation itself is almost invisible: In Condition 3, we presented the same stimulus as in Condition 2, but asked the observers to report the *retinotopic* rotation direction. Performance was very low overall, and hence cannot explain the large effect on nonretinotopic perception we observed in Condition 2 ($C3_{all}$: 57.5%; $C3_{all}$ vs. 50%: $p$ = .004). When only two disks were presented, the retinotopic rotation could easily be reported irrespective of the nonretinotopic rotation direction ($C4_{all}$: 94.4%, $C4_{same}$: 95.8%, $C4_{opposite}$: 93.0%).
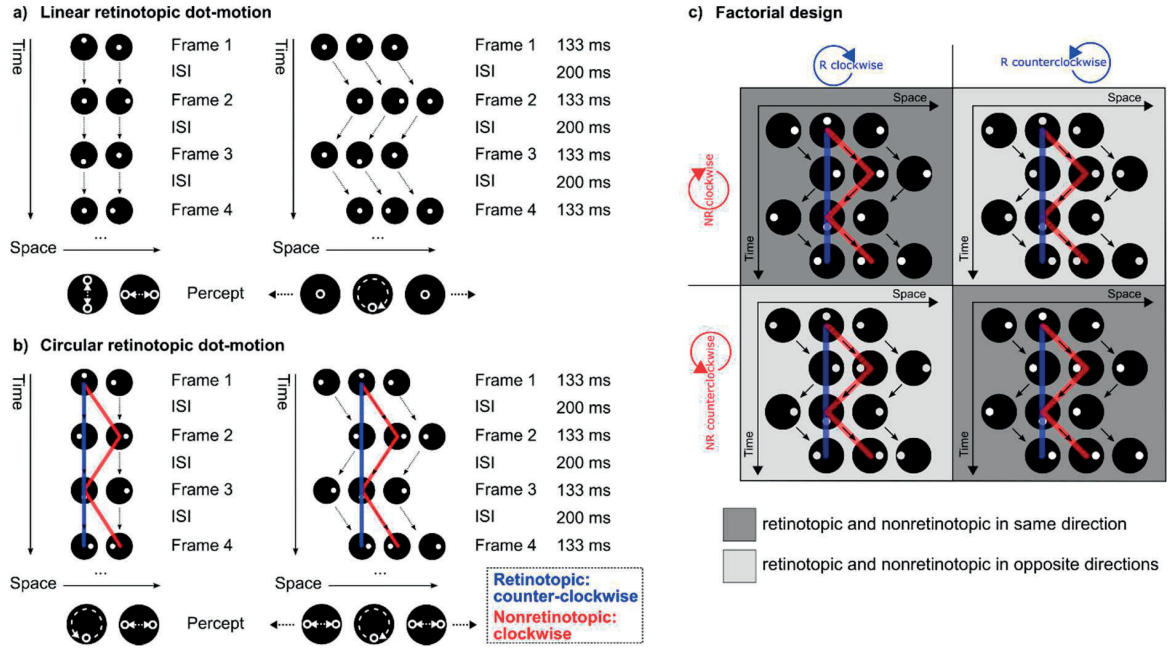
*Figure 11 Unconscious processing experiments: Stimulus and design. We adapted the stimulus of Boi et al. 2009 (Figure 11a), to contain a retinotopic rotation, in addition to the usual nonretinotopic rotation (Figure 11b). The rotation direction of the retinotopic and nonretinotopic rotation were varied in a balanced factorial design (Figure 11c), and could hence be in the same (dark grey background) or opposite direction (light grey background). When three disks were presented the nonretinotopic dot rotation was perceived in the middle disk, and the dots in the left and right disk jumped from left to right or right to left on every other frame. When only two disks were presented, a retinotopic rotation was perceived in the left disk and the nonretinotopic rotation was invisible. Only four stimulus frames with dot rotation were presented per trial, preceded by two and followed by one frame without dots. R - retinotopic rotation direction. NR - nonretinotopic rotation direction. The blue and red lines indicate the rotation and were not presented during the experiment. All stimuli were presented on a midlevel grey background. Adapted from Lauffs, Choung, Öğmen, & Herzog (submitted).*
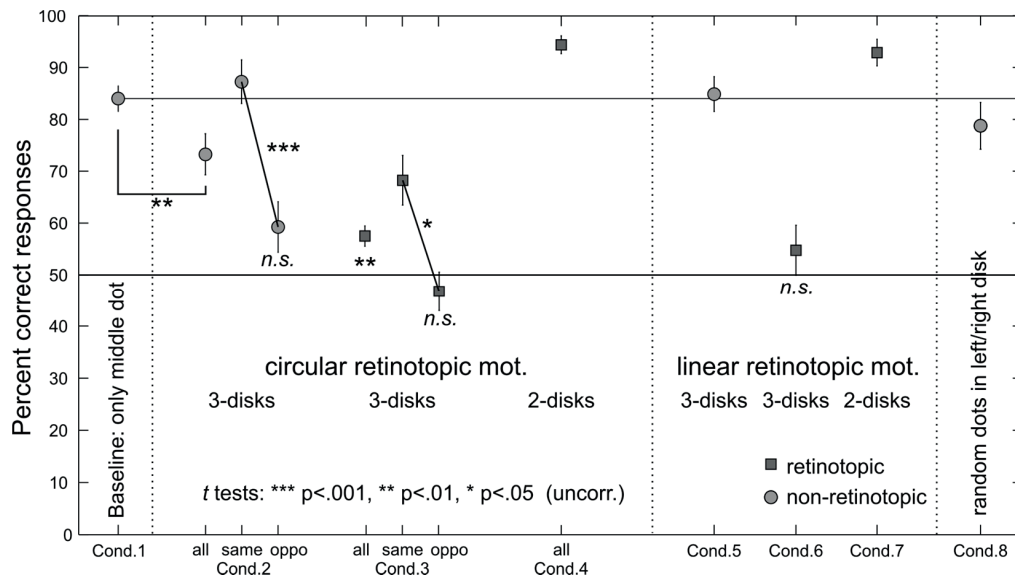
*Figure 12 Unconscious processing Experiment 1: Results. Light grey circles indicate that observers were instructed to report the nonretinotopic rotation in the middle disk. Dark grey squares indicate that observers were asked to report the direction of the retinotopic motion. As a baseline for comparison, we omitted the dots in the outer disks and observers reported the direction of the nonretinotopic dot rotation in the middle disk (C1). Performance was equally good when the dot positions in the outer disks were chosen randomly (C8), or in the "classic" configuration with linear up-down and left-right retinotopic dot motion (C5). However, perception of the nonretinotopic rotation was severely impaired when the retinotopic dot motion was a rotation in opposite direction (C2opposite). Performance was comparable to baseline when the retinotopic and nonretinotopic rotations were in the same direction (C2same). This strong effect of the retinotopic dot motion is surprising, because it is largely invisible (C3, C6). It can only reliably be perceived when the third disk is omitted (C4, C7). Error bars depict the standard error of the mean. Reprinted from Lauffs, Choung, Öğmen, & Herzog (submitted).*

23

For comparison, we next presented the classic Ternus-Pikler display with *linear* retinotopic dot motions (Figure 11a). The nonretinotopic rotation direction could be discriminated equally well as in the baseline condition (C5: 84.9%; C5 vs. C1: $p = .678$). The retinotopic motion direction, in this case linear up-down or left-right motion in a predefined retinotopic position, could not be reported above chance-level (C6: 54.7%; C6 vs. 50%: $p = .352$). However, it was clearly visible when only two disks were presented (C7: 92.9%). Finally, we presented a Ternus-Pikler display in which the dots in the left and right disk were placed randomly. Although this creates random dot motion percepts in the flanking disks, observers could discriminate the nonretinotopic dot rotation direction in the middle disk equally well as in the baseline condition (C8: 78.7%; C8 vs. C1: $p = .130$).

In a very similar second experiment (described in detail in Appendix 2) we replicated these effects in a new group of naïve observers: Incongruent retinotopic rotation interfered with the nonretinotopic rotation percept, even though it was almost invisible. Linear or random retinotopic dot motion, and congruent rotation did not impair the nonretinotopic rotation percept.

To examine whether low-level differences between the conditions can explain the large effect of incongruent retinotopic rotation on the nonretinotopic rotation percept, we split up the data depending on the start position of the dot (3, 6, 9, 12 o'clock), and start position of the disks (left/right). Noise levels were substantial due to the low number of trials, but we found no striking differences. This suggests that the impairment observed in the incongruent rotation case is not simply due to the low-level properties of the stimulus. Instead, it appears that the retinotopic dot motion is processed unconsciously and that it interferes with the conscious interpretation only when it is incompatible.

In a third experiment, we investigated whether *two* retinotopic rotations influence the nonretinotopic percept independently, or whether they interact before jointly influencing the percept. We added a new condition in which we presented not one, but two retinotopic rotations, in addition to the familiar nonretinotopic rotation (Figure 13a). If the retinotopic rotations interact at an early processing stage, we expected that the presence of a *congruent* retinotopic rotation would compensate for the negative effect of an *incongruent* retinotopic rotation. If, on the other hand, the retinotopic rotations influence the nonretinotopic percept independently, we expected a further decrease in performance when two, instead of one, incongruent retinotopic rotations are presented. Our data support the latter scenario: Performance decreased almost linearly when first one and then two incongruent retinotopic rotations were presented (Figure 14). This argues against a mechanism in which the retinotopic rotations are first pooled before they influence the nonretinotopic percept. Rather they seem to have a rather independent, additive, effect.

In another condition we investigated whether nonretinotopic rotations affect performance in a similar fashion as the retinotopic rotations did. We presented no retinotopic rotations. Instead, we presented nonretinotopic rotations not only in the middle disk, but also

in the left and right disk (Figure 13b). We found that whether these rotations were congruent or incongruent with the rotation in the middle disk only had a minor and insignificant influence on performance (Figure 14). Hence, only incongruent *retinotopic* rotations impaired the nonretinotopic percept, but not further nonretinotopic rotations.
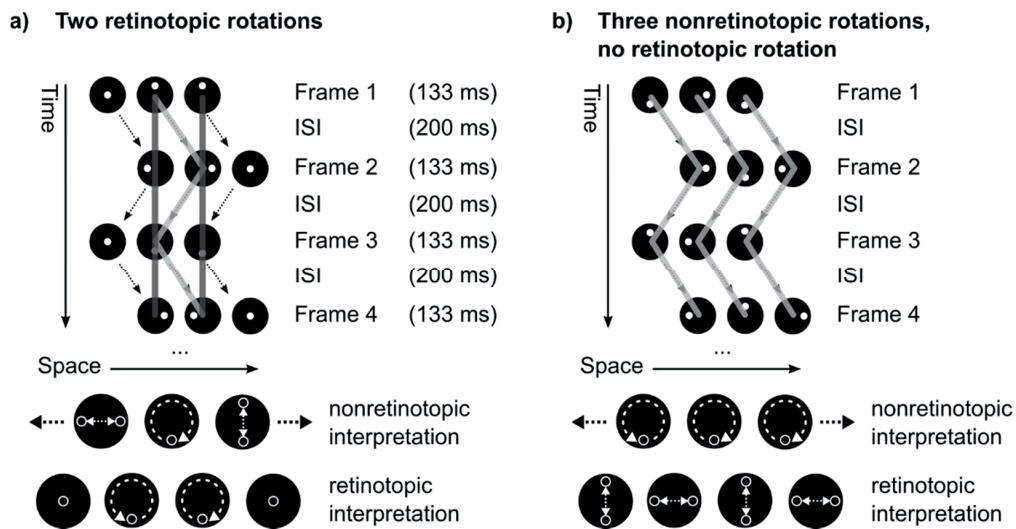


*Figure 13 Unconscious processing experiments: New stimulus in Experiment 3. A) Two disks contained retinotopic rotations, instead of one disk as in Experiments 1-2. The two retinotopic rotations could be in the same or opposite direction. In addition, each retinotopic rotation could be in the same or opposite direction of the nonretinotopic rotation in the middle disk. In the depicted example, the nonretinotopic rotation (light grey) and right retinotopic rotation (dark grey) are clockwise, and the left retinotopic rotation is counter-clockwise. Only the nonretinotopic interpretation is perceived: The dot in the middle disk rotates and the dots in the left and right disks move up-and-down and left-and-right, respectively. The retinotopic interpretation of the image was not perceived, but is depicted here to illustrate the two retinotopic rotations. B) No retinotopic, but three nonretinotopic rotations were presented. The observers reported the rotation in the middle disk. The outer disks could rotate both in the same direction as the middle disk, both in the opposite direction, or one in the same and one in the opposite direction. Retinotopically, the dots moved left-right or up-down every second frame. Rotation was perceived in all three disks (nonretinotopic interpretation). The retinotopic interpretation was not perceived and shown only for illustration. Reprinted from Lauffs, Choung, Öğmen, & Herzog (submitted).*
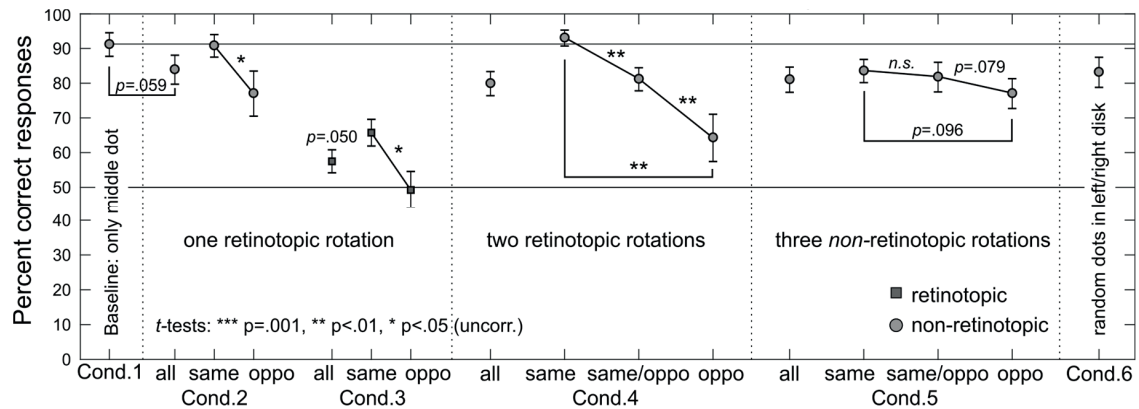
*Figure 14 Unconscious processing Experiment 3: Results. Conditions 1-3 replicate the findings of Experiment 1: Unconscious processing of an incongruent retinotopic rotation interferes with nonretinotopic rotation perception. In Condition 4, we used a stimulus with two retinotopic rotations. Performance decreased when one retinotopic rotation direction was incongruent with the nonretinotopic rotation, compared to when both retinotopic rotations were in the same direction as the nonretinotopic rotation. Performance decreased even more when both retinotopic rotations were incongruent with the nonretinotopic rotation. In Condition 5, we presented nonretinotopic rotations in all three disks and there was no retinotopic rotation. The participants perceived dot rotations in all three disks and reported the rotation in the middle disk. Performance was slightly lower than in the baseline condition (Condition 1), and differed only little depending on whether the left and right disks rotated in the same or opposite sense of the middle disk. In Condition 6, the dots in the left and right disk were placed randomly and the middle disk rotated. Performance was slightly lower than in the baseline condition. Error bars depict the standard error of the mean. Reprinted from Lauffs, Choung, Öğmen, & Herzog (submitted).*

## 2.3 Retinotopic perception despite group motion at intermediate ISI durations

*This subchapter summarizes unpublished data of an ongoing project.*

*Collaborators: Marc M. Lauffs, Haluk Öğmen, & Michael H. Herzog.*
*Contributions: I conceived of and designed the experiment, programmed the stimulus, and collected, analyzed, and interpreted the data.*

It has been proposed that nonretinotopic motion computation in general, and in the Ternus-Pikler display in particular, is a two-stage process (Johansson, 1973; Öğmen & Herzog, 2010; Clarke et al., 2016; Ağaoğlu et al., 2016). In the first stage, the correspondence problem is solved, yielding the motion vectors of the display elements in retinotopic coordinates. In the second stage, the retinotopic motion vectors are decomposed into motion that is shared between elements, and the residual motion vector that is unique to each element. The shared motion is perceptually subtracted from each element's motion, to yield the (nonretinotopic) residual motion.

This model can well explain perception in the Ternus-Pikler display (Clarke et al., 2016). When group motion is perceived, the left-and-right motion of the disks is perceptually subtracted from the retinotopic motion of the dots, yielding the nonretinotopic dot rotation percept as residual. When element motion is perceived, the two central disks appear stationary, so there is no motion to be subtracted from the dots: The retinotopic up-down and left-right dot motions are perceived. An implicit implication of this model is that whenever the disks appear to move, this motion is subtracted from the retinotopic dot motion, resulting in a circular dot motion percept.

In a series of experiments, I found that this is not necessarily the case. I presented the Ternus-Pikler display with white dots (Figure 7) and randomly varied the ISI duration from trial to trial. Observers indicated A) whether they perceived element motion or group motion, and B) whether they perceived linear or circular dot motion. The tasks were performed either in separate blocks (single-task), or both in the same trial (dual-task). In either case the order of the tasks was counterbalanced across participants. The response mapping, for example which button to press for group motion, was randomly counterbalanced within participants on a trial-by-trial basis. In dual-task blocks, eight instead of four stimulus frames were presented, so that there was enough time to pay attention to both movements. Order effects or differences due to limited attentional resources are hence controlled for as far as possible.

Under the assumption that the model is correct, rotation should be perceived in the exact same number of trials in which group motion is perceived (Figure 15a). This was indeed the case for very short (0 ms) and long ISIs (above ca. 100 ms). However, for ISI durations of intermediate length, group motion was consistently more frequently reported than circular dot motion (Figure 15b-d). This suggests that in some trials group motion was perceived in combination with linear dot motion. Theoretically, it is possible that differences in task

difficulty cause such a pattern of results. However, my personal subjective experience is that it is indeed possible to perceive linear dot motion, although the disks appear to move back and forth as a group. The exact ISI durations at which either percept switched are of no particular interest, because perception of the Ternus-Pikler display is determined by many factors, not just the ISI duration (Petersik & Rice, 2006, 2008).

According to our current model of nonretinotopic processing (Öğmen & Herzog, 2010; Clarke et al., 2016), this case should not occur, because when group motion is perceived it is subtracted from the linear (retinotopic) dot motion to yield a circular (nonretinotopic) motion percept. Does this mean that the model is wrong? Not necessarily. Both processing stages of the model require computations and hence take time to be completed. A long ISI between two stimulus frames leaves the system enough time to complete both stages: First, object correspondence is computed for the disks; second, their global motion is subtracted from the dot motions. With an ISI of 0 ms, there is no time to complete either stage, and only the raw retinotopic input is perceived (i.e., element motion and linear dot motion). At intermediate ISI durations, there is usually enough time to complete the first processing stage, leading to the perception of group motion. However, the disks reappear on the screen before the second processing stage is completed, in which the group motion is subtracted from the dot motion. The retinotopic, linear dot motion is perceived, because the nonretinotopic percept could not be computed in time. I suggest that the different report frequencies for disk and dot motions at intermediate ISIs are evidence for, rather than against our current model of nonretinotopic motion computation. However, this position can be debated, and further experiments are needed to verify that the difference is indeed caused by differences in perception, rather than task difficulty.
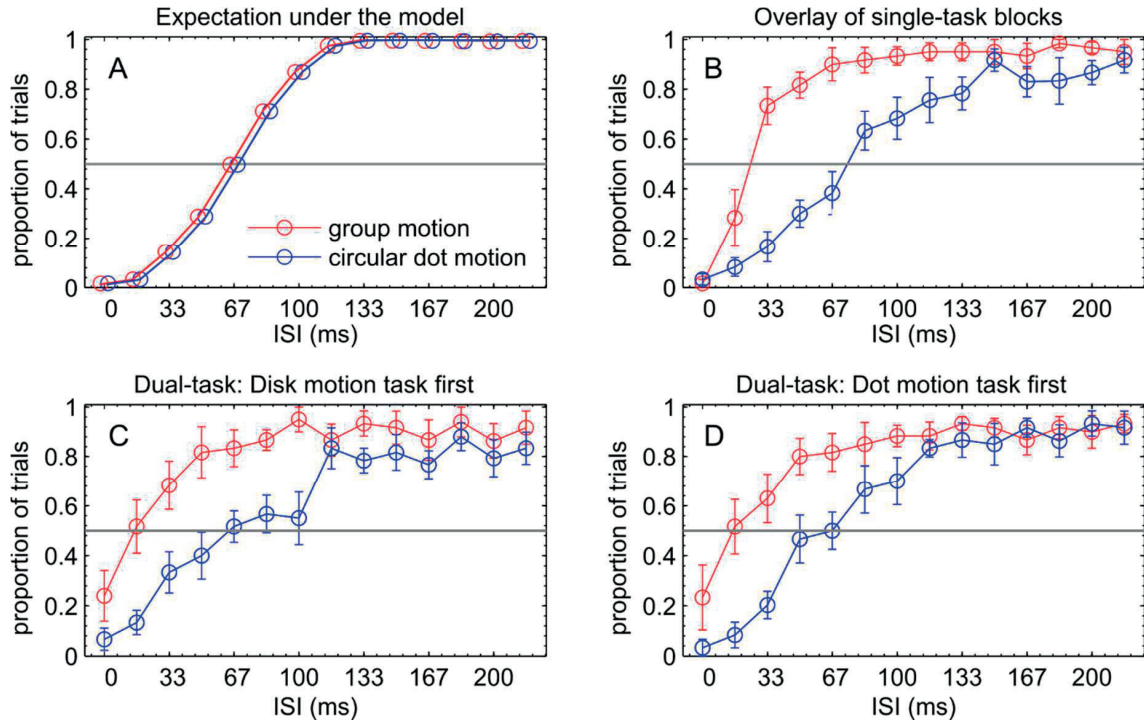
*Figure 15 Ternus-Pikler display with intermediate ISI durations: Results. In theory, whenever group motion is perceived, the nonretinotopic circular dot motion should be perceived. However, this does not seem to be the case: For ISI durations of intermediate length (here: roughly 17-100ms) group motion is more frequently perceived than dot rotation. A) Hypothetical psychometric functions under the assumption that the model is correct. B) Observers reported in separate blocks whether they perceived element or group motion, and whether they perceived linear (retinotopic) or circular (nonretinotopic) dot motion. C) Dual-task condition: After each trial, observers reported first whether they perceived element or group motion and second whether they perceived linear or circular dot motion. D) Same as c), but observers reported the dot motion percept first. The dual-task data (Figure 15c-d) are somewhat more noisy than the single-task data (Figure 15b), but show the same pattern. The increased noise may be expected due to higher attentional demands and an added memory component in the dual-task. Error-bars depict the standard error of the mean.*

## 2.4 Nonretinotopic motion processing in schizophrenia patients

*Contributions: I designed the experiment, programmed the stimulus, analyzed and interpreted the data, and wrote the manuscript.*

Schizophrenia is a devastating disease entailing severe cognitive and perceptual impairments. Past research has primarily focused on cognitive deficits, although perceptual deficits are among the most sensitive endophenotypes of schizophrenia (Chkonia et al., 2010). For example, the perception of *simple*, local motion in drifting gratings seems intact (Chen et al., 2003; Tibber et al., 2015). But patients are impaired in tasks involving *complex* motion, for example, when the coherent motion of a subset of dots has to be detected in a random dot kinematogram (Chen et al. 2003). Results on the perception of *apparent* motion are mixed (Saucer and Deabler, 1956; Saucer, 1958, 1959; Chambers and Wilson, 1986; Sanders et al., 2013; but see Sanders et al., 2012; Saucer, 1958; Tschacher et al., 2008). A deficit might be expected, because apparent motion perception requires complex processing, for example when inferring the motion path between the inducers.

We investigated nonretinotopic motion processing in schizophrenia using the Ternus-Pikler display. Due to the specific impairment of complex, but not simple motion perception, we hoped that schizophrenia patients would show only a partial deficit, that is specific enough to allow conclusions regarding the mechanisms underlying nonretinotopic perception. For example, a possible finding would be that the perception of group motion is relatively intact, but that perception of the nonretinotopic dot rotation is impaired due to its higher complexity. First, we presented a simple Ternus-Pikler display with three black disks without the white dots (Figure 6). We randomly varied the ISI duration and observers indicated whether they perceived element motion or group motion. Schizophrenia patients' responses followed the same pattern as those of healthy, age-matched controls (Figure 16a): Element motion was perceived at an ISI of 0 ms and group motion was perceived increasingly frequently as the ISI duration increased. We did not find significant differences between patients and controls at any of the used ISI durations. The perceptual organization of schizophrenia patients hence matched that of healthy controls. Next, we added white dots to the disks and observers indicated the nonretinotopic rotation direction of the dot in the middle disk (Figure 7). Again we did not find significant differences between patients and controls (Figure 16b). For an ISI of 0 ms, element motion and retinotopic up-down, left-right dot motions are perceived. The nonretinotopic rotation was not perceived and performance was close to chance-level for both groups. For an ISI of 200 ms, group motion and the nonretinotopic dot rotation in the middle disk are perceived. Performance was at the same level for patients and controls. As a control condition, instead of moving left-and-right, three

disks were flashed stationarily and the dot in the middle disk rotated. Performance in the control conditions was good in both groups, showing that the task was well understood and feasible.

In sum, the results suggest that nonretinotopic motion processing is not altered in schizophrenia patients. However, since the Ternus-Pikler display is a highly complex motion stimulus, the absence of an effect sheds doubt on the assumption that the *complexity* of random dot kinematograms alone causes schizophrenia patients' impaired performance. We suggest a moderating role of attention: Nonretinotopic motion computation, although complex, is a highly natural kind of processing, because the default perception of moving objects is nonretinotopic. It is therefore likely to be highly automatized and require little attention. The same is not true of random dot kinematograms, which are very artificial stimuli.
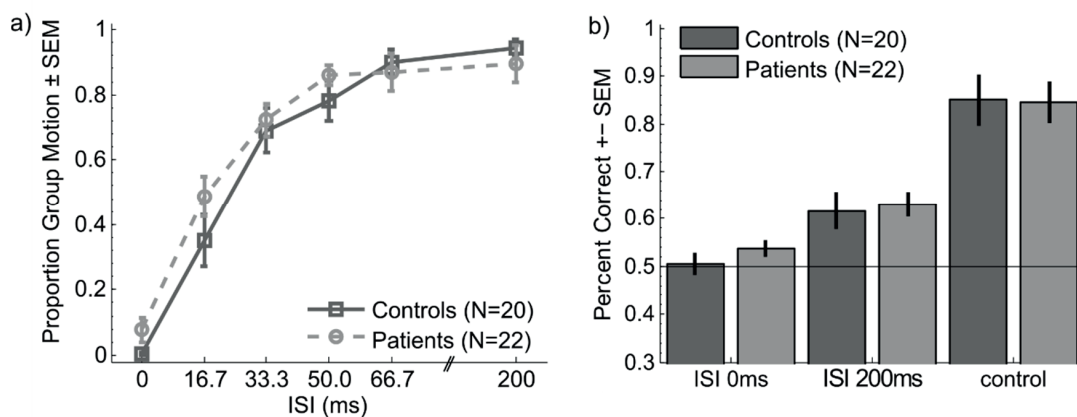


*Figure 16 Schizophrenia experiments: Results. A) A Ternus-Pikler display with three black disks (without white dots) was presented and the ISI was varied. For very brief ISIs, element motion was perceived. As the ISI duration increased, group motion perception increased. This pattern was identical for schizophrenia patients and healthy, age-matched control subjects. We did not find significant differences for any of the ISIs. B) A Ternus-Pikler display with white dots was presented and observers discriminated the nonretinotopic dot rotation direction. For an ISI of 0 ms, element motion and linear dot motion were perceived and rotation discrimination performance was near chance-level. For an ISI of 200 ms, group motion and the nonretinotopic dot rotation are perceived. In a control condition with a dot rotation in three stationary disks performance was good, confirming that the task was well understood. Again, we did not observe significant differences between patients and controls, suggesting that nonretinotopic motion perception in schizophrenia patients is intact. Error bars depict the standard error of the mean. Adapted from Lauffs et al. (2016) with permission © Elsevier*

## 2.5 Sustained spatial attention affects feature fusion through enhanced signal encoding

*This subchapter summarizes the manuscript*
*Hochmitz, I., Lauffs, M. M., Herzog, M. H., & Yeshurun, Y. (in preparation).*
*(Appendix 4)*

*Contributions: I programmed the stimulus, collected, analyzed, and interpreted the data, and wrote the manuscript.*

When two elements are presented in rapid succession, *feature integration* can occur: Not two elements with distinct features are perceived, but only one element that combines, or integrates, the features of both. For example, when a red and a green disk are briefly presented in rapid succession, a single yellow disk is perceived (Efron, 1967, 1973). The color of the second disk integrates with the color of the first disk, resulting in a fused color percept. In the fused percept, elements presented more recently are perceived slightly stronger than more temporally distant elements (Efron, 1973; Yund, Morgan, & Efron, 1983). For example, when first a red and second a green disk is presented, a single yellow disk with a slight green tint is perceived.

We investigated the influence of sustained spatial attention on feature integration. We used a Vernier stimulus, two abutting vertical lines separated by a small gap (Figure 17). The lower line is horizontally offset with respect to the upper line and the observer indicates whether this offset is to the left or to the right. When two Vernier stimuli are presented in rapid succession, only a single Vernier with fused offset is perceived (Hermens, Scharnowski & Herzog, 2009). If the offset of the second Vernier is in the opposite direction of the first Vernier (i.e., if it is an "Anti-Vernier"), the opposing offsets cancel each other and an almost aligned Vernier is perceived. Since more recent elements dominate the percept, a slight offset in line with the second Vernier is perceived, that is much smaller than the veridical offset. The percept is nonretinotopic because it pools information over retinal locations and strongly deviates from what was registered by the retina (two individual Verniers with opposite offsets).

We used spatial uncertainty to manipulate the amount of sustained/endogenous attention available for stimulus processing: Either the stimuli were always presented 2° to the right of a fixation point, or 2° to the left or right in random order. In the former case, the observer knows with certainty where the stimuli will appear and can focus all attention on stimulus processing in this location (spatial certainty / full attention condition). In the latter case, the observer does not know whether the stimuli will appear left or right, and has to split attention over both locations (spatial uncertainty / divided attention condition). The observer was informed before each block of trials where the stimuli could appear.

The stimuli were presented too briefly (2 * 30 ms = 60 ms) for eye movements to occur in response to the stimulus (even express saccades have reaction times in the range of

90 - 120 ms; Fischer & Ramsperger, 1984; Gezeck, Fischer, & Timmer, 1997; Rohrer & Sparks, 1993). However, when the position is known with certainty, the observer might be tempted to fixate on this position instead the fixation point. This would bring the stimulus closer to the fovea and allow the observer to get a sharper image. To prevent this, the stimulus was only presented after a central fixation was detected using eye tracking. In offline analyses, we re-confirmed that the average gaze position indeed did not differ between conditions (post-hoc power was 95% to detect 0.2° differences).
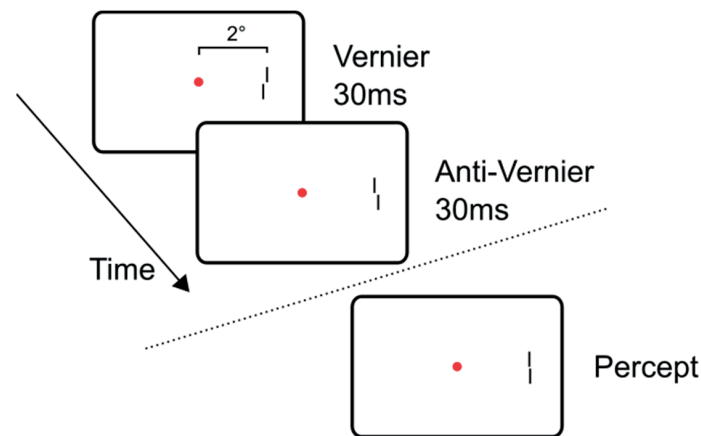


*Figure 17 Feature integration experiments: Stimulus. A Vernier stimulus is presented 2° in the periphery and immediately followed by an Anti-Vernier, i.e., a Vernier with offset in the opposite direction. A single Vernier is perceived, that appears almost aligned because the Vernier and Anti-Vernier offsets are perceptually integrated. The Anti-Vernier offset is perceived slightly more strongly in the fused Vernier. Endogenous spatial attention is manipulated by presenting the stimulus always on the right hand side, or left and right in random order, depending on the block. © CC-BY 4.0*

In Experiment 1, we first presented only a single Vernier, not followed by an anti-Vernier. Observers indicated the perceived offset. We used a staircase procedure to individually adjust the offset size for 75% correct responses when the stimulus location was known. We then presented Verniers with this offset under spatial certainty and uncertainty. Offset-discrimination performance was better when the observer knew where to expect the stimulus, than under spatial uncertainty ($p = .076$), replicating earlier results (Yeshurun & Carrasco, 1999, and Figure 18a, left). Next, we presented Vernier and Anti-Vernier pairs. We adjusted the offset size to yield 25% responses in line with the first Vernier (i.e., 75% anti-Vernier dominance) when the location of the stimulus is known. When the stimulus location was uncertain, the Anti-Vernier dominance of the integrated percept was *less* strong (Figure

18a, right; *p* < .001). This might suggest that integration operates differently under full and divided attention. Specifically, it is possible that the Vernier stimulus is given a higher relative weight in the integration when attention is divided. However, an alternative explanation is that performance is worse under divided attention, because of a lower resolution of the stimuli at the encoding stage (cf. Yeshurun & Carrasco, 1999).
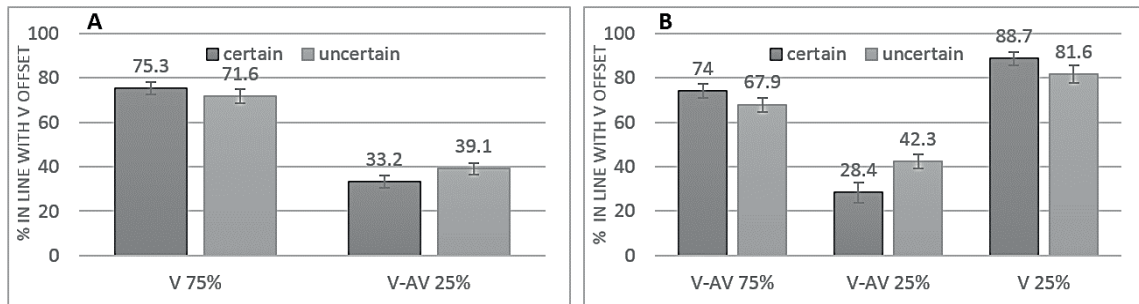


*Figure 18 Feature integration experiments: Results. Plotted is the proportion of responses in line with the (first) Vernier offset [%]. Values below 50% indicate anti-Vernier dominance, values above 50% indicate Vernier dominance. A) Experiment 1. V 75% - Vernier only, offset adjusted for 75% responses Vernier dominance when the location is known. V-AV 25% - Vernier anti-Vernier pair, offsets adjusted for 25% Vernier dominance, i.e., 75% anti-Vernier dominance. B) Experiment 2. V-AV 75% - Vernier anti-Vernier pair, offsets adjusted for 75% Vernier dominance. V-AV 25% - Vernier anti-Vernier pair, offsets adjusted for 25% Vernier dominance. V 75% Vernier only, with same offset size as in V-AV 25% condition. Error bars depict the standard error of the mean.*

In Experiment 2, we sought to differentiate whether the effect occurs at the encoding or integration stage. We tested the assumption that the first Vernier has a stronger weight on the integration under spatial uncertainty, than when the location is kept constant. We added a new condition in which the Vernier is perceived stronger than the Anti-Vernier. Since the Anti-Vernier is given higher weight in the integration, this necessitated reducing its bottom-up energy. By cutting the Anti-Vernier presentation duration in half, we were able to adjust the offsets for 75% Vernier dominance when the location was certain. If indeed attention affects the relative weighting of Vernier and Anti-Vernier in the integration, the Vernier should dominate even more under spatial uncertainty. This was not what we found. Performance was closer to chance under spatial uncertainty than certainty, irrespective of whether the offset was originally adjusted for Vernier or Anti-Vernier dominance (both *p* < .001; Figure 18b).

In sum, it appears that the effect is mostly driven by attention-dependent differences in spatial resolution at the encoding stage, rather than a differential weighting of temporally distal and proximate information at the integration stage.

# Chapter 3    Discussion

In the human brain, visual information is initially processed in retinotopic coordinates, that is, preserving the spatial layout of the retinal image. However, perception is nonretinotopic. For example, we perceive the world as stable and continuous, although the retinal input is constantly changing due to movements of the eye and body. Hence, the brain must compute a stable nonretinotopic percept from the initially retinotopic image. Visual stability across eye movements is typically explained by efference copies, copies of the motor command that predict the displacement of the retinal image (e.g., Wurtz, 2008). This prediction is discounted from the actual image shift, resulting in stable perception.

Nonretinotopic perception is equally important in the absence of eye movements. For example, objects are not perceived in retinal coordinates, but in relation to other objects, even if the eyes are kept perfectly still. In this situation, in which efference copies cannot play a role, it is largely unknown how the brain transforms its retinotopic input into nonretinotopic representations. Our current understanding is that this transformation is a two-stage process (Öğmen & Herzog, 2010): In the first stage, objects are detected and their retinotopic motion is inferred. In the second stage, the objects are grouped based on the similarity of their retinotopic motion paths. This shared motion component is perceived as the motion of the group. The group motion is then perceptually subtracted from the motion of its elements, who are perceived to move, nonretinotopically, within the moving group. This processing can be experienced in the Ternus-Pikler display (Boi et al., 2009). For example, three disks are presented on a computer screen and each disk contains a white dot that is displaced from frame to frame. When the disks appear stationary, a retinotopic up-down or left-right dot motion is perceived. But when the disks appear to move, their motion is subtracted from the retinotopic dot motions and a nonretinotopic dot rotation is perceived. This theoretical model has successfully been developed into computational models, that produce results in line with current psychophysical data (Clarke et al., 2016; Ağaoğlu et al., 2016). However, how the *brain* transforms the retinotopically encoded image into the nonretinotopic percept remains poorly understood.

An intuitive starting point for the investigation of the mechanisms underlying nonretinotopic perception *without* eye movements is the comparison with visual stability *across* eye movements, which is much better understood. Both appear similar in several respects. In visual stability, the retinal motion is discounted for an efference copy predicting the motion. Similarly, in nonretinotopic perception *without* eye movements, the retinotopic object motion is discounted for the motion of a reference object. For example, a reflector on

the wheel of a bicycle is discounted for the motion of the bicycle and perceived to rotate, although its motion is cycloidal in retinotopic coordinates. A crucial feature of efference copies is that they *predict* the retinal shift. This allows the processing of the retinal shift to commence already before it occurs. The use of predictions is not restricted to efference copies: The brain actively predicts future visual stimulation, making its processing faster and more efficient (Muckli et al., 2005; Alink et al., 2010; Vetter et al., 2015). The motion of real-world objects is normally relatively predictable, too. For example, the motion of a bicycle is usually fairly constant and predictable. There are hence obvious similarities between nonretinotopic processing across eye movements and without eye movements, and it is possible that both are supported by common mechanisms in the brain.

I investigated this possibility by asking whether nonretinotopic perception without eye movements requires predictability of the reference object, in the same way that visual stability requires efference copies. My results suggest that nonretinotopic perception is robust against unpredictability, because the nonretinotopic dot rotation percept was unimpaired in the face of unpredictable changes of the reference object's motion trajectory, as well as its contrast polarity and shape. Without a prediction of the reference object, the processing cannot benefit from the more efficient processing of predicted information. This suggests that nonretinotopic motion can be computed rapidly after the stimulus is presented, in our experiments at least within the 333 ms (120 ms stimulus duration + 213 ms ISI) before the next stimulus frame appeared. Nonretinotopic perception is a fundamental part of visual perception, so it is not surprising that its computation is very efficient. However, that predictability does not seem crucial contrasts with the finding that in monkeys visual stability breaks down when the predictive signaling of efference copies is interrupted by deactivating the underlying neural pathway (cf. Wurtz, 2008). This suggests that nonretinotopic perception with and without eye movements is achieved through largely separate mechanisms. To investigate this issue further, experiments are needed in which nonretinotopic perception with and without eye movements is compared directly. For example, experiments with the Ternus-Pikler display could include additional conditions, in which the disks are always presented in the same position and, instead of the disks, the fixation cross is moved left-and-right with every frame. The dot positions are changed while the observer's saccade is in flight, so that the retinal image displaces in the same way as in the conditions in which the gaze is kept stable and instead the stimulus moves. This would allow us to compare nonretinotopic perception with and without eye movements in a quantitative way, and answer, for example, if one is more efficient than the other.

Given that the nonretinotopic percept can apparently be computed very rapidly, the question arises as to how fast this computation can be. I explored the temporal limitations of nonretinotopic processing, by presenting the Ternus-Pikler display with randomly chosen ISI durations between 0 and 217 ms. The observers indicated whether they perceived A) group or element motion of the disks, and B) the nonretinotopic dot rotation or retinotopic linear dot motion. At the longest ISIs, group motion of the disks and the nonretinotopic dot rotation

were perceived in almost all trials. At the shortest ISI, element motion and the retinotopic dot motions were perceived in almost all trials. However, at ISI durations of intermediate length (ca. 17-100 ms) it was possible to perceive retinotopic dot motion, although the nonretinotopic group motion of the disks was perceived. This finding is at odds with our current model of nonretinotopic processing, in which the disk motion is subtracted from the retinotopic dot motions. Hence, when the disks are perceived to move left-and-right in group motion, this motion should be subtracted from the retinotopic dot motion, and the nonretinotopic dot rotation should be perceived. As I argued previously, my observations can be aligned with the model by acknowledging that each stage of the model involves neural computations that take time to complete: At intermediate ISIs, the system has enough time to complete the first stage of the model (computation of group motion, solving the correspondence problem), but not the second (subtraction of group motion from retinotopic dot motions). This means that in the discussion of temporal limitations, the object has to be considered separately from its features. This finding touches on the question how and when features are connected to an object, the so called binding problem (Treisman, 1996). I found differences between disk and dot motion perception up to ISI durations of at least 100 ms. Including the stimulus duration of 130 ms, this suggests that up to 230 ms post stimulus onset, a *stable* binding of the dot feature to the disk was not yet achieved.

A related question is how the objects in a visual scene are grouped more generally. Within a single frame, objects are likely grouped based on their proximity and similarity. But due to the dynamic nature of the Ternus-Pikler display, changes over time have to be taken into account. How the brain determines which disk in frame N corresponds to which disk in frame N+1, is a complex question, in which spatial features interact over time. For example, it is known that the similarity of the elements of a Ternus-Pikler display can bias perception towards a static or moving interpretation, by changing the spatio-temporal grouping of the elements (e.g**.,** Petersik & Rice, 2006, 2008; Hein & Cavanagh, 2012; Hein & Moore, 2012). In my experiments, random changes of the contrast polarity and shape of the stimulus did not affect performance, showing that dissimilarity between the elements did not change the percept. However, I found that nonretinotopic perception was severely impaired when the motion predictability of the individual elements *within* the moving group was decreased. Subjective ratings showed that the more dissimilar the motion of the individual elements was, the less they were perceived as a coherent group. These results support our theoretical model (Öğmen & Herzog, 2010), in which common motion is the main determinant of which elements group across space and time. The group's motion is subtracted from the individual elements' motions to compute the nonretinotopic percept. Here, the unpredictable motion of individual elements prevented them from being perceived as a group. Consequently, the reference motion could not be computed and nonretinotopic perception broke down, likely because the object correspondence problem could not be solved before the next stimulus frame was presented. That is, the visual system could not establish which disk in frame N is which disk in frame N+1. The results are in line with the assumption of the model, that shared

motion between the elements in the visual scene is the crucial grouping factor, rather than their appearance.

To understand the mechanisms involved in nonretinotopic perception, it is also important to understand how and at what stage it is determined whether the retinotopic and nonretinotopic interpretation of the visual scene is perceived. Early visual areas of the brain process visual information in retinotopic coordinates. So the retinotopic and nonretinotopic representation of the visual scene must exist in parallel, even when only the nonretinotopic one is actually perceived. Previous experiments (Boi et al., 2009, 2011) have found retinotopic aftereffects, although the feature in question was *perceived* nonretinotopically, indicating that the processing of retinotopic information continued unconsciously. For example, a Ternus-Pikler display with three Gabor gratings moved left-and-right in group motion. The gratings were perceived to invert their drift direction with every new frame. In retinal coordinates, the gratings drifted coherently up or down. The coherent drift produced strong retinotopic motion aftereffects, although it was likely not perceived. This would mean that the retinotopic interpretation of the image must have been processed unconsciously in the background. However, it is unsure whether the retinotopic drift was indeed not perceived, because the experiments used actual drift that continued within the individual frames.

I addressed this question using an adapted dot motion Ternus-Pikler display. The dots in the display change position from frame to frame, but stay in the same position within each frame. Hence, motion can only be perceived as apparent motion across frames, but never within one frame. The perceived object correspondence determines whether the stimulus is perceived in retinotopic or nonretinotopic coordinates. I adapted the display to contain a retinotopic rotation that could be in the same or opposite direction of the nonretinotopic rotation. My results confirmed that when the nonretinotopic rotation is perceived, the retinotopic motion is almost invisible, and vice versa. In addition, I discovered a complex interaction between the retinotopic and nonretinotopic motion: The nonretinotopic rotation could well be reported when a retinotopic rotation was rotating in the same direction. However, if the retinotopic rotation rotated in the opposite direction, the nonretinotopic rotation percept was severely impaired. Nonretinotopic perception was even more impaired when two, instead of one, retinotopic rotations turned in the opposite sense.

This shows that the invisible retinotopic interpretation of the visual scene continues to be processed unconsciously and that it can influence the conscious nonretinotopic percept. This influence cannot be explained at a local level, because the negative influence of a retinotopic rotation in opposite sense is unbroken when a second retinotopic rotation in the same sense as the nonretinotopic rotation is present. Hence, the retinotopic rotations cannot interact at a local level, for example by rotation direction sensitive cells that inhibit each other. The influence must hence take place at a high-level interpretation stage, that interprets the retinotopic motion signal depending on the context (in this case, whether the third disk is presented or not). This requires a flexibility that is hard to reconcile with retinotopic models

of visual processing. The results therefore suggest that the interaction takes place in nonretinotopic coordinates. In future experiments, I want to investigate whether the effect occurs also with simple linear dot motions. It is possible to design a Ternus-Pikler display with linear retinotopic and nonretinotopic dot motions to test this directly. The findings also have wide ranging implications for the study of unconscious processing, which is often explained within retinotopic models of vision. For example, in masking, the mask is assumed to overwrite the target in early visual areas (Enns & Di Lollo, 2000; Breitmeyer & Öğmen, 2006). My findings raise the question whether the interaction between target and mask is indeed taking place in an early retinotopic processing stage, or whether it can also be explained by later, nonretinotopic mechanisms.

Although the exact mechanisms underlying nonretinotopic perception are still not very well understood, it may be clear that its computation is highly complex. Solving the correspondence problem alone requires taking into account a host of interacting factors that have to be weighted and integrated over time. In schizophrenia, the integration of visual information over space and time is strongly impaired (e.g., Butler, Silverstein, & Dakin, 2008). As a result, the perception of certain types of motion is often deficient in schizophrenia patients. For example, the perception of coherent motion in random dot kinematograms is impaired in patients, but the perception of motion in drifting gratings is not (e.g., Chen et al., 2003; Chen, Bidwell, & Holzman, 2005). This finding has been explained by the complexity of the task: The perception of coherent motion in a random dot kinematograms is complex, because it requires the integration of the coherently moving dots and to perceptually separate it from the randomly moving context. The perception of motion in drifting gratings is simple, because it requires neither integration nor context suppression. Similarly, perception of *apparent* motion is complex, because it requires the integration of two or more inducers over space and time. Apparent motion perception should hence be impaired in schizophrenia patients, and deficits have indeed been reported in some (Saucer & Deabler, 1956; Saucer, 1958, 1959; Chambers & Wilson, 1986; Sanders et al., 2013), but not other studies (Sanders et al., 2012; Saucer, 1958; Tschacher et al., 2008).

The Ternus-Pikler display combines apparent motion with complex processing. According to the literature, its perception should hence be severely impaired in schizophrenia patients. However, I found that both nonretinotopic group motion and dot rotation perception was comparable to healthy, age-matched control participants. This finding questions the assumption that the perception of random dot kinematograms and similar stimuli is impaired because they require complex processing. We suggest that this controversy is explained by a defining difference between the Ternus-Pikler display and other complex motion stimuli: Nonretinotopic perception in the Ternus-Pikler display is a highly natural process, analog to the perception of any kind of moving object in the real world. On the other hand, random dot kinematograms and similar displays are very artificial stimuli, to which far fewer real world situations are comparable. We therefore suggest that not complex motion perception per se is deficient in schizophrenia, but rather that mediating factors are

impaired. One such factor might be attention. For example, although nonretinotopic processing is highly complex, it is so relevant to normal visual perception that it is likely to be implemented in a largely automatized fashion. Situations similar to the integration of coherently moving dots in a random dot kinematograms are far less frequently encountered in the real world, and might therefore be less automatized and require a higher degree of attention.

In summary, my results contribute to a better understanding of the mechanisms underlying nonretinotopic perception in humans. They support the theory of a two-stage computation, in which first the objects in a visual scene are grouped across space and time, primarily based on their common motion, rather than their similarity. Second, the common motion of the group is subtracted from the retinotopic motion of its elements to compute nonretinotopic representations. My results suggest that this computation is fundamentally different from the mechanisms involved in nonretinotopic perception across eye movements, because it is independent of predictability. Nonetheless, it is extremely fast and efficient, at least for the simple stimuli used in my experiments. The reason for the efficient implementation of nonretinotopic perception is probably its ubiquitous importance in everyday life vision. The retinotopic image is very unstable and therefore not very useful as a basis for interactions with the real world, or predictions of where moving objects, predators, or mates are headed. Evidently, these are important factors for the evolutionary success of an organism, that are greatly facilitated by stable nonretinotopic representations. Nonretinotopic perception is hence a fundamental aspect of natural visual perception, which is probably why it is spared in schizophrenia.

The drive to understand nonretinotopic processing is not only fueled by curiosity and the motivation to understand our most important sense. It also has direct implications for other fields of research, like engineering and computer science. For example, computer vision has made great progress in recent years, but remains a bottleneck in the automation of even simple work. Most models of vision, either implicitly or explicitly, rely on a retinotopic architecture. Famous examples include models of visual search (Treisman, 1988), attention (Itti & Koch, 2001), and object recognition (Serre & Poggio, 2007). For example, Treisman's visual search model assumes that the features of an object are presented in separate retinotopic feature maps, and that they are linked through a master map on which attention operates. The other models employ an architecture in which neurons in higher areas pool information from neurons in lower areas in a hardwired, retinal-position dependent fashion. While these models are based on a state-of-the-art understanding of visual processing in the early visual areas of the brain, they fail to capture the higher visual processing leading to nonretinotopic perception, that I have studied here. This might be one of the reasons why computer models of vision are still failing to achieve human-like performance.

# Chapter 4   Achievements in other fields

## 4.1   Prospective Memory Impairment in Healthy Aging

*This subchapter summarizes the submitted manuscript Ballhausen*, N., Lauffs*, M. M., Herzog, M. H., & Kliegel, M. (submitted). Investigating prospective memory through the eyes: No evidence for a monitoring deficit in older adults.*
*(Appendix 5)*

*Contributions: I designed the experiment, programmed the stimulus, analyzed, and interpreted the data, and wrote the manuscript. (*equal contributions)*

### 4.1.1   Introduction

Remembering future intentions, while being engaged in an ongoing activity, is referred to as *prospective memory* (Kliegel et al., 2016). Most actions we plan to perform in the future are prospective memory tasks. For example, we might want to remember attaching a file while writing an email, to post a letter while walking home from work, to buy toilet paper next time we're grocery shopping, to pay the electricity bill next time we're logged in to our online banking system.

Especially in high age, prospective memory is a major determinant of the quality of life, personal independence, and physical well-being of the individual. He who forgets to pay the electricity bill will find himself in the dark. She who forgets to buy food will go hungry. In addition, the often higher need of medical attention required by an aging body increases both the number and importance of prospective memory tasks that an individual has to deal with: Body functions have to be controlled regularly, doctor's appointments are not to be missed, and medicine has to be taken at prescribed moments.

While even younger adults sporadically forget to perform a planned action, prospective memory becomes progressively worse with age, leading to an impairment of old (above 65-70 years) compared to younger adults (Zimmermann & Meier, 2006; Kliegel, Mackinlay, & Jäger, 2008). It is important to study this age-related decline to inform the design of tools and trainings that allow the aging population to maintain an independent life-style for as long as possible. Unfortunately, the mechanisms behind the age-effect in prospective memory are still little understood.

The literature describes two complementary routes to successful prospective memory task performance (McDaniel & Einstien, 2007). The first is to *monitor* the environment for the right occasion to perform the intended action. For example, you could

check the clock every 5 mins to see if it is time to take your medicine. Or if your car is running low on fuel, you might actively start searching for gas stations while you drive. Second, a *spontaneous retrieval* of the intended action can occur when an occasion to perform the action is encountered. For example, seeing a gas station might remind you of your empty gas tank and your intention to refuel, even though you were not actively looking for one.

Active monitoring of the environment takes mental effort and competes with the ongoing task for cognitive resources. For example, the time and attention spent on looking for gas stations is not available for driving. Since, famously, the pool of cognitive resources becomes increasingly limited with age, it was suggested that a monitoring deficit lies at the heart of the age-related prospective memory decrease. More specifically, it is assumed that the ongoing task uses up the available resources, so that they are unavailable for the concurrent monitoring for action cues. On the other hand, spontaneous retrieval mechanisms are thought to effortlessly and automatically trigger the recall of the intended action. Since automatic mechanisms require only a very limited amount of cognitive resources, no age-effect should occur when spontaneous retrieval mechanisms can be used.

We tested these predictions in 31 old (M = 69.9 years, SD = 5.2) and 30 young observers (M = 22.6 years, SD = 2.7) using eye tracking. In each trial, three colored letters were presented: One in the screen center, one horizontally offset 17° to the left, and one 17° to the right (Figure 19). Each of the three letters was surrounded by different letters in different colors, inducing "visual crowding" of the central letter (e.g., Whitney & Levi, 2011). As a consequence, the central letters could only be identified when looked at directly, but not when viewed peripherally. We ensured that at stimulus onset the gaze of the observer was always directed at the letter in the screen center, by presenting the stimulus only after the eye tracker detected a fixation on the central fixation point. The ongoing task of the participant was to indicate whether this letter was a consonant or vowel.

First, we trained participants in the ongoing task and established individual baseline performance levels with only the ongoing task. Each participant then performed four randomly ordered blocks in which a prospective memory task was added. The prospective memory task consisted in the detection of a target attribute, that had to be signaled by pressing a third button *before* performing the ongoing task. When no target was presented, no button had to be pressed and the ongoing task had to be performed immediately. Only 12% of the trials contained a target, so that in most trials only an ongoing task response was required, which promoted a forgetting of the prospective memory task.
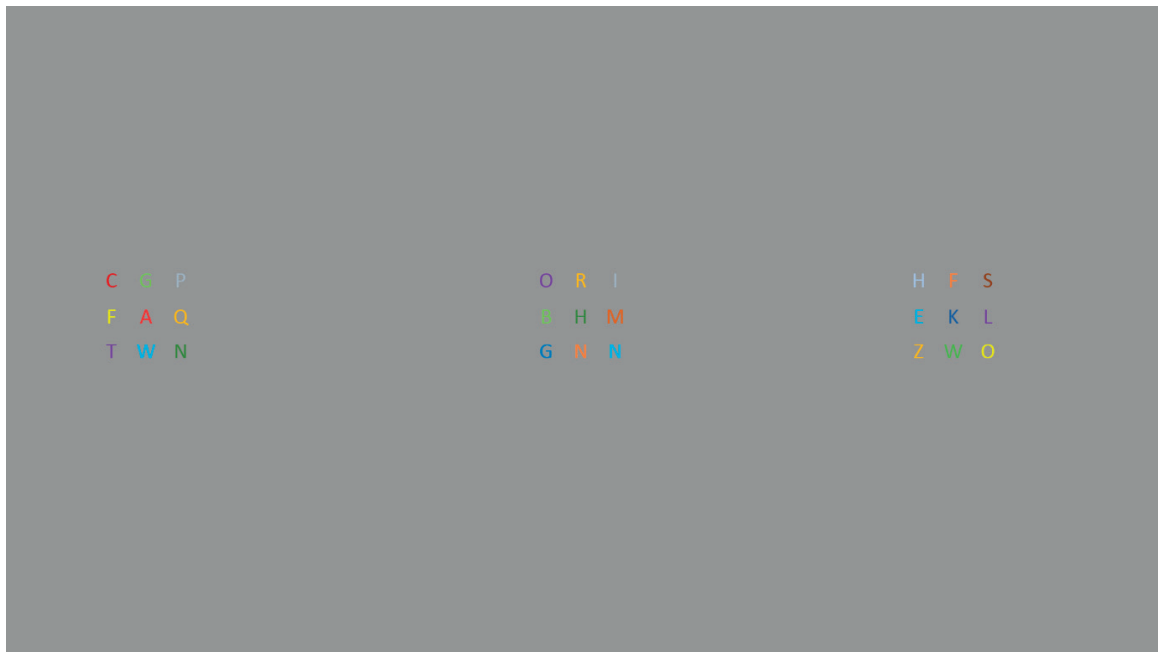
*Figure 19 Prospective memory experiment: Stimulus. Three colored letters, surrounded by other letters, were presented in the screen center and 17° to the left and right. As an ongoing task, the observer indicated if the letter in the center of the screen (here: a green H) was a consonant or a vowel. As prospective memory task, the observers first had to push a third button when they detected the presence of a target attribute. To qualify as target, the attribute had to be present in the same letter that was relevant for the ongoing task, or in the letter on the right-hand side (here: a blue K). At stimulus onset the observer's gaze was always directed at the letter in the screen center. The letter on the right, as well as its color, could not be identified unless looked at directly, because it was crowded by the surrounding letters. Hence, to check for targets on the right, an eye movement had to be made. Using eye tracking, we observed and quantified whether the observers monitored the right-hand side letter for target attributes. The letters on the left-hand side were fully irrelevant distractors. © CC-BY 4.0*

Targets were defined differently depending on the block. In one block, a target was said to be present if the letter in the screen center was a U or an R, but not if it was any other letter. In another block, a target was said to be present if the letter in the screen center was either black or white, but not if it had any other color. We expected a more pronounced performance decrement in older observers, because the color is fully irrelevant to the ongoing task. It is hence less likely to trigger spontaneous retrieval of the prospective memory task and requires active monitoring. In additional blocks, instead of the letter in the screen center, we defined the letter in the right periphery as relevant for the prospective memory task (detection of U or R, or black or white, respectively). This letter could only be identified when looked at directly, hence requiring an eye movement away from the initial central gaze position, that we could easily detect in the eye tracking data. The relevant letter for the ongoing task was still in the screen center, so that there was no other sensible reason to look to the right than the monitoring of targets in that position.

When faced with monitoring demands, we expected older adults to be more impaired than younger adults. Differences between young and old observers should be small when letters have to be detected in the central position, because the letter is also processed in the ongoing task and is therefore likely to trigger spontaneous retrieval of the prospective memory task. However, we expected older adults to be more impaired than younger adults, when the prospective memory task required monitoring of either a stimulus attribute (color) or stimulus position (right periphery) that is irrelevant to the ongoing task. Performance should be worst, and the age-difference greatest, when both the stimulus attribute and position were irrelevant to the ongoing task, and therefore had to be monitored.

## 4.1.2 Results

In the baseline block with only the ongoing task, both young and old adults performed at near 100% accuracy. We then looked at the change in ongoing task accuracy when a prospective memory task was added (Figure 20a). The younger subjects showed stable ongoing task performance, independent of which prospective memory task was performed. However, ongoing task performance of the older adults decreased in the blocks where the right-hand side position had to be monitored.

Next, we looked at prospective memory task performance, by analyzing the percentage of missed targets (Figure 20b). The younger observers had a stable, low miss rate across conditions. For older adults, the percentage of missed targets increased in accordance with the amount of required monitoring. When letters had to be detected in the central position and the monitoring need was low, the older observers' miss rate was as low as that of younger observers. However, when either the color or the position of the target had to be monitored, the older observers' miss rate increased, and reached its peak when both had to be monitored.
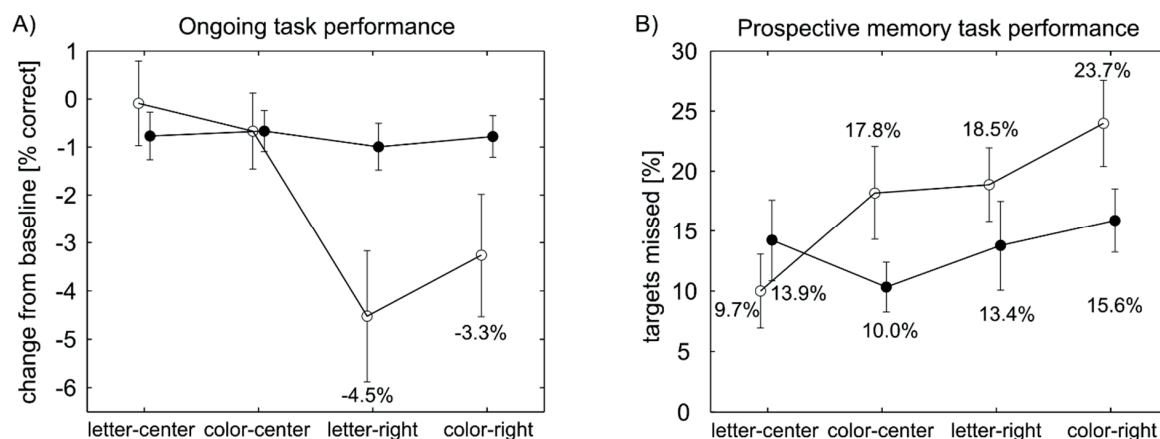
*Figure 20 Prospective memory experiment: Results. A) Ongoing task performance in the blocks with prospective memory task. Depicted is the change in accuracy compared to the baseline block in which only the ongoing task, and no prospective memory task, was performed. B) Prospective memory performance. Depicted is the percentage of targets that were not reported by the observer. The abscissa indicates what stimulus attribute and which stimulus position was relevant in the respective block. Error bars depict the within-subject standard error of the mean (Cousineau, 2005). © CC-BY 4.0*

Hence, both ongoing task and prospective memory performance was stable in the young group, but decreased in the old group when monitoring was required. This pattern of results fits the prediction of the literature, that older adults dispose of insufficient free cognitive resources to initiate efficient monitoring behavior. However, an analysis of the eye tracking data refutes this interpretation: In the conditions where targets were presented on the right, we were able to detect fixations on the target location in more than 95% of all trials in both young and old participants. In addition, 90% of the targets that were missed had been looked at. Hence, participants of both groups monitored very well, and missed targets although they had looked at them.

### 4.1.3 Conclusions

The literature predicted that older adults would perform worse than younger adults on prospective memory tasks comprising a monitoring need. This suggestion rested on the assumption that the initiation of monitoring behavior requires cognitive resources, which are more restricted in high age.

Using similar *behavioral* measures as previous studies, we could replicate the behavioral findings that have been cited as evidence for this suggestion in the past. However, unlike previous studies, we added conditions in which monitoring was overtly observable as an eye movement, rather than being purely cognitive in nature. We found that both young and old adults monitored for targets in almost all trials. Nonetheless, individuals of both groups, but older individuals in particular, failed to respond to the presence of some of the

targets. The majority of these misses (90%) occurred although the participant had looked at the target.

Our findings therefore refute the suggestion that age-related differences in prospective memory performance are rooted in an inability to initiate monitoring behavior. Rather, it suggests that post-monitoring processes fail. These might include the retrieval of the prospective memory task ("I remember I have to look here, but why again?"), action selection ("I know this is a target, but what am I supposed to do again?"), inhibition of ongoing task execution, and task switching between the ongoing and the prospective memory task. Future work will have to determine which of these processes are responsible for the decline of prospective memory performance in aging and whether the same processes cause prospective memory failure in old and young.

## 4.2 Neural Correlates of Risk Prediction Error

*This subchapter summarizes the manuscript*
*Lauffs, M. M., Geoghan, S., Favrod, O., Herzog, M. H., & Preuschoff, K. (in preparation).*

*Contributions: I designed the experiment, programmed the stimulus, collected, analyzed, and interpreted the data, and wrote the manuscript.*

### 4.2.1 Introduction

In today's world, the need to make choices is more ubiquitous than ever. For example, the simple choice of a breakfast cereal has to be made from dozens of competing brands. To choose sensibly, it is important to learn from past experiences. How satisfied was I with brand X last time, should I buy the more expensive brand Y this time? This kind of learning is hardwired in the primate brain: Neurophysiological studies in macaque monkeys have revealed that the firing of midbrain dopamine neurons correlates strongly with the difference between the magnitude of reward that the animal expected and the magnitude that it received (Schultz, 1998). This *reward prediction error* signal is the basis of reinforcement learning and motivated behavior. It occurs not only in response to the reward (or punishment) itself, but transfers to the earliest stimulus *predicting* it. In humans, dopamine projection sites can be visualized in fMRI recordings.

However, neither human nor nonhuman primates simply choose the most rewarding option. Instead they balance their choices against other factors, such as the associated risks. In a series of fMRI experiments, Preuschoff and colleagues unveiled the correlates of reward and risk processing in the human brain (Preuschoff, Bossaerts, & Quartz, 2006; Preuschoff, Quartz, & Bossaerts, 2008). Participants played a simple card game (Figure 21). In each trial, two cards, randomly chosen from a deck of ten cards with the values 1-10, were uncovered one after another. Before the first card was uncovered, the participant bet if the second card would be higher or lower than the first. Since the first card had not been seen yet, the

probability of winning was 50% for either bet at that point. When the first card was uncovered, the probability of winning changed depending on the chosen bet and the value of the first card. Expected reward magnitude and risk depend on the probability of winning. In economic theories, expected reward is computed as "probability of winning (times) reward magnitude". Since in the experiment 1 CHF could be won in every trial, expected reward increased *linearly* with the probability of winning. For example, when the first card had a value of 2 and the participant bet that the second card will be higher, eight of the nine remaining cards would lead her to win (probability of winning $\frac{8}{9}$, i.e., 89%). Risk, on the other hand, is defined as *outcome uncertainty*. It was hence minimal when the probability of winning was either 0 or 1 (Card 1 values 1 or 10, i.e., certain win or loss), and maximal when it was 0.5 (Card 1 values 4 or 5, i.e., 44% or 56% of winning or losing). Risk hence varied *parabolically* with the probability of reward. The second card resolved the gamble.

By analyzing the BOLD signal in the period between the first and second card, it was found that activity in the ventral striatum reflected both risk and expected reward, albeit at different timeframes. During the first second after Card 1, the signal increased linearly with the probability of winning, hence correlating with expected reward. However, after the first second the signal varied parabolically, correlating strongly with risk (Preuschoff et al., 2006). Activity in insular cortex reflected risk prediction error magnitude, that is, the difference between the expected and actual risk (Preuschoff et al., 2008).

Here, we adapted the paradigm for use with EEG (Figure 21). We recorded high-density EEG with 192 active electrodes from 20 healthy, right-handed participants and analyzed the data from a mid-central electrode cluster, that was formed by averaging electrode Cz and the five surrounding channels.

## 4.2.2 Results

Both card presentations were followed by pronounced ERPs (Figure 22a). We found that the amplitude of a late sustained positive potential was highly sensitive to risk prediction error magnitude (Figure 22b-c). We plotted the mean amplitudes of this component as a function of the probability of winning (Figure 22d-e). After Card 1, mean amplitudes varied parabolically with the probability of winning, hence showing the characteristic pattern of risk prediction errors that was also observed using fMRI. After Card 2, the risk prediction error magnitude depends on both the probability of winning and the outcome - unlikely events entail a larger and likely events a smaller risk prediction error. Analog to risk prediction error magnitude, mean amplitudes increased linearly with the probability of winning on loss trials and decreased linearly on win trials.
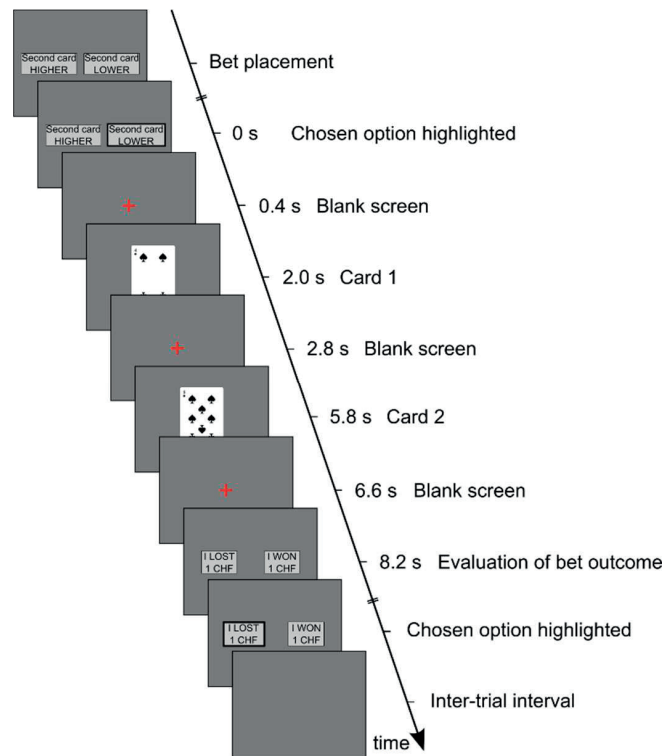
*Figure 21 Neural correlates of risk anticipation: Paradigm. In each trial, two cards were drawn randomly without replacement from ten cards with the values 1-10. Before the first card was shown, participants made a bet whether the second card would be higher or lower than the first card. Each Card 1 value occurred equally often per block, so the probability of winning at the moment of the bet was 50%. The participant received a brief visual confirmation of the chosen bet (0.4 s), followed by a blank screen (1.6 s), during which only a central fixation cross was presented. Card 1 was then presented (0.8 s) and the probability of winning changed depending on the bet and card value. For example, if Card 1 was a 4, the probability of winning changed from 5/10 to 3/9 for a "Second card lower" and to 5/9 for a "Second card higher" bet. After a prolonged blank-screen interval (3.0 s), during which the participant anticipated the outcome, Card 2 was presented (0.8 s), resolving the gamble. After another blank screen (1.6 s), the participant evaluated whether she won or lost 1 CHF.*

*Reprinted from Lauffs, Geoghan, Favrod, Preuschoff, & Herzog (in prep.).*

Apart from the late positive component, first amplitude differences between uncertain (and hence risky) and certain events were reflected as early as 200 ms after presentation of Card 1 and 150 ms after Card 2. This suggests that the brain rapidly differentiates between risky and non-risky options. In addition, this early response might be part of a salience signal, that has been reported in the early responding of dopamine cells for rewards.

We computed inverse solutions to estimate the neural sources giving rise to this activity (Figure 23). For both early and late differences, we found sources in or near insular cortex and the superior temporal lobe, which is in line with the fMRI correlates of risk prediction error magnitude (Preuschoff et al., 2008).

Finally, we correlated the differences in EEG amplitude with concurrently recorded pupil size measurements (Figure 24). Pupil size has been shown to correlate with risk prediction error magnitude in an auditory version of our paradigm (Preuschoff, t' Hart, & Einhäuser, 2011). The presentation of each card caused large fluctuations in pupil size due to the pupillary light response, in this case constrictions in response to the white cards that had a higher luminance than the grey background (Figure 24a). However, the change in pupil size also depended on risk prediction error magnitude (Figure 24 b-c). At the time point where this influence was most significant (Figure 24 b-c, lower), we computed the mean pupil size change per probability of winning. Except for loss trials following Card 2, the mean change in pupil size correlated strongly with the mean amplitudes observed in the EEG.

### 4.2.3 Conclusions

While the EEG correlates of reward processing are rather well investigated, very little is known on how risk processing is reflected in the EEG. With the current study we close this gap. We found that the amplitude of a late, sustained ERP component correlates with risk prediction error magnitude. The reconstructed sources of this activity match the sources observed in fMRI studies using the same paradigm. A correlation with pupil size, which has been previously shown to signal risk prediction error magnitude, provides further evidence that we indeed identified the EEG correlates of risk prediction error.
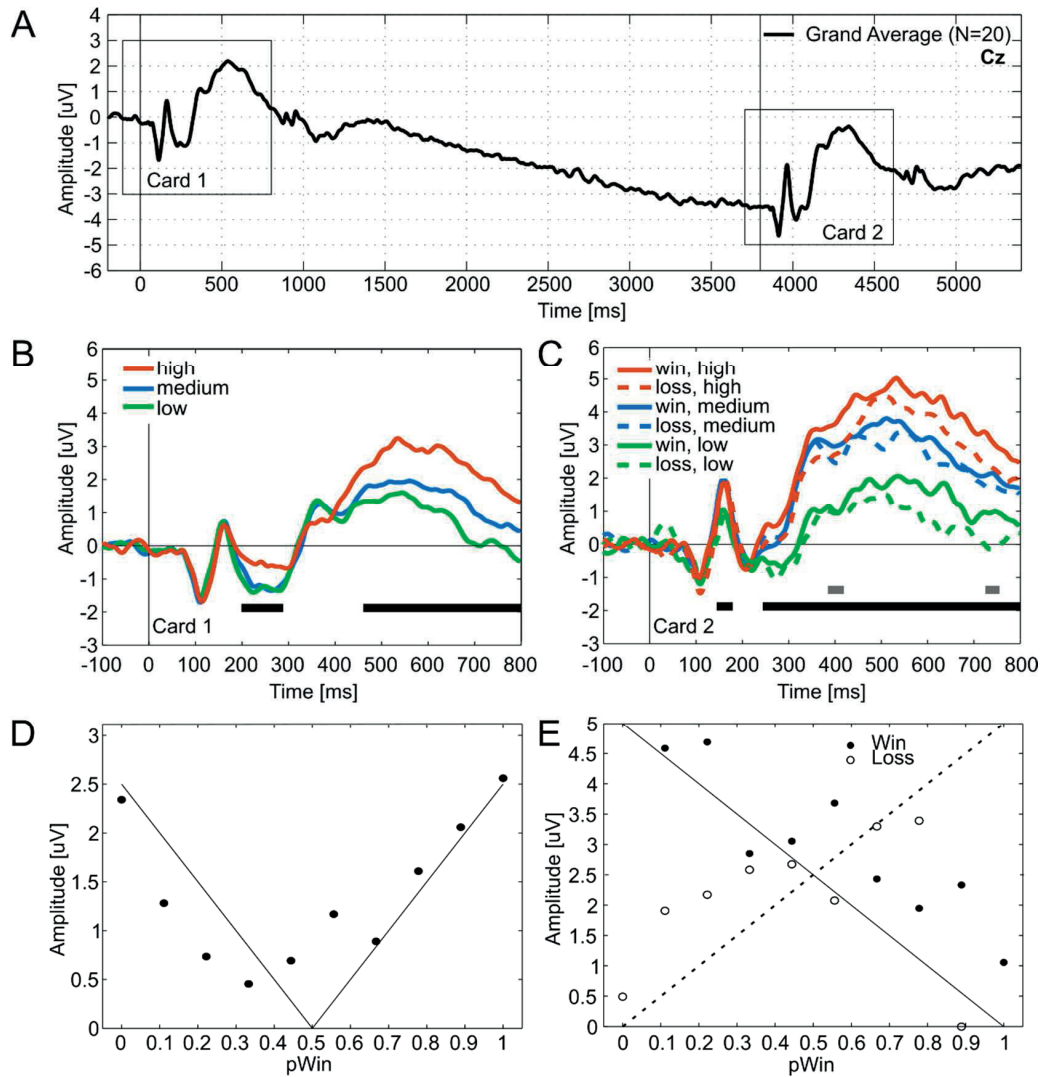
*Figure 22 Grand average ERP waveforms, recorded from an electrode cluster around Cz and baseline corrected to the depicted pre-stimulus interval (i.e., 200 ms in A, 100 ms in D-G). **A.** ERP activity from -200 to +5400 ms relative to the onset of Card 1, comprising both card presentations. Both card presentations were followed by pronounced event-related potentials. Anticipation of Card 2 was reflected in a sustained negativity that ramped up until the presentation of Card 2. **B.** Card 1 (-100 – 800 ms): The amplitudes of an early negative and a late positive component varied as a function of risk prediction error magnitude. Amplitudes were more positive on trials with high (red), than both medium (blue) and low (green) risk prediction error. Black bar: Significant main effect of risk prediction error magnitude (high, medium, low). **C.** Card 2 (-100 – 800 ms): The amplitudes of an early positive and a late positive component were enhanced for trials with high (red), compared to medium (blue), and low (green) risk prediction error. The effect of the bet outcome is comparably minor, with amplitudes being only slightly more positive on win (solid lines) than loss trials (broken lines). Black bar: Significant main effect of risk prediction error magnitude (high, medium, low). Grey bar: Significant main effect of outcome (win, loss). **D.** Mean EEG amplitude during the late significant interval after Card 1 (460-800 ms), plotted against the probability of winning after Card 1. The V-shaped model of risk prediction error is overlaid as solid lines and fitted well by the data. **E.** Mean EEG amplitude during the late significant interval after Card 2 (243-800 ms), plotted separately for win (filled markers) and loss trials (empty markers) against the probability of winning after Card 1. The model or risk prediction error is overlaid as solid (win) and broken (loss) lines and is again well fitted by the data. Reprinted from Lauffs, Geoghan, Favrod, Preuschoff, & Herzog (in prep.).*
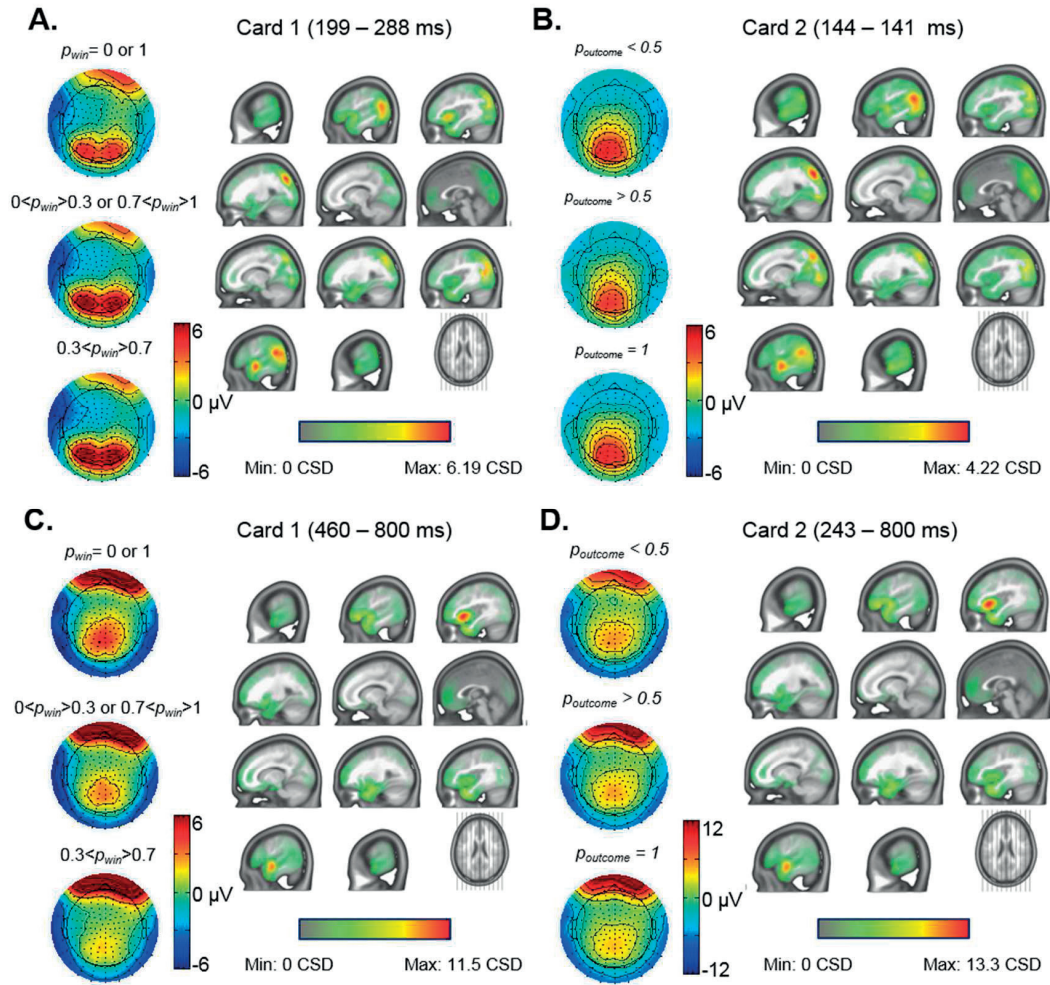
*Figure 23 Left: Scalp topographies for trials with high (top), medium (middle), and low (bottom) risk prediction error. In all time-windows, differences are only quantitative. Right: Mathematical source reconstructions (inverse solutions) of the grand average EEG activity observed on the scalp. **A.** Early significant interval (199-288ms) after Card 1 onset: Scalp topographies have lateralized occipital foci. The inverse solutions show maximum activation in the left and right middle temporal lobe, right superior occipital lobe, the left superior temporal lobe, and in the right insula. **B.** Late significant interval (460-800ms) after Card 1 onset: Scalp topographies have a post-central focus. Inverse solutions show maximum activation in the left superior temporal lobe and the right insula. **C.** Early significant interval (144-181 ms) after Card 2 onset: Scalp topographies have a central occipital focus. Inverse solutions show maximum activation in the left precuneus and the left calcarine sulcus. **D.** Late significant interval (243-800 ms) after Card 2 onset: Scalp topographies have a post-central focus. Inverse solutions show maximum activation in the left and right insula and the frontal lobe. Reprinted from Lauffs, Geoghan, Favrod, Preuschoff, & Herzog (in prep.).*
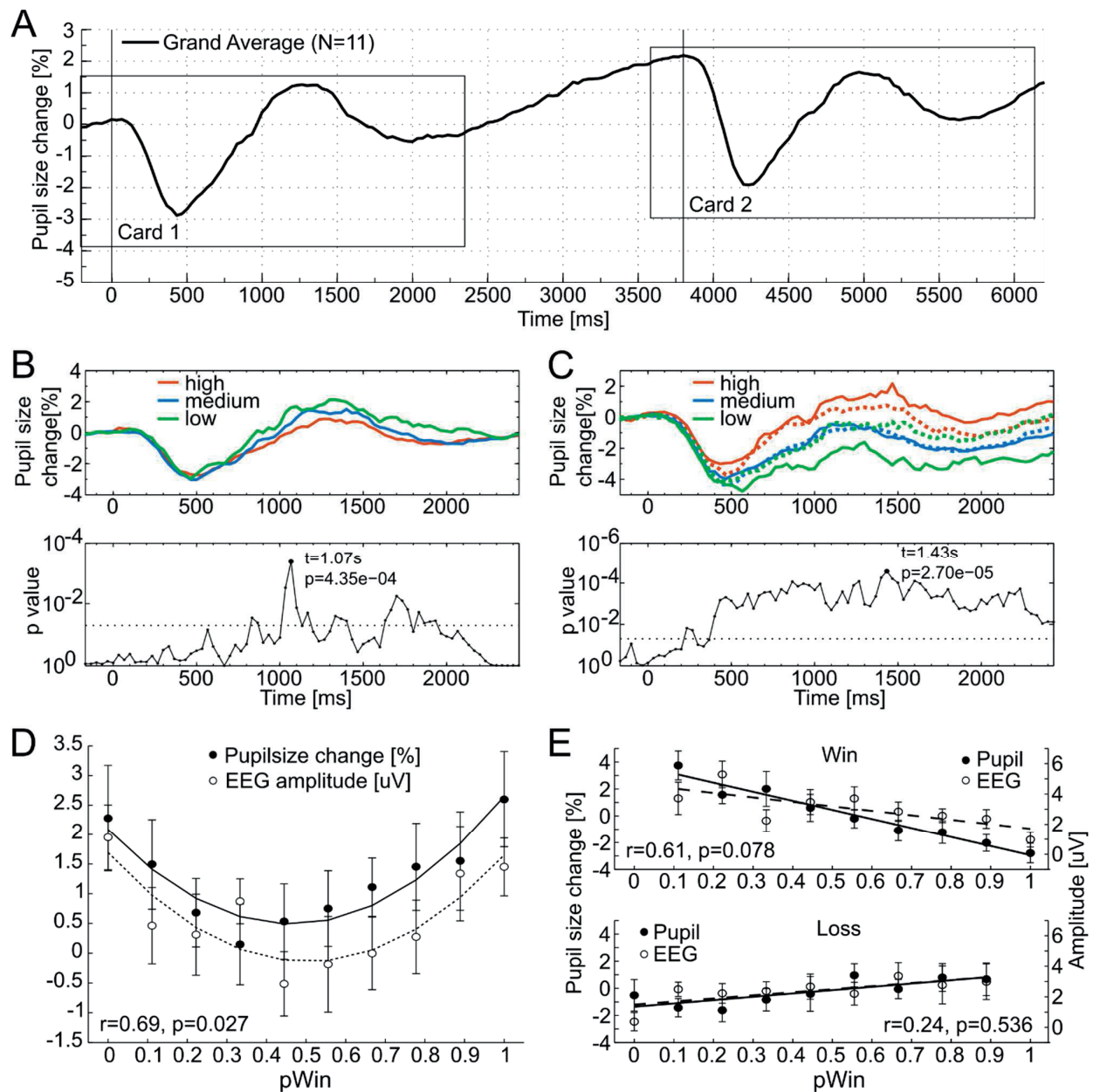
*Figure 24 **A.** Pupil size from -200 to +6200 ms relative to the onset of Card 1, comprising both card presentations. Upon the presentation of both cards, the pupil initially constricted, but quickly started to dilate again. **B.** Card 1 (-200 – 2400 ms). Top: Pupil size change over time. The pupil dilates slightly more in high (red) than medium (blue), and low (green) risk prediction error trials. Bottom: p-values for a main-effect of risk prediction error (sample-wise RM-ANOVAs).The sample with the lowest p-value is selected for further analysis. **C.** Card 2 (-200 – 2400 ms). Top: On win trials (solid lines) the pupil is more dilated on high (red), than medium (blue), and low (green) risk prediction error trials. Pupil size varied comparably little on loss trials (broken lines). Bottom: Significance as described in B. **D-E.** Pupil size at the most significant sample (filled markers) correlates strongly with the mean EEG amplitude of the late significant interval for the same subjects (empty markers). Error bars are +/- 1 SEM. Reprinted from Lauffs, Geoghan, Favrod, Preuschoff, & Herzog (in prep.).*

## 4.3 Eye tracking of Renaissance art

*This subchapter summarizes unpublished data of an ongoing project. I provide a preliminary version of the methods section in the appendix.*
*(Appendix 6)*

*Collaborators: Marc M. Lauffs, Darius Karacsony, & Michael H. Herzog.*
*Contributions: I designed the experiment, programmed the stimulus, collected, analyzed, and interpreted the data, and wrote the manuscript.*

### 4.3.1 Introduction

Many of the most highly appreciated paintings in art history were created during the Renaissance period. At the time, artists and painters were schooled according to strict guidelines. One of these guidelines concerned the embedding of visual narratives, so called *historiae* in the painting. The observer should discover these narratives bit by bit when viewing the image. Our collaborator from EPFL's architecture department hypothesized that there is a link between the enduring appreciation of the paintings from this era and the visual narratives. He suspected a similar link between highly appreciated works in other art forms, notably architecture and abstract painting, and the sequence in which they are viewed. His goal was to train young architects to include visual narratives in their designs, to produce better, more appealing work.

Here, we tested if there are indeed visual narratives embedded in Renaissance paintings that guide the viewers gaze. To that end, we searched for communalities in the gaze paths of different observers viewing the same painting.

### 4.3.2 Methods

Eight observers freely viewed ten paintings on a computer screen for 30 secs each. All observers were students in an architecture Master's course taught by our collaborator. In later sessions, we also presented images of architecture, abstract paintings, and the students' own architectural models, respectively. We used eye tracking to record the viewers' gaze position binocularly at 500 Hz and detected fixations and saccades to reconstruct the gaze path.

Next, I compared the gaze paths using the Matlab toolbox ScanMatch (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010). ScanMatch computes a similarity score for a pair of gaze paths based on the position, duration, and order of fixations (see Figure 25 and Appendix 6 for details). Per pair of gaze paths, a single similarity score is obtained, which can then easily be compared to other pairs.

The most straightforward approach would have been to compare the recorded gaze paths to the historia of the image directly. However, this was not an option because clear predictions of the gaze path could not be derived from the literature.
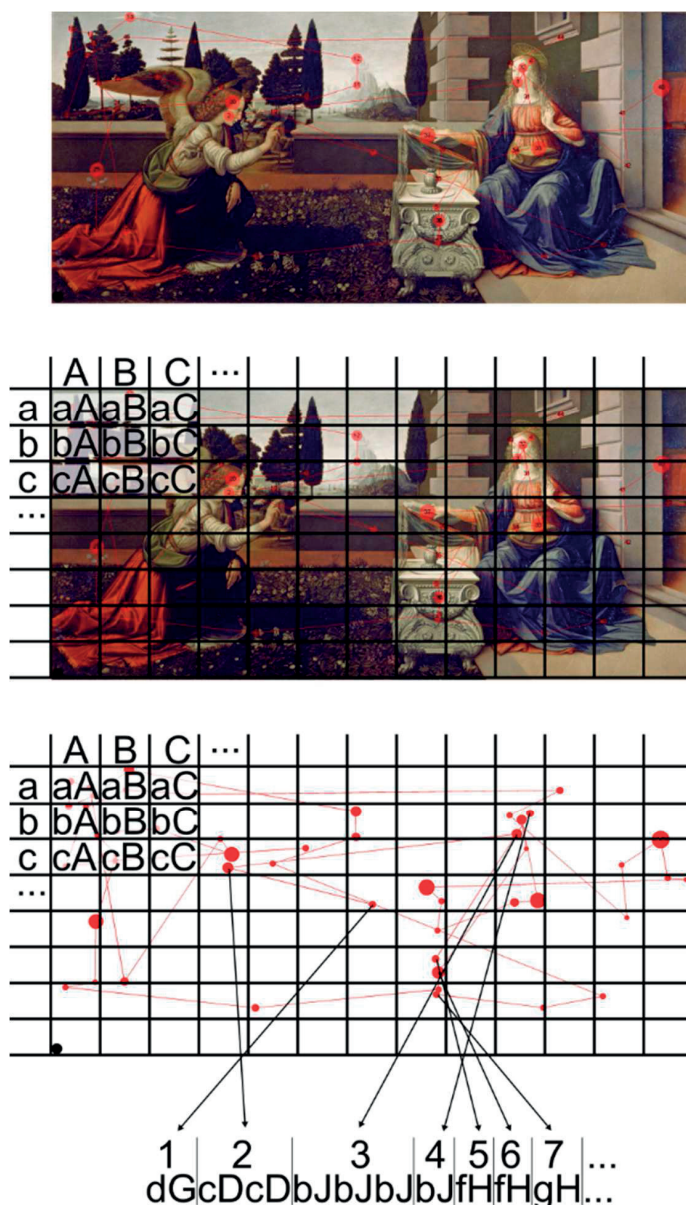
*Figure 25 Renaissance art eye tracking: Schematic of the analysis. Top: Fixations were detected in the raw eye tracking data. Middle: The image was subdivided in a grid of 20 x 11 cells (1.5 x 1.5 ° per cell). Each cell was named with a combination of an upper-case letter indexing the column and a lower-case indexing the row of its position. Bottom: The fixation positions were then coded with the letter indices of the grid cells. For example, the first fixation usually fell in cell 'dG' (row d, column G), the center of the screen, because a central fixation cross was presented before the image. Fixations 3 and 4 are separate fixations in different positions (here, lips and halo of the same person), and are hence coded separately. However, since both positions fall in the same cell they are coded with the same letters. Pairs of gaze paths can then be compared using the ScanMatch plugin for Matlab. ScanMatch computes a single-digit similarity score, which can then be used to determine if certain pairs of gaze paths are more or less similar to each other than other pairs. To compute the similarity score, ScanMatch uses the Needleman-Wunsch algorithm, an algorithm typically used for the sequencing of DNA data. The algorithm determines the most parsimonious way to match the two gaze paths by means of alterations, deletions, and insertions of letters. The relative costs of each type of action have to be set by the user. For more information see Cristino et al. (2010). © CC-BY 4.0*

.

I therefore tried to determine whether the assumption that the gaze is guided by an embedded visual narrative is justified at all. I assumed that if the artist achieved to guide the gaze of the observer, this should overrule idiosyncrasies in the viewing behavior of the individual. Therefore, the gaze path should be influenced more by which image somebody looks at than by who is looking at it. To test this hypothesis, I computed the average gaze path similarity within images, across observers and compared it to the average similarity within observers, across images

Second, I tested whether the similarity of our observers' gaze paths is *higher* than would be expected by chance. To determine the baseline similarity, I generated a large number of gaze paths, in which the direction of the saccades was determined purely by chance. To keep the simulated data comparable to the human data, I constrained the simulated gaze paths to have the same number of fixations, similar fixation durations, similar saccade amplitudes, and the same starting point. For a second set of simulations, I constructed random gaze paths by sampling from the human fixation positions and durations with replacement (bootstrapping). By computing the gaze path similarity of all unique pairs of *human* observers and all unique pairs of *simulated* observers, I could then compare the average similarity of the human gaze paths against two different models of randomness. If the gaze of the observers is indeed driven by some stylistic element that the author embedded in the painting, the gaze paths of different observers should be more similar to each other than gaze paths that are random and, hence, not driven by the image at all.

### 4.3.3 Preliminary results

The results of this analysis showed that the similarity within images is not significantly greater than the similarity within subjects. Neither was the similarity between observers greater than the similarity between the chance models. The results were robust to the parameter changes of the algorithms that I explored so far. Nonetheless, interpretation of these results are not straight-forward. It is tempting to conclude that the images do not contain the hypothesized visual narratives that guide the gaze and the hypotheses are hence wrong. Especially, since the analyses confirm the subjective impression that the gaze paths of different observers on the same image are very dissimilar. On the other hand, the absence of a difference is an extraordinary claim that requires extraordinary evidence. In future analyses, I therefore plan to include artificially altered versions of the observed gaze paths, to demonstrate that the algorithm is working as intended, and indeed capable of detecting similarities between gaze paths. Then, I will rerun the analysis, taking into account only a restricted number of fixations instead of the entire 30 secs recording, to exclude that later data points introduce variability that clouds differences that are present early on. Finally, I will explore further parameters of the ScanMatch algorithm, to ensure as far as possible that the null result is not caused by factors other than the absence of an effect.

# References

Ağaoğlu, M. N., Clarke, A. M., Herzog, M. H., & Öğmen, H. (2016). Motion-based nearest vector metric for reference frame selection in the perception of motion. *Journal of Vision*, *16*(7), 14-14.

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, *30*(8), 2960-2966.

Arcaro, M. J., McMains, S. A., Singer, B. D., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *Journal of Neuroscience*, *29*(34), 10638-10652.

Boi, M., Öğmen, H., & Herzog, M. H. (2011). Motion and tilt aftereffects occur largely in retinal, not in object, coordinates in the Ternus–Pikler display. *Journal of Vision*, *11*(3), 7-7.

Boi, M., Öğmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino-vs. nonretinotopic processing. *Journal of Vision*, *9*(13), 5-5.

Boi, M., Vergeer, M., Ogmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology*, *21*(20), 1732-1737.

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*. Oxford, UK: Oxford University Press.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature*, *371*(6497), 511.

Butler, P. D., Silverstein, S. M., & Dakin, S. C. (2008). Visual perception and its impairment in schizophrenia. *Biological Psychiatry*, *64*(1), 40-47.

Cavanagh, P., Hunt, A. R., Afraz, A., & Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences*, *14*(4), 147-153.

Cavanaugh, J., Berman, R. A., Joiner, W. M., & Wurtz, R. H. (2016). Saccadic corollary discharge underlies stable visual perception. *Journal of Neuroscience*, *36*(1), 31-42.

Chambers, J. L., & Wilson, W. T. (1968). Perception of apparent motion and degree of mental pathology. *Perceptual and Motor Skills*, *26*(3), 855-861.

Chen, Y., Bidwell, L. C., & Holzman, P. S. (2005). Visual motion integration in schizophrenia patients, their first-degree relatives, and patients with bipolar disorder. *Schizophrenia Research*, *74*(2), 271-281.

Chen, Y., Nakayama, K., Levy, D., Matthysse, S., & Holzman, P. (2003). Processing of global, but not local, motion direction is deficient in schizophrenia. *Schizophrenia Research, 61*(2), 215-227.

Chkonia, E., Roinishvili, M., Makhatadze, N., Tsverava, L., Stroux, A., Neumann, K., Herzog, M. H., & Brand, A. (2010). The shine-through masking paradigm is a potential endophenotype of schizophrenia. *PLoS One*, *5*(12), e14268.

Clarke, A. M., Öğmen, H., & Herzog, M. H. (2016). A computational model for reference-frame synthesis with applications to motion perception. *Vision Research*, *126*, 242-253.

Clarke, A. M., Repnow, M., Öğmen, H., & Herzog, M. H. (2013). Does spatio-temporal filtering account for nonretinotopic motion perception? Comment on Pooresmaeili, Cicchini, Morrone, and Burr (2012). *Journal of Vision*, *13*(10), 19-19.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42-45.

Dawson, M. R. (1991). The how and why of what went where in apparent motion: modeling solutions to the motion correspondence problem. *Psychological Review*, *98*(4), 569.

Descartes, R., & Hall, T. S. (1972). *Treatise of man*. Cambridge, MA: Harvard University Press.

Drissi Daoudi, L., Öğmen, H., & Herzog, M. H. (2016). Mandatory feature integration across retinotopic locations. Talk presented at the 39th European Conference on Visual Perception, Barcelona, Spain. Abstract published in *Perception, 45*(2)*, 359.

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.

Duncker, K. (1929). Über induzierte Bewegung. *Psychological Research*, *12*(1), 180-259.

Efron, R. (1967). The duration of the present. *Annals of the New York Academy of Sciences*, *138*(1), 713-729.

Efron, R. (1973). Conservation of temporal information by perceptual systems. *Attention, Perception, & Psychophysics*, *14*(3), 518-530.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex (1991)*, *7*(2), 181-192.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345-352.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, *57*(1), 191-195.

Gezeck, S., Fischer, B., & Timmer, J. (1997). Saccadic reaction times: a statistical analysis of multimodal distributions. *Vision Research*, *37*(15), 2119-2131.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20-25.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649-677.

Hein, E., & Cavanagh, P. (2012). Motion correspondence in the Ternus display shows feature bias in spatiotopic coordinates. *Journal of Vision*, *12*(7), 16-16.

Hein, E., & Moore, C. M. (2012). Spatio-temporal priority revisited: The role of feature identity and similarity for object correspondence in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 975.

Hermens, F., Scharnowski, F., & Herzog, M. H. (2009). Spatial grouping determines temporal integration. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 595.

Johansson, G. (1950). Configurations in the perception of velocity. *Acta Psychologica*, *7*, 25-79.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201-211.

Johansson, G. (1974). Vector analysis in visual perception of rolling motion. *Psychological Research*, *36*(4), 311-319.

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, *38*(4), 379-393.

Kliegel, M., Ballhausen, N., Hering, A., Ihle, A., Schnitzspahn, K. M., & Zuber, S. (2016). Prospective memory in older adults: Where we are now and what is next. *Gerontology, 62*(4), 459-466.

Kliegel, M., Mackinlay, R., & Jäger, T. (2008). A lifespan approach to the development of complex prospective memory. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective Memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives* (pp. 187-216). Mahwah, NJ: Erlbaum.

Kuba, M., Kubová, Z., Kremláček, J., & Langrová, J. (2007). Motion-onset VEPs: characteristics, methods, and diagnostic use. *Vision Research*, *47*(2), 189-202.

Lauffs, M. M., Shaqiri, A., Brand, A., Roinishvili, M., Chkonia, E., Öğmen, H., & Herzog, M. H. (2016). Local versus global and retinotopic versus nonretinotopic motion processing in schizophrenia patients. *Psychiatry Research*, *246*, 461-465.

Lauffs, M. M., Öğmen, H., & Herzog, M. H. (2017) Unpredictability does not hamper nonretinotopic motion perception. *Journal of Vision, 19*(7):6, 1-10.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychological Bulletin*, *81*(12), 899.

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Thousand Oaks, CA: Sage.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research*, *6*(1), 57-77.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS biology*, *3*(8), e265.

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.

Öğmen, H. (2007). A theory of moving form perception: Synergy between masking, perceptual grouping, and motion computation in retinotopic and nonretinotopic representations. *Advances in Cognitive Psychology, 3*, 67-84.

Öğmen, H., & Herzog, M. H. (2010). The geometry of visual perception: Retinotopic and nonretinotopic representations in the human visual system. *Proceedings of the IEEE*, *98*(3), 479-492.

Otto, T. U., Öğmen, H., & Herzog, M. H. (2006). The flight path of the phoenix—The visible trace of invisible elements in human vision. *Journal of Vision*, *6*(10), 7-7.

Otto, T. U., Öğmen, H., & Herzog, M. H. (2009). Feature integration across space, time, and orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1670.

Peterson, M. A., & Kim, J. H. (2001). On what is bound in figures and grounds. *Visual Cognition*, *8*(3-5), 329-348.

Peterson, M. A., & Skow, E. (2008). Inhibitory competition between shape properties in figure-ground perception. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 251.

Petersik, J. T., & Rice, C. M. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the Ternus effect. *Perception*, *35*(6), 807-821.

Pikler, J., (1917). *Sinnesphysiologische Untersuchungen*. Leipzig, Germany: Barth.

Pooresmaeili, A., Cicchini, G. M., Morrone, M. C., & Burr, D. (2012). "Nonretinotopic processing" in Ternus motion displays modeled by spatiotemporal filters. *Journal of Vision*, *12*(1), 10-10.

Pooresmaeili, A., Cicchini, G. M., Morrone, M. C., & Burr, D. (2013). Spatiotemporal filtering and motion illusions. *Journal of Vision*, *13*(10), 21-21.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3-25.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160.

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*(3), 381-390.

Preuschoff, K., Marius't Hart, B., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*.

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, *28*(11), 2745-2752.

Rohrer, W. H., & Sparks, D. L. (1993). Express saccades: the effects of spatial and temporal uncertainty. *Vision Research*, *33*(17), 2447-2460.

Riggs, L. A., Merton, P. A., & Morton, H. B. (1974). Suppression of visual phosphenes during saccadic eye movements. *Vision Research*, *14*(10), 997-1011.

Rolfs, M., Jonikaitis, D., Deubel, H., & Cavanagh, P. (2011). Predictive remapping of attention across eye movements. *Nature Neuroscience*, *14*(2), 252-256.

Rolfs, M., & Szinte, M. (2016). Remapping attention pointers: linking physiology and behavior. *Trends in Cognitive Sciences*, *20*(6), 399-401.

Ross, J., Burr, D., & Morrone, C. (1996). Suppression of the magnocellular pathway during saccades. *Behavioural Brain Research*, *80*(1), 1-8.

Sanders, L. L. O., de Millas, W., Heinz, A., Kathmann, N., & Sterzer, P. (2013). Apparent motion perception in patients with paranoid schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, *263*(3), 233-239.

Sanders, L. L. O., Muckli, L., de Millas, W., Lautenschlager, M., Heinz, A., Kathmann, N., & Sterzer, P. (2012). Detection of visual events along the apparent motion trace in patients with paranoid schizophrenia. *Psychiatry Research*, *198*(2), 216-223.

Saucer, R. T. (1958). A further study of the perception of apparent motion by schizophrenics. *Journal of Consulting Psychology*, *22*(4), 256.

Saucer, R. T. (1959). Chlorpromazine and apparent motion perception by schizophrenics. *Journal of Consulting Psychology*, *23*(2), 134.

Saucer, R. T., & Deabler, H. L. (1956). Perception of apparent motion in organics and schizophrenics. *Journal of Consulting Psychology*, *20*(5), 385.

Saygin, A. P., & Sereno, M. I. (2008). Retinotopy and attention in human occipital, temporal, parietal, and frontal cortex. *Cerebral Cortex*, *18*(9), 2158-2168.

Scharnowski, F., Rüter, J., Jolij, J., Hermens, F., Kammer, T., & Herzog, M. H. (2009). Long-lasting modulation of feature integration by transcranial magnetic stimulation. *Journal of vision*, *9*(6), 1-1.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, *80*(1), 1-27.

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature, 444*(7117), 374-377.

Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience*, *31*, 317-338.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*(6), 482.

Stevens, J. K., Emerson, R. C., Gerstein, G. L., Kallos, T., Neufeld, G. R., Nichols, C. W., & Rosenquist, A. C. (1976). Paralysis of the awake human: visual perceptions. *Vision Research*, *16*(1), 93-IN9.

Ternus, J. (1926). Experimentelle Untersuchung über phänomenale Identität. *Psychologische Forschung, 7*, 81–136.

Thorpe, S. J., & Fabre-Thorpe, M. (2001). Seeking categories in the brain. *Science*, *291*(5502), 260-263.

Thunell, E., Plomp, G., Ögmen, H., & Herzog, M. H. (2016a). EEG correlates of relative motion encoding. *Brain Topography*, *29*(2), 273-282.

Thunell, E., van der Zwaag, W., Öğmen, H., Plomp, G., & Herzog, M. H. (2016b). Retinotopic encoding of the Ternus-Pikler display reflected in the early visual areas. *Journal of Vision*, *16*(3), 26-26.

Tibber, M. S., Anderson, E. J., Bobin, T., Carlin, P., Shergill, S. S., & Dakin, S. C. (2015). Local and global limits on visual processing in schizophrenia. *PloS one, 10*(2), e0117951.

Tolias, A. S., Moore, T., Smirnakis, S. M., Tehovnik, E. J., Siapas, A. G., & Schiller, P. H. (2001). Eye movements modulate visual receptive fields of V4 neurons. *Neuron*, *29*(3), 757-767.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*(2), 171-178.

Tschacher, W., Dubouloz, P., Meier, R., & Junghan, U. (2008). Altered perception of apparent motion in schizophrenia spectrum disorder. *Psychiatry Research*, *159*(3), 290-299.

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology*, *78*(3), 1373-1383.

Vetter, P., Grosbras, M. H., & Muckli, L. (2013). TMS over V5 disrupts motion prediction. *Cerebral Cortex*, *25*(4), 1052-1059.

Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik* (Vol. 9). Hamburg, Germany: Voss.

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, *37*(20), 464-476.

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of Neurophysiology*, *73*(5), 1988-2003.

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, *56*(2), 366-383.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160-168.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*(20), 2070-2089.

Wutz, A., Drewes, J., & Melcher, D. (2016). Nonretinotopic perception of orientation: Temporal integration of basic features operates in object-based coordinates. *Journal of Vision*, *16*(10), 3-3.

Yund, E. W., Morgan, H., & Efron, R. (1983). The micropattern effect and visible persistence. *Attention, Perception, & Psychophysics*, *34*(3), 209-213.

Zimmermann, T. D., & Meier, B. (2006). The rise and decline of prospective memory performance across the lifespan. *The Quarterly Journal of Experimental Psychology*, *59*(12), 2040-2046.

Zirnsak, M., & Moore, T. (2014). Saccades and shifting receptive fields: anticipating consequences or selecting targets? *Trends in Cognitive Sciences*, *18*(12), 621-628.

Zirnsak, M., Steinmetz, N. A., Noudoost, B., Xu, K. Z., Moore, T. (2014). Visual space is compressed in prefrontal cortex before eye movements. *Nature*, *507*(7493), 504.

# Appendix

# Appendix 1

Lauffs, M. M., Öğmen, H., & Herzog, M. H. (2017).
Uncertainty does not hamper nonretinotopic motion perception.
*Journal of Vision, 19*(7):6, 1-10.

Marc M. Lauffs[1], Haluk Öğmen [2], Michael H. Herzog[1]

[1] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] Department of Electrical and Computer Engineering, University of Denver, CO, USA

Contributions: ML and MH designed the experiment. ML programmed the stimulus, collected the data, and analyzed the data. ML, HÖ, and MH interpreted the data and wrote the manuscript.

# Local versus global and retinotopic versus non-retinotopic motion processing in schizophrenia patients

Marc M. Lauffs[a,*], Albulena Shaqiri[a], Andreas Brand[b], Maya Roinishvili[c,d], Eka Chkonia[c,e], Haluk Öğmen[f], Michael H. Herzog[a]

[a] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[b] Institute for Psychology and Cognition Research, University Bremen, Bremen, Germany
[c] Institute of Cognitive Neurosciences, Agricultural University of Georgia, Tbilisi, Georgia
[d] Vision Research Laboratory, Beritashvili Centre of Experimental Biomedicine, Tbilisi, Georgia
[e] Department of Psychiatry, Tbilisi State Medical University, Tbilisi, Georgia
[f] Department of Electrical & Computer Engineering, University of Denver, Denver, CO, USA

## A R T I C L E   I N F O

## A B S T R A C T

Schizophrenia impairs cognitive functions as much as perception. For example, patients perceive global motion in random dot kinematograms less strongly, because, as it is argued, the integration of the dots into a single Gestalt is complex and therefore deteriorated. Similarly, the perception of apparent motion is impaired, because filling-in of the illusory trajectory requires complex processing. Here, we investigated very complex motion processing using the Ternus-Pikler display. First, we tested whether the *perception* of global apparent motion is impaired in schizophrenia patients compared to healthy controls. The task requires both the grouping of multiple elements into a coherent Gestalt and the filling-in of its illusory motion trajectory. Second, we tested the perception of rotation in the same stimulus, which *in addition* requires the computation of non-retinotopic motion. Contrary to earlier studies, patients were not impaired in either task and even tended to perform *better* than controls. The results suggest that complex visual processing itself is not impaired in schizophrenia patients.

## 1. Introduction

Schizophrenia impairs perception to the same extent as cognition. While cognitive deficits are well investigated, less is known about visual deficits, even though they are amongst the most sensitive endophenotypes of schizophrenia (Chkonia et al., 2010).

Interestingly, simple local motion processing is not impaired in schizophrenia, whereas global and complex motion processing are disturbed. For example, discriminating the motion direction of simple drifting gratings is equally good in both patients and controls (Chen et al., 2003; Tibber et al., 2015). However, discriminating the coherent motion direction of a small subset of dots within a cloud of randomly moving dots, which requires both motion integration and segregation across dots, is deteriorated by a factor of about two (Chen et al., 2003; for a review, see Chen, 2011; Butler et al., 2008).

A special case of motion perception is apparent motion, where two elements are flashed at different locations separated by an inter-stimulus interval (ISI). Observers perceive motion from one element to the other, rather than two consecutively presented elements

(Wertheimer, 1912). Apparent motion processing has been suggested to be impaired in schizophrenia patients. However, experimental results are mixed (Saucer and Deabler, 1956; Saucer, 1958, 1959; Chambers and Wilson, 1986; Sanders et al., 2013; but see Sanders et al., 2012; Saucer, 1958; Tschacher et al.,2008). A deficit might be expected, because apparent motion needs complex processing, such as filling in the illusory motion trajectory between the elements.

Here, we investigated even more complex motion processing in schizophrenia by using the Ternus-Pikler display (TPD), combining apparent motion processing and non-retinotopic interpretation. Non-retinotopic processing is a fundamental aspect of visual perception. For example, if visual information would be processed strictly retinotopically, eye movements could not be differentiated from real motion in the world, which would arguably make it impossible to create a stable image of the world. In the TPD, three disks are briefly presented on a computer screen (Fig. 1). Following an ISI of variable duration, the disks reappear shifted by one position to the right. After another ISI, the disks reappear in their original position and the sequence starts over again (see animation on http://lpsy.epfl.ch/). The ISI determines

---

* Correspondence to: Laboratory of Psychophysics, Brain Mind Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Station 19, CH-1015 Lausanne, Switzerland.
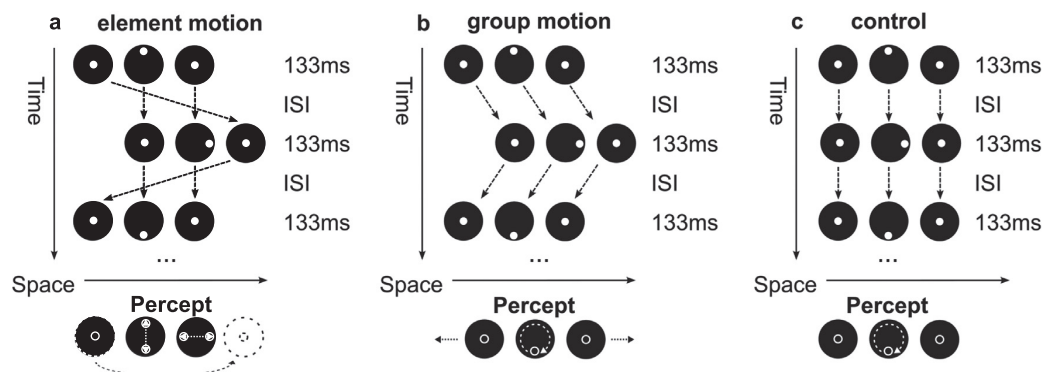E-mail address: marc.lauffs@epfl.ch (M.M. Lauffs).

**Fig. 1.** Ternus-Pikler display. Three black disks with white dots are presented (in Cond. 1, the dots were not presented). After an ISI, the disks reappear shifted to the right by one position, so that the leftmost disk in frame 2 spatially overlaps with the middle disk in frame 1. Following another ISI the sequence starts over again. a) For short ISIs, the two central disks are perceived as stationary and the dots appear to move linearly up-down (left disk) and left-right (right disk). The outermost disk appears to "jump" back and forth. The arrows indicate the perceived object correspondence (arrows were not shown in the experiment). b) For long ISIs, the three disks are perceived to move left-right in tandem. In the middle disk, a dot appears to rotate. The dot rotation is a non-retinotopic combination of the linear up-down and left-right motions, which are themselves not visible. c) As a control condition, we presented the three disks without lateral displacement and with a long ISI of 200 ms. The dot in the middle disk rotates either clockwise or counterclockwise. This rotation is retinotopic. © Marc M. Lauffs.

the global motion percept: If the ISI is long (e.g., 100 ms), the three disks form a group that is perceived to move back and forth horizontally (*group motion*; Fig. 1b). If the ISI is short (e.g., 0 ms), two stationary disks are perceived, with the outermost disk appearing to jump from left to right and back (*element motion*; Fig. 1a).

First, we parametrically varied the ISI and participants reported whether they perceived element or group motion. Next, we added a dot to each disk. The motion of the dot is determined by the ISI: When element motion is perceived, the dots appear to move linearly up-down in the left central disk and left-right in the right central disk (Fig. 1a). In the "jumping" outermost disk, the dot is always perceived in the center. The perceived motion is *retinotopic*, i.e., the percept is in accordance with the stimulation on the retina. In the group motion condition, the dots in the two outer disks appear stationary, whereas the dot in the middle disk appears to rotate either clockwise or counter-clockwise (Fig. 1b). The rotation percept is *non-retinotopic*, because there is no rotation on the retina. The rotation is a combination of the retinotopic left-right and up-down motions (see arrows in Fig. 1), which are invisible themselves.

Global motion processing *precedes* local motion processing in the Ternus-Pikler display because the global disk motion must first be computed before the local dot motion can be computed relative to it (Clarke et al., 2013; but see Pooresmaeili et al., 2012). Hence, patients should be strongly deteriorated with the Ternus-Pikler display, if global motion processing is deficient.

## 2. Methods

### 2.1. Participants

Twenty-two adults diagnosed with schizophrenia and 20 healthy controls matched in age and education participated in the experiment. All participants had normal or corrected-to-normal vision, with a visual acuity of ≥0.8 (corresponding to 20/25) at least in one eye, as determined with the Freiburg Visual Acuity Test (Bach, 1996).

Schizophrenia patients were recruited from the Tbilisi Mental Health Hospital. Among the patients, 8 were inpatients and 14 outpatients. Healthy controls were recruited from the general population. General exclusion criteria were drug or alcohol abuse, neurological or other somatic illnesses. Participants were no older than 53 years.

Ethics approval was obtained in Tbilisi from the Georgian National Council on Bioethics. All participants signed informed consent and were informed that they could quit the experiments at any time.

Patients were diagnosed according to DSM-IV by means of an interview based on the SCID, information of the staff, and the study of the records. Psychopathology of schizophrenia patients was assessed by an experienced psychiatrist (EC) by Scales for the Assessment of Negative Symptoms and Scales for the Assessment of Positive Symptoms (SANS, SAPS; Andreasen, 1983, 1984). All patients were receiving neuroleptic medication. Chlorpromazine equivalents and group characteristics are depicted in Table 1. Severity of positive and negative symptoms and CPZ dose were comparable to earlier studies with subjects from the same hospital (e.g., Plomp et al., 2013) and a hospital in Bremen, Germany (Herzog et al., 2004; Grimsen et al., 2013). In- and outpatients did not fundamentally differ in terms of age, illness duration, education, negative symptoms (SANS) and medication dose. Positive symptoms (SAPS) were higher in the inpatient group ($M$(SD) =11.0(3.6) vs. 8.1(3.0); $t$(20) =2.05, $p$=0.053), who on average also received higher medication doses (chlorpromazine $M$(SD) =697(429) vs. 556(366); $t$(19) =0.8, $p$=0.434). The medication data for one patient were not available.

### 2.2. Stimuli and task

Participants were tested in a dimly lit room. Stimuli were presented at a distance of 0.6 m on a Dell Latitude E5540 laptop (1920*1080 px, 60 Hz) and were programmed with Matlab and Psychtoolbox (Brainard, 1997; Pelli, 1997).

We used a variant of the Ternus-Pikler display (Fig. 1; Ternus, 1926; Pikler, 1917). In each stimulus frame, two horizontally aligned, equidistant black disks were presented for 133 ms, followed by a blank-screen ISI of variable duration. Participants were instructed to fixate on a fixation point presented between and below the two disks. A third disk was displayed either to the left or the right, alternating position

**Table 1**
Demographic data (mean ± S.D.) of schizophrenia patients and healthy controls.

|  | Schizophrenia patients | Healthy controls |
| --- | --- | --- |
| *N* | 22 | 20 |
| **Age (years)** | 39.5 ± 9.3 | 40.4 ± 7.6 |
| **Gender (f/m)** | 5/17 | 6/14 |
| **Education level (years)** | 13.3 ± 2.4 | 14.85 ± 2.4 |
| **Duration of illness (years)** | 15 ± 9.1 | |
| **SANS** | 10.23 ± 5.5 | |
| **SAPS** | 9.4 ± 3.45 | |
| **CPZ** | 609.82 ± 387.1 | |
| **Handeness (R/L)** | 22/0 | 20/0 |

with every frame. The start side of the third disk was chosen randomly with equal probability. Depending on the ISI, there were two percepts: At short ISIs, the two central disks are perceived as stationary at one location, while the third disk is "jumping" from left to right and vice versa (element motion, EM, Fig. 1a). At long ISIs, the three disks appear to move in tandem from left to right and back (group motion, GM, Fig. 1b).

Each trial comprised seven stimulus frames, interleaved by ISIs, and followed by a blank screen of 0.75 s before the next trial started. In Condition 1, the ISI was chosen randomly with equal probability from 0, 17, 33, 50, 67, 200 ms for each trial. In the case of an ISI of 0 ms, the stimulus frames followed each other without intermittent blank screen and the stimulus onset asynchrony was equal to the duration of one stimulus frame (133 ms). Ten trials for each ISI were presented randomly (60 trials in total). The task was to indicate, via push-button presses, whether element motion or group motion prevailed.

In Condition 2, a dot was added to each disk (Fig. 1). Only four frames with dots were presented, preceded by two and followed by one frame without dots. We used only ISIs of 0 and 200 ms, which elicit strong element and group motion percepts, respectively. For each ISI, 24 trials were presented in random order. Trials with clockwise and counterclockwise rotation were presented in random order with equal probability. Participants indicated, via push-button presses, whether they perceived clockwise or counterclockwise rotation. In condition 2, we randomly interleaved 12 control trials in which the three disks were presented in the same position in each frame, resulting in a retinotopic rotation of the middle disk (Fig. 1c). For these trials, an ISI of 200 ms was used.

### 2.3. Stimulus specifics

The disks had a diameter of 1.6° and were presented 0.8° above a central fixation point (red square with 0.07° side lengths). The center-to-center inter-disk distance was 1.9°. The white dots had a diameter of 0.25° and were presented halfway between the disk's center and rim. Background color of the screen was midlevel gray. The start orientation of the rotating dot was chosen randomly to be at 0, 90, 180 or 270 degrees. When a participant failed to respond within 3 s, a beep indicated the omission and the trial was repeated at a random later moment. No other feedback was given.

### 2.4. Statistical analysis

All statistical tests were performed in the free and open-source JASP software (Love et al., 2015). Mixed ANOVAs were Greenhouse-Geisser corrected where appropriate. T-tests for unequal variances (Welch's test) were used where appropriate. Uncorrected tests led to the same conclusions in all cases.

## 3. Results

### 3.1. Condition 1: Group vs. element motion

In Condition 1, the proportion of trials in which group motion was reported increased monotonically with ISI, in both patients and healthy controls (Fig. 2). Overall, neither group reported significantly more group motion than the other ($F(1,40)=0.710$, $p=0.404$, $\eta^2=0.017$, 95% CI for $M_{control} - M_{patient} =[-13.5\%; 5.6\%]$). There was also no significant interaction between *group* and *ISI* in a mixed ANOVA design ($F(5, 200)=1.367$, $p=0.256$, $\eta^2=0.008$, $\varepsilon=0.601$). This shows that the increase in group motion percepts in patients was not significantly different from that in controls. Even without correcting for multiple comparisons, *t*-tests per ISI were all statistically non-significant.

As can be seen in Fig. 2, patients reported slightly more group motion percepts than controls for the trials with ISIs between 0 and
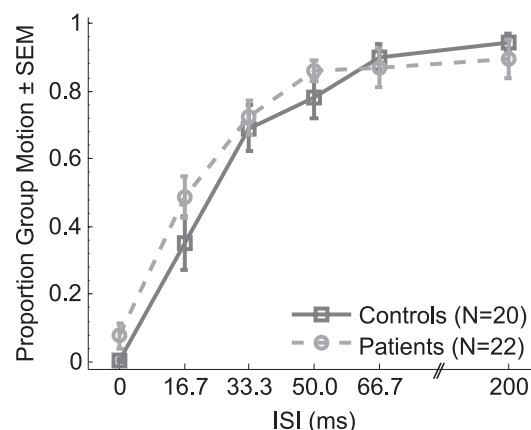


**Fig. 2.** Results Condition 1. Three disks (without the white dots) were shown and we varied the ISI. The proportion of group motion percepts increased with ISI for both groups. We did not find any significant differences between patients and controls..

50 ms and slightly less at the longest ISI of 200 ms. We suggest that these differences are mainly explained by patients' tendency to be more variable in their responses. For example, at the shortest ISI, control subjects reported group motion in only 0.5% (SD=2.24) of the trials. Since it can hardly be lower, the higher variance in patients (SD=17.44) necessarily leads to a higher mean. The inverse holds for the longest ISI, where controls reported group motion perception in 94.5% of the trials, and more variable responding can only lead to a lower mean. Supporting our suggestion, Levene's test for equality of variances was highly significant for ISIs 0 and 50 ms ($p \le 0.001$) and close to significance for ISI 200 ms ($p=0.055$).

### 3.2. Condition 2: Rotation discrimination

In Condition 2, white dots were added to the disks and participants indicated whether the central dot rotated clockwise or counterclockwise. We computed $d'$, a bias-free measure of sensitivity (Fig. 3a; Abdi, 2007; Macmillan and Creelman, 2005) but we also show results as the percentage of correct responses (Fig. 3b).

Patients' and controls' sensitivity did not differ significantly overall, nor was there an interaction between group and ISI ($p=0.616$ and $p=0.683$, respectively). For the 0 ms ISI, element motion was perceived and the rotation was invisible. Clockwise/counterclockwise discriminations were at chance level. Patients had a non-significantly *higher* sensitivity than healthy controls ($t(40)=-2.02$, $p=0.051$, $M_{control} - M_{patient} =-0.30 \pm 0.15$, 95% CI [-0.59, 0.00], Cohen's d=0.6). For trials with 200 ms ISI and the control trials with retinotopic rotation, sensitivity did not differ significantly between the two groups ($p=0.962$ and $p=0.990$, respectively).

Likewise, the percentage of correct responses did not differ significantly between the groups and there was no interaction between group and ISI ($p=0.652$ and $p=0.858$, respectively). All planned comparisons were non-significant (all $ps > 0.24$).

### 3.3. Correlations

Overall, we did not find any coherent relation between performance and psychopathology. We correlated the positive and negative symptoms scores (SAPS/SANS) with group motion perception per ISI using Pearson's *r*. None of the correlations was significant (Cond. 1: all $ps > 0.567$ for positive symptoms, all $ps > 0.197$ for negative symptoms; Cond. 2: all $ps > 0.540$ for positive, all $ps > 0.202$ for negative symptoms). Illness duration did not correlate with the subjective ratings in Condition 1 (all $ps > 0.371$). In Condition 2, illness duration only correlated with sensitivity for the ISI of 0 ms ($r =-0.489$, $p=0.021$; all
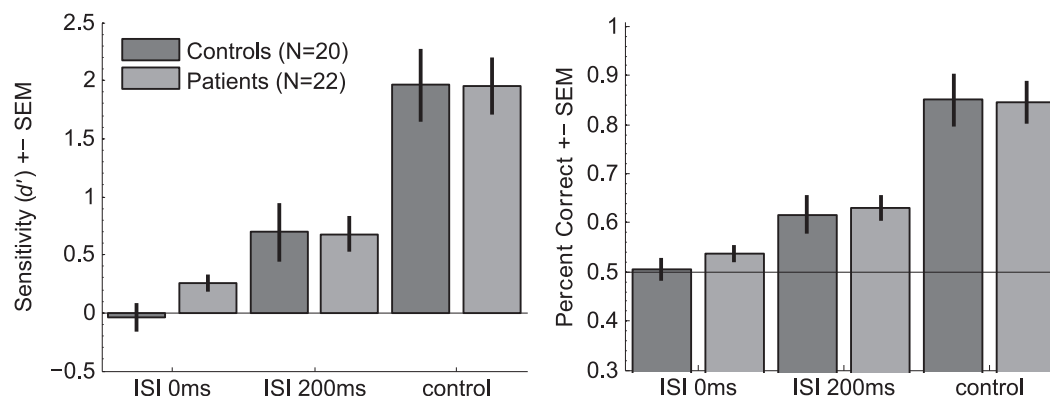
**Fig. 3.** Results Condition 2. Participants discriminated the rotation direction of the dot in the middle disk. Results are shown in terms of sensitivity d′ (left) and percent correct responses (right). Performance is at chance level for an ISI of 0 ms, because no group motion is perceived and the dots appear to move linearly up-down or left-right. Performance improves for an ISI of 200 ms and in the control condition.

other $p$s > 0.690). Chlorpromazine dose tended to correlate with group motion perception in Condition 1 for the ISI of 0 ms ($r$ =0.416, $p$=0.06), but not for the other ISIs (all $p$s > 0.133). In Condition 2, chlorpromazine dose only correlated with sensitivity for the 200 ms ISI ($r$ =−0.54, $p$=0.011; all other $p$s > 0.449).

## 4. Discussion

Non-retinotopic processing is critical in most real world situations. For example, a reflector appears to move on an orbital trajectory on the wheel of a moving bicycle. However, the "true" trajectory is very different, namely, cycloidal (Fig. 4). We cannot perceive the true motion trajectory because we subtract the horizontal motion of the bike from the motion trajectory of the reflector. The bicycle motion serves as a reference for the reflector motion (Duncker, 1929; Johansson, 1974). Motion processing in the Ternus-Pikler display follows exactly the same principles as in the bike example. Clearly, non-retinotopic motion processing is as important as it is complex. For this reason, we expected schizophrenia patients to be highly disturbed in the Ternus-Pikler display, because it involves (a) global motion processing, (b) the computation of apparent motion trajectories, and (c) the computation of *non-retinotopic* motion. However, we found

patients to perform almost identical to controls.

First, we tested the *subjective* perception of the patients by changing the ISI. In line with previous studies, which found diminished apparent motion perception in patients (Sanders et al., 2013; Chambers and Wilson, 1968; Saucer and Deabler, 1956; Saucer, 1958), we expected that the change from element to group motion occurs at a much longer ISI in the patients, because the local, retinotopic motion dominates the processing of the global, non-retinotopic percept. However, if anything, patients in our study perceived group motion for *shorter* ISIs than controls. Hence, there is no evidence for a global motion deficit on the phenomenal level.

Next, we tested accuracy. We added dots to the disks and used two ISIs, eliciting either strong element motion or strong group motion percepts. In the case of element motion, we expected that all participants would be unable to see the non-retinotopic and illusory dot rotation because the retinotopic up-down, left-right dot motion percepts dominate. This was indeed the case. In the group motion condition, we expected patients to be much worse than controls, because complex global motion processing is required to establish group motion, which in turn is needed to compute the non-retinotopic dot rotation. However, performance in patients and controls was very similar. This is also true for the control condition, where perceiving the
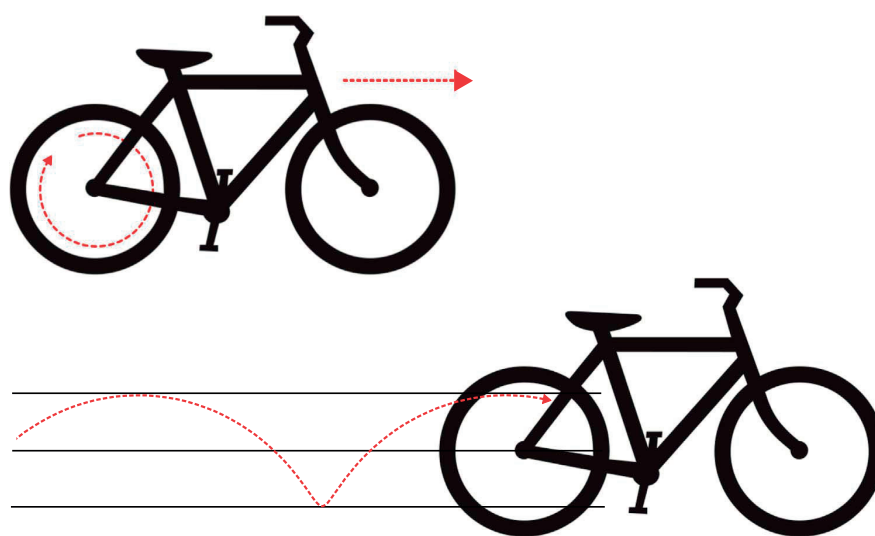


**Fig. 4.** The parts of a moving object are not perceived in retinal coordinates, but relative to the object. For example, a reflector on the wheel of a moving bicycle is perceived as circling, although its trajectory is cycloidal in retinal coordinates. It is perceived to circle, because the linear bicycle motion is subtracted from the cycloid motion. © Marc M. Lauffs.

dot rotation did not require non-retinotopic integration.

We like to mention that with an average $d'$ of around 0.6 (62–63% correct responses), dot rotation discrimination performance was much lower than in previous studies with healthy students (Boi et al., 2009). Experimental procedure, stimulus size, timing and other stimulus parameters were comparable across studies. Potentially, the difference is explained by the fact that the participants of the current study were older and had a more diverse educational background.

As with the Ternus-Pikler display, studies with other apparent motion paradigms did not find differences between patients and controls (Tschacher et al., 2008; Sanders et al., 2012; Saucer, 1958). Sanders and colleagues (2013) did not find a significant difference between the mean data of patients and controls. However, when they determined the alternation rate which yielded strongest apparent motion for each participant, significant differences between patients and controls were found. This analysis is not possible with the Ternus-Pikler display, because strongest group-motion percepts occur at the longest ISI for all observers alike.

We can only speculate about why the schizophrenia patients are not impaired in the Ternus-Pikler display. Potentially the Ternus-Pikler display is less attention demanding than other paradigms, such as the random dot kinematogram, where very close attention is needed. On the contrary, both group and dot motion are easily detected in the Ternus-Pikler display.

Whatever the exact mechanisms are, we can conclude that complexity of visual processing itself is not disturbed in the patients because the TPD requires the solution of many complex and interacting tasks.

### Conflicts of interest

All authors declare that this work was done free of conflicts of interest.

### Funding

### Contributors

MML, AS, AB, MR, EC, HO, and MMH initiated and designed the experiment and wrote the paper. MR and EC collected the data. MML programmed the experiment and analyzed the data.

All authors contributed to and have approved of the final manuscript.

### References

Abdi, H., 2007. Signal detection theory. In: Salkind, N. (Ed.), Encyclopedia of Measurement and Statistics. Sage, Thousand Oaks, CA.

Andreasen, N.C., 1983. Scale for the Assessment of Negative Symptoms (SANS). University of Iowa, Iowa City.

Andreasen, N.C., 1984. The Scale for the Assessment of Positive Symptoms (SAPS). University of Iowa, Iowa City.

Bach, M., 1996. The Freiburg Visual Acuity Test-automatic measurement of visual acuity. Optom. Vis. Sci. 73 (1), 49–53.

Boi, M., Öğmen, H., Krummenacher, J., Otto, T.U., Herzog, M.H., 2009. A (fascinating) litmus test for human retino-vs. non-retinotopic processing. J. Vis. 9 (13), (5-5).

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Butler, P.D., Silverstein, S.M., Dakin, S.C., 2008. Visual perception and its impairment in schizophrenia. Biol. Psychiatry 64 (1), 40–47.

Chambers, J.L., Wilson, W.T., 1968. Perception of apparent motion and degree of mental pathology. Percept. Mot. Skills 26 (3), 855–861.

Chen, Y., Nakayama, K., Levy, D., Matthysse, S., Holzman, P., 2003. Processing of global, but not local, motion direction is deficient in schizophrenia. Schizophr. Res. 61 (2), 215–227.

Chen, Y., 2011. Abnormal visual motion processing in schizophrenia: a review of research progress. Schizophr. Bull. 37 (4), 709–715.

Chkonia, E., Roinishvili, M., Makhatadze, N., Tsverava, L., Stroux, A., Neumann, K., Herzog, M.H., Brand, A., 2010. The shine-through masking paradigm is a potential endophenotype of schizophrenia. *PLoS ONE* 5 (12), e14268.

Clarke, A.M., Repnow, M., Öğmen, H., Herzog, M.H., 2013. Does spatio-temporal filtering account for nonretinotopic motion perception? Comment on Pooresmaeili, Cicchini, Morrone, and Burr (2012). J. Vis. 13 (10), 1–14.

Duncker, K., 1929. Uber induzierte Bewegung [On induced motion]. Psychol. Forsch. 12, 180–259.

Grimsen, C., Brand, A., Fahle, M., 2013. No evidence for prolonged visible persistence in patients with schizophrenia. PloS One 8 (3), e58940.

Herzog, M.H., Kopmann, S., Brand, A., 2004. Intact figure-ground segmentation in schizophrenia. Psychiatry Res. 129 (1), p55–p63.

Johansson, G., 1974. Vector analysis in visual perception of rolling motion. Psychol. Forsch. 36 (4), 311–319.

Love, J., Selker, R., Marsman, M., Jamil, T., Verhagen, A.J., Ly, A., Wagenmakers, E.J., 2015. JASP (Version 0.6. 5). Computer software

Macmillan, N.A., Creelman, C.D., 2005. Detection Theory: A User's Guide. Lawrence Erlbaum Associates, New York.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10 (4), 437–442.

Pikler, J., 1917. Sinnesphysiologische Untersuchungen [Physiological investigations of the senses]. J.A. Barth, Leipzig.

Plomp, G., Roinishvili, M., Chkonia, E., Kapanadze, G., Kereselidze, M., Brand, A., Herzog, M.H., 2013. Electrophysiological Evidence for Ventral Stream Deficits in Schizophrenia Patients. Schizophr. Bull. 39 (3), p547–p554.

Pooresmaeili, A., Cicchini, G.M., Morrone, M.C., Burr, D., 2012. "Non-retinotopic processing" in Ternus motion displays modeled by spatiotemporal filters. J. Vis. 12 (1), 1–15.

Sanders, L.L.O., Muckli, L., de Millas, W., Lautenschlager, M., Heinz, A., Kathmann, N., Sterzer, P., 2012. Detection of visual events along the apparent motion trace in patients with paranoid schizophrenia. Psychiatry Res. 198 (2), 216–223.

Sanders, L.L.O., de Millas, W., Heinz, A., Kathmann, N., Sterzer, P., 2013. Apparent motion perception in patients with paranoid schizophrenia. Eur. Arch. Psychiatry Clin. Neurosci. 263 (3), 233–239.

Saucer, R.T., Deabler, H.L., 1956. Perception of apparent motion in organics and schizophrenics. J. Consult. Psychol. 20 (5), 385.

Saucer, R.T., 1958. A further study of the perception of apparent motion by schizophrenics. J. Consult. Psychol. 22 (4), 256.

Saucer, R.T., 1959. Chlorpromazine and apparent motion perception by schizophrenics. J. Consult. Psychol. 23 (2), 134.

Ternus, J., 1926. Experimentelle Untersuchungen über phänomenale Identität [Experimental investigations of phenomenal identity]. Psychol. Forsch. 7, 81–136.

Tibber, M.S., Anderson, E.J., Bobin, T., Carlin, P., Shergill, S.S., Dakin, S.C., 2015. Local and global limits on visual processing in schizophrenia. PloS One 10 (2), e0117951.

Tschacher, W., Dubouloz, P., Meier, R., Junghan, U., 2008. Altered perception of apparent motion in schizophrenia spectrum disorder. Psychiatry Res. 159 (3), 290–299.

Wertheimer, M., 1912. Experimentelle Studien über das Sehen von Bewegung [Experimental studies on the seeing of movement]. JA Barth.

# Appendix 2

Lauffs, M. M., Choung, O. H., Öğmen, H., & Herzog, M. H. (*submitted*). Unconscious retinotopic motion processing affects nonretinotopic motion perception.

Marc M. Lauffs[1], Oh-hyeon Choung[1], Haluk Öğmen [2], Michael H. Herzog[1]

[1] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] Department of Electrical and Computer Engineering, University of Denver, CO, USA

# Unconscious retinotopic motion processing affects non-retinotopic motion perception

Marc M. Lauffs[a*], Oh-hyeon Choung[a], Haluk Öğmen[b], & Michael H. Herzog[a]

[a] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, http://lpsy.epfl.ch, marc.lauffs@epfl.ch, oh-hyeon.choung@epfl.ch, michael.herzog@epfl.ch

[b] Department of Electrical and Computer Engineering, University of Denver, CO, USA, haluk.ogmen@du.edu

*Corresponding author


**Postal adresses**

Marc M. Lauffs, Oh-hyeon Choung, & Michael H. Herzog

Ecole polytechnique fédérale de Lausanne (EPFL), Brain Mind Institute, Laboratory of Psychophysics

EPFL SV BMI LPSY, AI3102, Station 19, 1015 Lausanne, Switzerland

Phone 0041216937229

marc.lauffs@epfl.ch, oh-hyeon.choung@epfl.ch, michael.herzog@epfl.ch


Haluk Öğmen

Department of Electrical and Computer Engineering, University of Denver, CO, USA,

Daniel Felix Ritchie School of Engineering and Computer Science

University of Denver

2155 E. Wesley Ave., Room 481, Denver, CO 80208

Phone: (303) 871-2621

haluk.ogmen@du.edu

## Highlights

- Visual motion is usually perceived in non-retinotopic, object-centered coordinates.
- The *retinotopic* motion trajectory is invisible, but it is processed unconsciously.
- Invisible retinotopic motion affects conscious non-retinotopic motion perception.
- The Ternus-Pikler display is a versatile new tool to study unconscious processing.

## Abstract

In masked priming, unconscious processing of a masked prime can speed up processing of a later stimulus, although it is invisible. Similarly, the invisible interpretation of a bistable figure can affect the processing of a later stimulus. In the Ternus-Pikler display, three disks with white dots move left-and-right in concert. The dot in the middle disk is perceived to rotate (non-retinotopic percept). The retinotopic dot motion is invisible. Here, we asked the question whether unconscious processing of invisible retinotopic motion can affect non-retinotopic motion perception. Perception of the non-retinotopic rotation was severely degraded when an invisible retinotopic rotation was in the opposite direction. The effect cannot be explained by retinotopic models of vision, because all elements of the display are visible at all times, and form and distance of the elements can be flexibly adjusted. Only the interpretation changes, depending on non-retinotopic processes such as spatio-temporal object grouping.

148 / 150 characters incl. spaces

## Keywords

non-retinotopic processing; invisibility; ambiguous figures; consciousness

# Unconscious retinotopic motion processing affects non-retinotopic motion perception

## 1. Introduction

Conscious and unconscious perception are usually investigated with techniques such as backward masking (Bachmann & Francis, 2013; Breitmeyer & Öğmen, 2006), binocular rivalry (Wheatstone, 1838; Blake, 2001), or (continuous) flash suppression (Wolfe, 1984; Tsuchiya & Koch, 2004). Interestingly, even when an element is invisible, it can influence the processing of visible elements. For example, a masked, invisible prime can speed up the responses to an element presented later (Breitmeyer & Öğmen, 2006). Invisibility of the stimulus is usually explained by inhibition between neurons sensitive to the target and neurons sensitive to the mask or stimuli in the other eye. In line with the organization of the early visual areas in cortex, inhibition is either explicitly or implicitly assumed to be *retinotopically* organized (Wandell, Dumoulin, & Brewer, 2008; Engel, Glover, & Wandell, 1997). For example in binocular rivalry, neurons processing information originating from the left and right eye mutually suppress each other, by each inhibiting the neuron of the other eye coding for the same position in the visual field (Blake, 1989; however, see Leopold & Logothetis, 1996, Kovacs, Papathomas, Yang, & Fehér, 1996). In ambiguous figures, one of two rivaling interpretations is perceived. For example, in Rubin's vase, the silhouette of either a goblet or two faces is perceived, depending on which one is interpreted as figure and which one as ground. Unlike in masking and rivalry, all elements of the display are fully visible, but one *interpretation* is suppressed. As in masking, the suppressed interpretation of the image can influence stimuli presented later (e.g., Peterson & Kim, 2001; Peterson & Skow, 2008). Also for ambiguous figures, retinotopic inhibition may explain why stimuli or interpretations are invisible, for example, by mutually

inhibition boundary-ownership neurons in V2 (Zhou, Friedman, & von der Heydt, 2000; Zhaoping, 2005).

Perception is usually *non-retinotopic*. The parts of a moving object are perceived relative to the object, rather than in retinal coordinates (Öğmen & Herzog, 2010; Duncker, 1929; Johansson, 1950, 1974, 1976; Clarke, Öğmen, & Herzog, 2016; Ağaoğlu, Clarke, Herzog, & Öğmen, 2016). For example, the reflector on the wheel of a bicycle is perceived to rotate, although its motion is cycloidal in retinotopic coordinates. The bicycle motion is perceptually discounted from the cycloidal retinotopic motion of the reflector. Only the circular non-retinotopic reflector motion is perceived consciously, whereas the cycloidal retinotopic motion is invisible. Similar to ambiguous figures, all elements are visible, only one interpretation is suppressed (Herzog, Hermens, & Öğmen, 2014). Obviously, the percept cannot be explained by retinotopic inhibition because it depends on non-retinotopic information, namely whether the reflector is perceived as belonging to the bicycle or not. Almost nothing is known about the unconscious processing of invisible retinotopic motion and its influences on the consciously perceived non-retinotopic motion.

Here, we used an adapted version of the Ternus-Pikler display (Ternus, 1926; Pikler, 1917). In our version of the Ternus-Pikler display (Boi, Öğmen, Krummenacher, Otto, & Herzog, 2009), two black disks with white dots are briefly presented on a computer screen. After a brief inter-stimulus interval (ISI), the disks reappear in the same position and the sequence starts over. The white dots are displaced from frame to frame, creating an apparent motion percept of up-and-down motion in one, and left-and-right motion in the other disk (retinotopic percept; Figure 1a left and Movie 1). Then a third disk is added alternately to the left and right (Figure 1a right and Movie 2). The addition of the third disk changes the perceptual organization: The

three disks group and are perceived as moving left-and-right in concert. The dot in the middle

disk is perceived to rotate, while the white dots in the outer disks appear stationary (non-

retinotopic percept). The dot motion percept is fundamentally different although the retinal

image is unchanged, except for the addition of the third disk (compare left and right part of

Figure 1). Hence, invisibility of the retinotopic motion is not due to retinotopic information

being suppressed, but because the addition of the third disk changes its interpretation.

## a) Linear retinotopic dot-motion
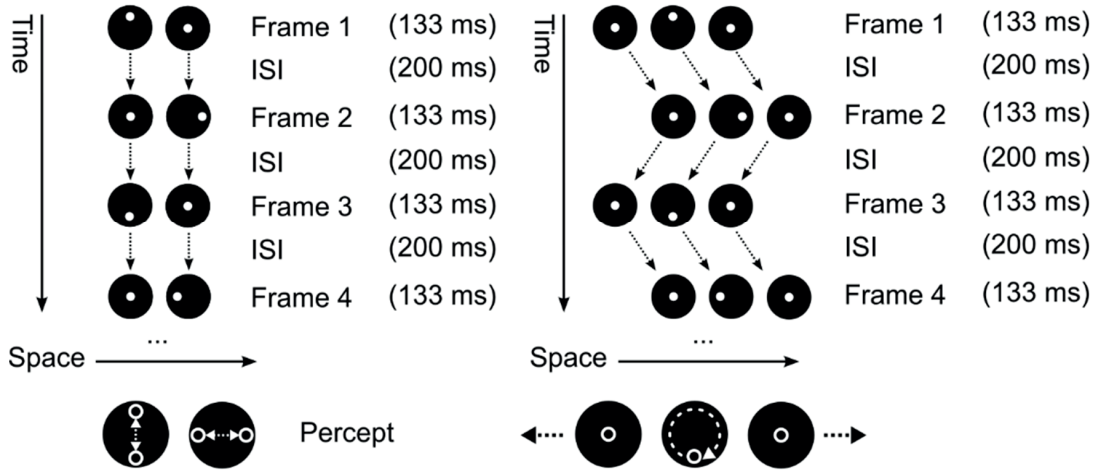


## b) Circular retinotopic dot-motion
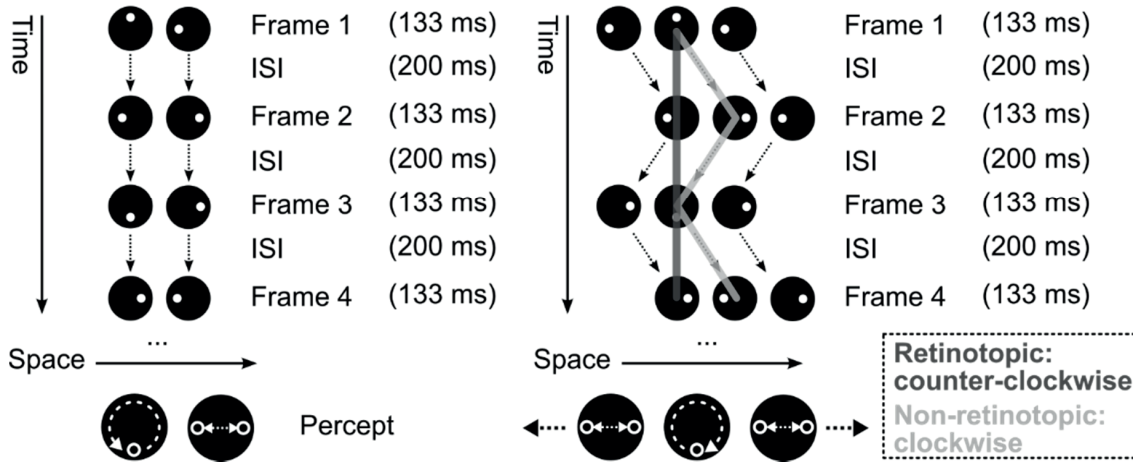


*Figure 1 a) Ternus-Pikler display with linear retinotopic and circular non-retinotopic dot motion (as introduced by Boi et al., 2009). Left: When two disks are presented, the dot in one disk is perceived to move up-and-down in one and left-and-right in the other disk (retinotopic percept). Right: When a third disk is added alternately to the left and right, the three disks are perceived to move left-and-right in concert ("group motion" percept for the disks) and the dot in the middle disk is perceived to rotate (non-retinotopic percept). b) Ternus-Pikler display with circular retinotopic and non-retinotopic dot motion. Left: When only two disks are presented, in the left disk the dot is perceived to rotate and in the right disk to jump left-and-right every second frame (retinotopic percept). Right: When a third disk is added alternately to the left and right, the dot in the middle disk is perceived to rotate and the dots in both outer disks jump left-and-right every second frame (non-retinotopic percept). Although the retinal image is identical apart from the third contextual disk, only the non-retinotopic rotation is perceived with three disks. With two disks, only the retinotopic rotation is perceived. Dark and light grey lines point out the retinotopic and non-retinotopic rotation, respectively, and were not presented during the experiment. We show here an example with clockwise non-retinotopic and counter-clockwise retinotopic rotation, but the rotations could be in both the same and opposite directions. Dots that are not involved in either rotation (first and third in Frame 1, third in frame 2, etc.) can be placed arbitrarily. We chose the depicted dot positions because the symmetry of the outer disks enhances the group-motion percept. Broken arrows indicate the perceived object correspondence and motion direction of the disks and were not presented during the experiment.*

## 2 Experiment 1

In Experiment 1, we adapted the Ternus-Pikler display to contain both retinotopic and non-retinotopic rotations (Figure 1b and Movies 3-6). The retinotopic and non-retinotopic rotation could either be in the same direction (*congruent* rotation, e.g., both clockwise) or in opposite directions (*incongruent* rotation; e.g., retinotopic clockwise and non-retinotopic counter-clockwise). The white dots were positioned so that when only two disks are presented, retinotopic rotation was perceived in the left disk, and the dot in the right disk jumped left-and-right or up-and-down every second frame (Figure 1b left and Movie 3). When the third disk was added, non-retinotopic rotation was perceived in the middle disk, and the dots in the outer disks jumped left-and-right or up-and-down every second frame (Figure 1b right and Movie 4). The retinal image was identical in both cases, except for the added third disk. The observers were asked to report either the retinotopic or non-retinotopic rotation direction (clockwise/counter-clockwise).

To test the influence of the retinotopic rotation on the non-retinotopic percept we presented three disks and had observers report the non-retinotopic rotation direction. We then compared performance between trials in which the retinotopic rotation was in the same (Movie 5) versus the opposite direction of the non-retinotopic rotation (Movie 6). We hypothesized that incongruent retinotopic rotation might impair the non-retinotopic rotation percept, whereas congruent retinotopic rotation might facilitate it. For comparison, we used the classic Ternus-Pikler display with retinotopic up-down and left-right dot motion. As additional control condition, we presented a Ternus-Pikler display with a rotating dot in the middle disk, but randomly placed dots in the left and right disk (Movie 7). As a baseline

condition, we first presented a Ternus-Pikler display with only the rotating dot in the middle disk, but no dots in the flanking disks (Movie 8).

## *2.1 Methods*

### 2.1.1 Observers

Sixteen naïve observers took part in the experiment, but three observers were excluded from the analysis due to inferior performance in the baseline condition (60% correct responses or less). Three other observers were excluded because they were unable to maintain stable central fixation. Hence, 10 observers were available for analysis (Mean age 23.3 years, *SD* = 2.3 years, 5 female, 1 left-handed, 1 wore glasses). None of the observers had participated in earlier experiments with the Ternus-Pikler display and all were naïve to the purpose of the experiments. All observers had normal or corrected-to-normal visual acuity, as indicated by a binocular value ≥ 1.0 (corresponding to 20/20) in a program similar to the Freiburg visual acuity test (Bach, 1996). The experiments were approved by the local ethics committee and performed in accordance with the Declaration of Helsinki (World Medical Association, 2013). All observers gave written informed consent prior to the experiment. The observers were recruited from the EPFL student population and paid 20 CHF/h for their participation.

### 2.1.2 Apparatus

Stimuli were programmed in Matlab with Psychtoolbox (Brainard, 1997; Pelli, 1997) and presented on a 24.5 inch BenQ XL2540 LCD monitor (1920x1080 pixels, 60 Hz, http://display-corner.epfl.ch). Observers were positioned in the head rest of a SMI iViewX Hi-Speed 1250 eye tracker (Sensomotoric Instruments, Teltow, Germany) and viewed the stimuli from a

distance of 0.66 m. Eye tracking data were recorded binocularly at 500 Hz and immediately averaged over both eyes to reduce noise. The room was well lit to facilitate the eye tracking.

**2.1.3 Stimulus**

Depending on the condition, either two or three black disks with a white dot were presented in each frame for 133ms on a grey background. Each stimulus frame was followed by an ISI of 200ms. The black disks were 2.0° in diameter and separated horizontally by 0.5° (2.5° center-to-center). When three disks were presented, the stimulus shifted back and forth by one inter-stimulus distance (2.5°) per frame, so that two stimulus positions overlapped in all frames (cf. Figure 1). The white dots were 0.25° in diameter and positioned in the center of the disk or halfway between the disk's center and rim at the 3, 6, 9, or 12 o'clock position. The stimulus was presented 2 disk diameters (4.0°) above a central fixation point (red, r = 0.025°), whose horizontal position was midway between the two central disks. In each trial only four frames were presented, preceded by two and followed by one frame where the black disks were shown without the white dots. Trials were separated by an inter-trial interval of at least 1s, but started only after the observer fixated the fixation point.

In the conditions with linear retinotopic motion, the start position of the disks (left/right), the start position of the white dot in the middle disk (3, 6, 9, 12 o'clock), and its rotation direction (clockwise/counter-clockwise) were combined in a full-factorial design and presented in random order. In the conditions with circular retinotopic motion, the retinotopic and non-retinotopic rotation directions (clockwise/counter-clockwise) were combined in a full-factorial design and presented in random order. The retinotopic and non-retinotopic rotations could hence be in the same and in opposite directions with equal probability. Start

position of the disks (left/right) and start position of the dot in the middle disk (3, 6, 9, 12 o'clock) were each chosen randomly with equal probability.

Depending on the condition, the observers were asked to report the non-retinotopic rotation direction (clockwise vs. counter-clockwise) or the retinotopic motion direction (clockwise vs. counter-clockwise, or up-down vs. left-right) via handheld push-buttons. The observers were re-instructed before each block with a demonstration of the stimulus they were about to see. Each observer performed one block of 80 trials per condition. The conditions were presented in the following order:

### Baseline condition (Movie 8)

o Condition 1: Three black disks moved left-and-right in concert. Only the middle disk carried a rotating dot. The observers were asked to report the direction of the (non-retinotopic) rotation of the middle disk (clockwise vs. counter-clockwise).

### Ternus-Pikler display with retinotopic and non-retinotopic rotation (Movies 3-6)

o Condition 2: In Condition 2, three disks were presented and the observers were instructed to report the non-retinotopic rotation direction.

o Condition 3: Condition 3 was identical to Condition 2, but the observers were instructed to *ignore* the non-retinotopic rotation and report only the retinotopic rotation.

o Condition 4: The observers were again instructed to report the retinotopic rotation direction, but only two disks were presented.

### Ternus-Pikler display with retinotopic linear motion and non-retinotopic rotation (Movies 1-2)

- o Condition 5: In Condition 5, three disks were presented and the observers were instructed to report the non-retinotopic rotation direction.

- o Condition 6: Condition 6 was identical to Condition 5, but the observers were instructed to *ignore* the non-retinotopic rotation and report whether the retinotopic rotation *in the left disk* is going up-and-down or left-and-right.

- o Condition 7: The observers were again instructed to report the retinotopic rotation direction, but only two disks were presented.

***Control condition (Movie 7)***

- o Condition 8: As a final control condition, three disks with rotating dot in the middle disk were presented. The dot positions in the left and right disk were each chosen randomly from the 3, 6, 9, 12 o' clock positions and the center of the disk.

**2.1.4 Analysis**

We analyzed performance in terms of percent correct rotation direction discriminations. We performed pre-planned two-sided paired-samples *t* tests to compare conditions with each other, and two-sided one-sample *t* tests for comparisons with 50% chance-level. All tests were performed in the open-source JASP software (http://jasp-stats.org).

**2.1.5 Fixation control**

Fixation was automatically controlled after each trial. Fixation was considered broken if the gaze deviated more than 1.5° from the fixation point during a period longer than 20ms, and when the signal was lost during a period of 150ms or longer. Problematic trials were discarded and repeated in the same block at a random later moment. The observer received audio-visual feedback.

## 2.2 Results and discussion

In the baseline condition (Condition 1), observers discriminated the rotation direction very well (84.0%; Figure 2). In Conditions 2-4, we used the stimulus with retinotopic *and* non-retinotopic rotations (Figure 1b). Phenomenologically, only the non-retinotopic rotation was perceived and the retinotopic rotation was largely invisible. In Condition 2, observers reported the non-retinotopic dot rotation in the middle disk. Overall performance was significantly lower than in the baseline condition (C2$_{all}$: 73.3%; C2$_{all}$ vs. C1: $t(9) = 4.40$, $p = .002$, mean diff. = 10.7%). This difference was driven by trials in which the retinotopic rotation direction was opposite of the non-retinotopic rotation direction: Performance was strongly deteriorated (C2$_{opposite}$ vs. C1: $t(9) = 6.66$, $p < .001$, mean diff. = 24.7%; C2$_{opposite}$ vs. C2$_{same}$: $t(9) = 6.27$, $p < .001$, mean diff. = 28.0%) and barely exceeded the 50% chance-level (59.3%; C2$_{opposite}$ vs. 50%: $t(9) = 1.90$, $p = .090$). When the retinotopic rotation direction matched the non-retinotopic rotation direction, performance was equally good as in the baseline condition, but not significantly better (87.3%; C2$_{same}$ vs. C1: $t(9) = 1.14$, $p = .284$, mean diff. = 3.3%). Given the high baseline performance level of 84%, it cannot be ruled out that this is partly due to a ceiling effect. But if a congruent retinotopic rotation would benefit non-retinotopic perception as much as an incongruent rotation impairs it (-24.7% compared to baseline), we would have expected a larger performance increase than the 3.3% we observed here. Therefore, it is likely that congruent retinotopic rotation has little to no effect on the non-retinotopic motion percept.

In Condition 3, the observers were asked to report the retinotopic rotation direction. Performance was very low, albeit significantly above chance-level (C3$_{all}$: 57.5%; C3$_{all}$ vs. 50%: $t(9) = 3.84$, $p = .004$). In trials where the retinotopic rotation direction matched the visible

non-retinotopic rotation direction, performance was reasonably good (C3$_{same}$: 68.2%). However, when the retinotopic and non-retinotopic rotation direction did not match, performance was at chance-level (C3$_{opposite}$: 46.8%; C3$_{opposite}$ vs. 50%: $t(9) = 0.86$, $p < .410$). This pattern of results is compatible with the interpretation that the retinotopic rotation was phenomenologically invisible and that, given the absence of a clear retinotopic percept, the responses were biased by the non-retinotopic rotation. As we saw in Condition 2, the non-retinotopic rotation is clearly perceived when combined with a retinotopic rotation in the same sense, but not when combined with an incongruent retinotopic rotation. This might explain why the responses in Condition 3 were more biased by the (clearly visible) congruent non-retinotopic rotation than the (almost invisible) incongruent non-retinotopic rotation. This suggestion is difficult to test in the case of congruent rotation, where performance was well above chance-level, which could also indicate that the retinotopic rotation was visible. However, the rotation direction of the incongruent retinotopic rotation could *not* be reported with higher than chance accuracy, showing that it was effectively invisible. The strong effect of incongruent retinotopic rotation on the non-retinotopic percept in Condition 2 must hence have been due to *unconscious* processing of the retinotopic rotation.

In Condition 4, we presented only two disks. Now the retinotopic, but not the non-retinotopic rotation was perceived. Observers reported the retinotopic rotation direction. Performance was very good, irrespective of whether the now invisible non-retinotopic rotation direction was in the same direction or not (C4$_{all}$: 94.4%, C4$_{same}$: 95.8%, C4$_{opposite}$: 93.0%; C4$_{same}$ vs. C4$_{opposite}$: $t(9) = 2.18$, $p = .057$, mean diff. = 2.8%).

In Conditions 5-7, we presented the Ternus-Pikler display with linear retinotopic dot motions (Figure 1a). In Condition 5, we presented three disks and only the non-retinotopic dot rotation

in the middle disk was perceived. The retinotopic up-down dot motion in one, and left-right dot motion in the other disk were invisible. Observers were asked to report the non-retinotopic rotation direction and performed equally well as in the baseline condition without dots in the left and right disk (C5: 84.9%; C5 vs. C1: $t(9)$ =0.43, $p$ = .678, mean diff. = 0.9%) and as in the condition where the retinotopic rotation direction matched the non-retinotopic direction (C5 vs. C2$_{same}$: $t(9)$ =0.95, $p$ = .367, mean diff. = 2.4%). Hence, linear retinotopic motion did not affect the non-retinotopic motion percept. In Condition 6, we presented three disks and asked the observers to report the invisible retinotopic motion direction (up-down vs. left-right). Performance was at chance-level (C6: 54.7%; C6 vs. 50%: $t(9)$ = 0.98, $p$ = .352), confirming that the retinotopic motion was indeed invisible. In Condition 7, we presented only two disks, and the retinotopic rotation could be readily reported (C2: 92.9%).

Finally, in Condition 8, we presented a Ternus-Pikler display with a non-retinotopic rotation in the middle disk. The dot positions in the left and right disk were chosen at random, creating random motion percepts in these disks. Observers reported the non-retinotopic rotation direction. Performance was good (78.7%) and not significantly different from the baseline condition (C8 vs. C1: $t(9)$ =1.67, $p$ = .130, mean diff. = 5.3%), although the random dot motions in the left and right disks increased the complexity of the stimulus. However, performance in Condition 8 was slightly lower than when the retinotopic and non-retinotopic rotations were in the same direction (C8 vs. C2$_{same}$: $t(9)$ =2.77, $p$ = .022, mean diff. = 8.6%).
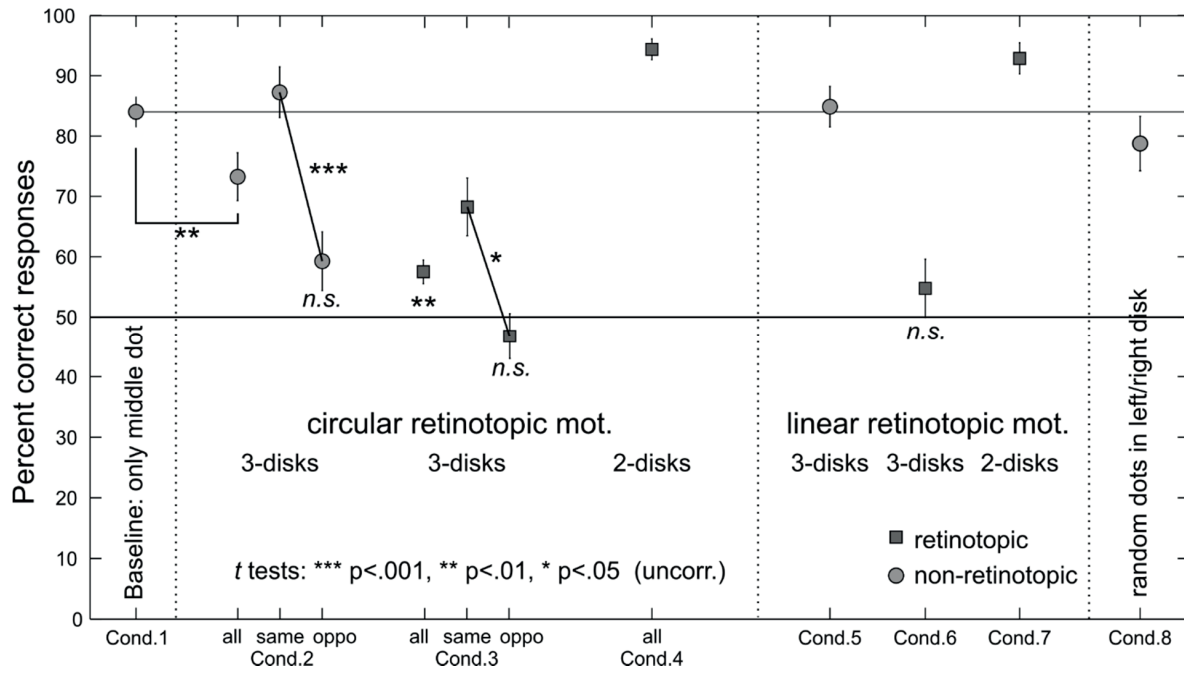
*Figure 2 Motion direction discrimination performance in Experiment 1. When three disks are presented, the non-retinotopic motion is perceived (C1, C2, C5, C8). The non-retinotopic motion percept is impaired when the stimulus with circular retinotopic motion is used (C2). The impairment occurs because performance is not significantly different from chance-level when the retinotopic rotation direction is incompatible with the non-retinotopic rotation direction (C2$_{opposite}$). This is highly surprising, because the retinotopic motion is largely invisible: When three disks are presented, the non-retinotopic motion percept dominates the percept (C3, C6). The retinotopic motion is visible when only two disks are presented (C4, C7). Error bars depict one standard error of the mean.*

## 3 Experiment 2

We performed a second experiment, in which we added a condition where a black frame enclosed the retinotopic rotation (Movie 9). Typically, the retinotopic dot motion is invisible because the dots are processed relative to the moving disks, rather than in retinal coordinates (Öğmen & Herzog, 2010; Clarke et al., 2016). We hypothesized that by providing the frame as a stationary reference, observers could prevent the dot from being processed relative to the moving disks, rendering the retinotopic dot motion visible. In addition, we included a condition with two disks in which we asked the observers to report the non-retinotopic rotation direction.

### 3.1 Methods

The methods for Experiment 2 are identical to Experiment 1 unless noted otherwise below.

Ten new, naïve observers participated (mean age 23.6 years, *SD* = 3.0 years, all male, 2 left-handed, 1 wore glasses, no exclusions). Stimuli were presented on a 24inch Asus VG24248QE LCD monitor (1920x1080 pixels, 60 Hz, http://display-corner.epfl.ch).

In Conditions 1-5 we used the Ternus-Pikler display with retinotopic *and* non-retinotopic rotations (Figure 1b and Movies 3-6). In Conditions 6-8 we used the Ternus-Pikler display with linear retinotopic motions (Figure 1a and Movies 1-2). Each observer performed one block of 80 trials per condition, except for Condition 3 (2 blocks of 80 trials) and Conditions 4 and 8 (each 1 block of 48 trials). The conditions were presented in the following order:

***Ternus-Pikler display with retinotopic and non-retinotopic rotation (Movies 3-6)***

- Condition 1: Three disks were presented and the observers were instructed to report only the non-retinotopic rotation direction.

- Condition 2. Three disks were presented and the observers were asked to report only the retinotopic rotation direction and ignore the non-retinotopic rotation. The retinotopically rotating disk was enclosed in a black frame (line width 0.1°, distance to disk 0.25°), that was presented throughout the entire trial including the ISIs.

- Condition 3: Condition 3 was identical to Condition 2, but the frame was omitted.

- Condition 4: The observers were again instructed to report the retinotopic rotation direction, but only two disks were presented.

- Condition 5: Only two disks were presented and the observers were instructed to report the direction of the *non-ret*inotopic rotation.

***Ternus-Pikler display with retinotopic linear motion and non-retinotopic rotation (Movies 1-2)***

- Condition 6: Three disks were presented and the observers were instructed to report the non-retinotopic rotation.

- Condition 7: Condition 7 was identical to Condition 6, but the observers were instructed to report whether the retinotopic motion *in the left disk* was going up-and-down or left-and-right.

- Condition 8: Condition 8 was identical to Condition 7, except that two instead of three disks were presented.

***3.2 Results and discussion***

We first presented the Ternus-Pikler display with three disks and invisible retinotopic rotation. In Condition 1, observers reported the non-retinotopic rotation direction. Overall performance was good (C1$_{all}$: 79.2%; Figure 3). As in the first experiment, performance was strongly impaired when the retinotopic rotation was in the opposite direction, compared to when it was in the same direction as the non-retinotopic rotation (C1$_{same}$: 89.5%, C1$_{opposite}$: 69.0%; C1$_{same}$ vs. C1$_{opposite}$: $t(9) = 6.49$, $p < .001$, mean diff. = 20.5%). Again performance levels were comparable when the retinotopic rotation direction was the same as for the non-retinotopic rotation and when a linear retinotopic motion was used (C6: 85.6%; C6 vs. C1$_{same}$: $t(9) = 1.62$, $p = .140$, mean diff. = 3.9%).

In Condition 2 and 3, we presented three disks and observers reported the retinotopic rotation direction. In Condition 2, the retinotopic rotation was enclosed in a black frame that remained on the screen during the ISIs. We expected that observers could anchor their perception to the frame and report the retinotopic motion direction. This was not the case.

Performance was low overall (C2$_{all}$: 62.8%), and hardly better than chance-level when the retinotopic did not match the non-retinotopic rotation direction (C2$_{opposite}$: 58.8%; C2$_{opposite}$ vs. 50%: $t(9) = 2.07$, $p = .069$). When the retinotopic and non-retinotopic rotation direction matched, performance was better (C2$_{same}$. 67.0%; C2$_{same}$ vs. C2$_{opposite}$: $t(9) = 2.72$, $p = .023$, mean diff. = 8.3%). We found the same pattern in Condition 3, where the frame around the retinotopic rotation was omitted (C3$_{same}$: 73.6%, C3$_{opposite}$: 41.4%; $t(9) = 5.12$, $p < .001$, mean diff. = 32.3%). Performance in the opposite rotation case was even significantly *below* chance-level ($t(9) = 2.85$, $p = .019$). As in Experiment 1, this pattern of results is compatible with the interpretation that the retinotopic rotation is invisible and responding is biased by the non-retinotopic rotation.

In Condition 4 and 5, only two disks were presented and the retinotopic, but not the non-retinotopic rotation was perceived. In Condition 4, observers reported the retinotopic rotation direction and performance was very good (C4$_{all}$: 92.9%), irrespective of whether the non-retinotopic rotation was in the same or the opposite direction (C4$_{same}$: 93.3%, C4$_{opposite}$: 92.5%; $t(9) = 0.36$, $p = .726$). In Condition 5, the observers were asked to report the (invisible) non-retinotopic rotation direction. Similar to the pattern observed in Condition 3, performance was reasonably good when the retinotopic and non-retinotopic rotation were in the same direction (C5$_{same}$: 74%) and below chance when not (C5$_{opposite}$: 34.5%), indicating that the non-retinotopic rotation was invisible and observers' responses were biased by the visible retinotopic rotation.

In Conditions 6-8, we used the Ternus-Pikler display with linear retinotopic motion. As in Experiment 1, non-retinotopic rotation discrimination performance was good and comparable to trials of Condition 1, where the retinotopic and non-retinotopic rotations were

in the same direction. The retinotopic motion direction (up-down vs. left-right) could *not* be

reported above chance-level when three disks were presented (C10: 55.7%; C10 vs. 50%: $t(9)$

= 1.31, *p* = .222). The retinotopic motion could easily be reported when only two disks were

presented (C11: 88.1%).
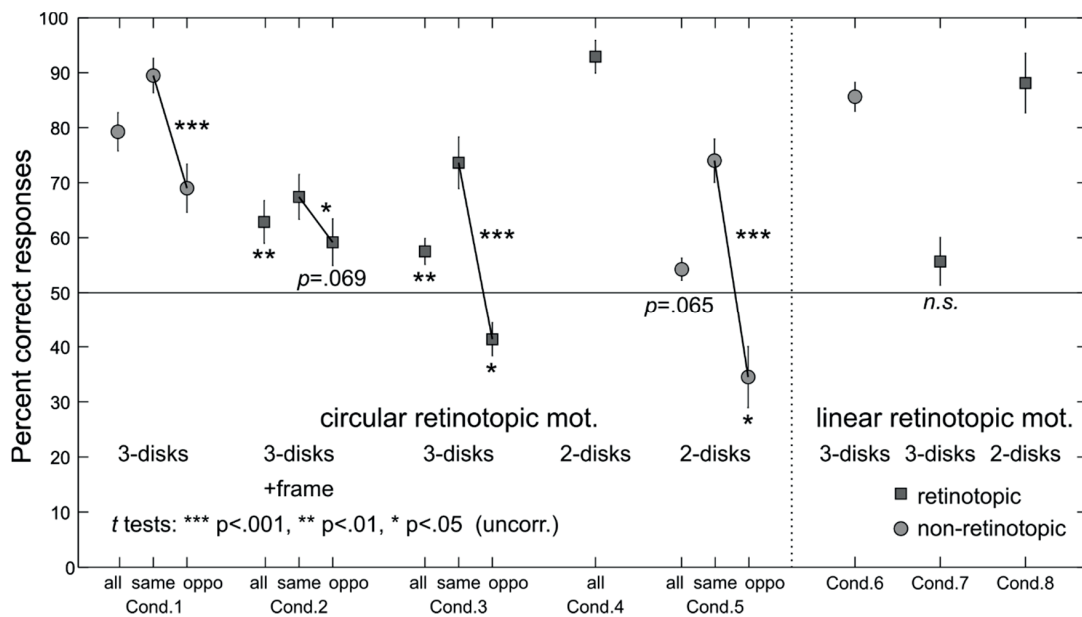


*Figure 3 Motion direction discrimination performance in Experiment 2. In line with the first experiment, when three disks are presented the non-retinotopic motion is perceived (C1, C6) and the retinotopic motion cannot reliably be reported (C3, C7). A black frame enclosing the retinotopic rotation improves performance only marginally (C2). When only two disks are presented, the retinotopic motion is perceived (C4, C8) and non-retinotopic motion cannot reliably be reported (C5). When the stimulus with retinotopic and non-retinotopic rotation is used, observers tend to answer in line with the visible rotation even if it is not task relevant, leading to above chance-level performance when the rotation directions are in the same direction (C3$_{same}$, C5$_{same}$) and below chance-level performance when not (C3$_{opposite}$, C5$_{opposite}$). Replicating the first experiment, the non-retinotopic motion direction can equally well be reported then the retinotopic rotation direction matches the non-retinotopic rotation direction (C1$_{same}$) and when the retinotopic motion is linear, rather than circular (C6). However, the non-retinotopic motion percept is strongly degraded if the retinotopic rotation is in the opposite direction. Error bars show one standard error of the mean.*

## 4 Experiment 3

In Experiment 1 and 2, we found that conscious non-retinotopic rotation perception is impaired by unconscious processing of a retinotopic rotation in the opposite direction. Linear retinotopic motion and a congruent retinotopic rotation did not affect the percept. In contrast, when we presented only two disks, retinotopic motion was perceived and was unaffected by the invisible non-retinotopic rotation.

In Experiment 3, we tested whether multiple retinotopic rotations interact. We added a condition in which there was not one, but two retinotopic rotations (Figure 4a and Movies 10-13). We were primarily interested if one congruent and one incongruent retinotopic rotation would cancel each other, and whether the detrimental effects of two incongruent retinotopic rotations would add up, leading to even worse performance. In addition, we investigated whether visible *non-ret*inotopic rotations affect the percept in a similar fashion as invisible retinotopic rotations. To this end, we added a condition in which we presented non-retinotopic rotation in each of the three disks, but no retinotopic rotations (Figure 4b and Movies 14-17).
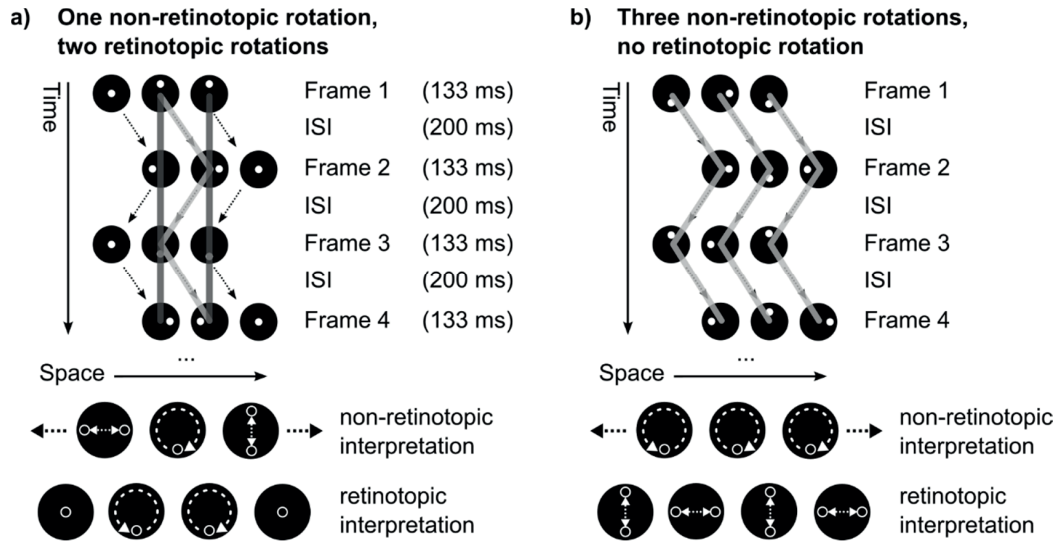
*Figure 4 a) In Experiment 3 we added a condition in which we presented two retinotopic rotations, instead of one as in Experiments 1 and 2. The two retinotopic rotations could be in the same or opposite direction, and each retinotopic rotation could be in the same or opposite direction of the non-retinotopic rotation in the middle disk. In the depicted example, the non-retinotopic rotation (light grey) and right retinotopic rotation (dark grey) are clockwise, and the left retinotopic rotation is counter-clockwise. Only the non-retinotopic interpretation is perceived: The dot in the middle disk rotates and the dots in the left and right disk move up-and-down and left-and-right, respectively. The retinotopic interpretation of the image was invisible and is shown only to illustrate the two retinotopic rotations. b) We also added a condition in which no retinotopic, but three non-retinotopic rotations were presented. The observers reported the rotation in the middle disk. The outer disks could rotate both in the same direction as the middle disk, both in the opposite direction, or one in the same and one in the opposite direction. Retinotopically, the dots moved left-right or up-down every second frame. In the depicted example, the dot rotates clockwise in the middle and right disk and counter-clockwise in the left disk. Again, only the non-retinotopic interpretation was perceived. The retinotopic interpretation was invisible and is shown only for illustration.*

## 4.1 Methods

The methods for Experiment 3 were identical to Experiment 1 unless noted otherwise below. Ten new, naïve observers participated (mean age 24.4 years, *SD* = 1.6 years, 6 female, 2 left-handed, 1 with glasses, no exclusions).

Conditions 1-3 were identical to Experiment 1. In Condition 1, we determined the baseline performance level using a Ternus-Pikler display with non-retinotopic rotation in the middle disk, and no dots in the left and right disk (Movie 8). In Conditions 2 and 3, we used the Ternus-Pikler display with retinotopic rotation in one disk and linear up-down or left-right retinotopic motion in the other disk (cf. Experiment 1, Figure 1b, and Movies 3-6). The observers reported the non-retinotopic rotation direction in Condition 2 and the retinotopic rotation in Condition 3. In Condition 4, there were retinotopic rotations in *two* disks, and the dot in the third disk was always in the center (Figure 4a and Movies 11-13). The directions of both retinotopic and the non-retinotopic rotations were combined in a factorial fashion and presented in random order. In Condition 5, we presented visible *non-ret*inotopic rotations in all three disks and observers reported the rotation direction of the middle disk (Figure 4b and Movies 14-17). The rotation direction in the left and right disk could either be in line with or in opposite direction of the middle risk rotation. There was no retinotopic rotation. Retinotopically the dots moved linearly up-down or left-right in every second frame. In Condition 6, there was a non-retinotopic rotation in the middle disk and the dots in the left and right disk were positioned randomly, with the restriction that only linear motion could occur between the randomly chosen dot positions (Movie 9). This was done to control whether the trend towards decreased performance in a similar condition in Experiment 1 can be explained by an incidental occurrence of incongruent rotation in the randomly placed dots.

Observers performed one block of 80 trials per condition. Other than in Experiment 1, all conditions were run in random order and with three disks. Before the experiment, each observer performed training blocks with the baseline stimulus of 20 trials each, until performance levels were above 70%. Seven observers met this criterion already in the first training block, two observers in the second training block, and one observer in the third training block.

## *4.2 Results and discussion*

In Experiment 3, Conditions 1-3 were identical to Experiment 1 and showed the same effects. In Condition 1, observers reported the direction of a non-retinotopic dot rotation in the middle disk and all other dots were omitted. Performance was very good (C1: 91.1%). In Conditions 2-3, we used the Ternus-Pikler with one invisible retinotopic rotation (Figure 1b). In Condition 2, observers reported the non-retinotopic rotation direction. Performance was worse than in Condition 1 (C2$_{all}$: 83.9%; C2$_{all}$ vs. C1: $t(9)$ = 2.16, $p$ = .059, mean diff. = 7.3%), which was again mainly due to worse performance in trials in which the retinotopic and non-retinotopic rotation direction were in opposite directions (C2$_{opposite}$: 77.0%; C2$_{opposite}$ vs. C2$_{same}$: $t(9)$ = 2.30, $p$ = .047, mean diff. = 13.8%). When the retinotopic and non-retinotopic rotation were in the same direction, performance was equally good as in Condition 1 (C2$_{same}$: 90.7%; C2$_{same}$ vs. C1: $t(9)$ = 0.13, $p$ = .898, mean diff. = 0.4%). As in Experiment 1, a ceiling effect cannot be ruled out given the good baseline performance. In Condition 3, observers were asked to report the invisible retinotopic rotation direction and performance was hardly above chance-level (C3$_{all}$: 57.5%; C3$_{all}$ vs. 50%: $t(9)$ = 2.26, $p$ = .050). Performance was at chance-level when the direction of the retinotopic and non-retinotopic rotation were

incongruent ($C3_{opposite}$: 49.2%; $C3_{opposite}$ vs. 50%: $t(9) = 0.14$, $p = .891$), and moderately good

when they were congruent ($C3_{same}$: 65.7%; $C3_{opposite}$ vs. 50%: $t(9) = 4.10$, $p = .003$).

In Condition 4, we added a second invisible retinotopic rotation to the stimulus (Figure 4a).

Observers reported the non-retinotopic rotation direction. Overall performance was worse

than in Condition 1 ($C4_{all}$: 79.9%; $C4_{all}$ vs. C1: $t(9) = 3.94$, $p = .003$, mean diff. = 11.3%). When

both retinotopic rotations were in the same direction as the non-retinotopic rotation,

performance was equally good as in Condition 1 ($C4_{same}$: 93.0%; $C4_{same}$ vs. C1: $t(9) = 0.74$, $p =$

.481, mean diff. = 1.9%). However, when one of the two retinotopic rotations was in the

opposite direction and the other in the same direction of the non-retinotopic rotation,

performance was lower compared to when both were in the same direction ($C4_{same/opposite}$:

81.1%; $C4_{same/opposite}$ vs. $C4_{same}$: $t(9) = 4.16$, $p = .002$, mean diff. = 11.9%) and compared to

Condition 1 ($C4_{same/opposite}$ vs. C1: $t(9) = 3.49$, $p = .007$, mean diff. = 10.0%). When both

retinotopic rotations were in the opposite direction of the non-retinotopic rotation,

performance was even lower ($C4_{opposite}$: 64.2%; $C4_{opposite}$ vs. $C4_{same/opposite}$: $t(9) = 3.67$, $p = .005$,

mean diff. = 16.9%). Hence, only the negative effect of the opposite direction retinotopic

rotation added up, but one same and one opposite direction rotation did not cancel each

other.

In Condition 5, we presented non-retinotopic rotations in all three disks, but no retinotopic

rotation (Figure 4b). All three rotations were perceived and observers reported the non-

retinotopic rotation in the middle disk. Overall, performance in this condition was slightly

worse than in Condition 1 ($C5_{all}$: 81.0%; $C5_{all}$ vs. C1: $t(9) = 3.42$, $p = .008$, mean diff. = 10.1%).

But performance differed only very slightly depending on whether the non-retinotopic

rotations in the flanking disks were both in the same ($C5_{same}$: 83.5%) or both in the opposite

direction of the middle disk (C5$_{opposite}$: 77.0%; C5$_{same}$ vs. C5$_{opposite}$: $t(9)$ = 1.86, $p$ = .096, mean diff. = 6.5%). When one of the rotations was in the same and one in the opposite direction of the middle disk rotation, performance did not differ significantly compared to when both were in the same or in the opposite direction (C5$_{same/opposite}$: 81.8%; C5$_{same/opposite}$ vs. C5$_{same}$: $t(9)$ = 0.46, $p$ = .655, mean diff. = 1.7%; C5$_{same/opposite}$ vs. C5$_{opposite}$: $t(9)$ = 1.98, $p$ = .079, mean diff. = 4.8%).

Finally, in Condition 6 we presented the non-retinotopic rotation in the middle disk with randomly placed dots in the left and right disks. Performance was slightly lower than in Condition 1, where the left and right disk had no dots at all (C6: 83.1%; C6 vs. C1: $t(9)$ = 2.79, $p$ = .021, mean diff. = 8.0%). This is somewhat surprising, since performance did not decrease significantly in a similar condition of Experiment 1, where we did not prevent the incidental occurrence of rotations in the random dots.
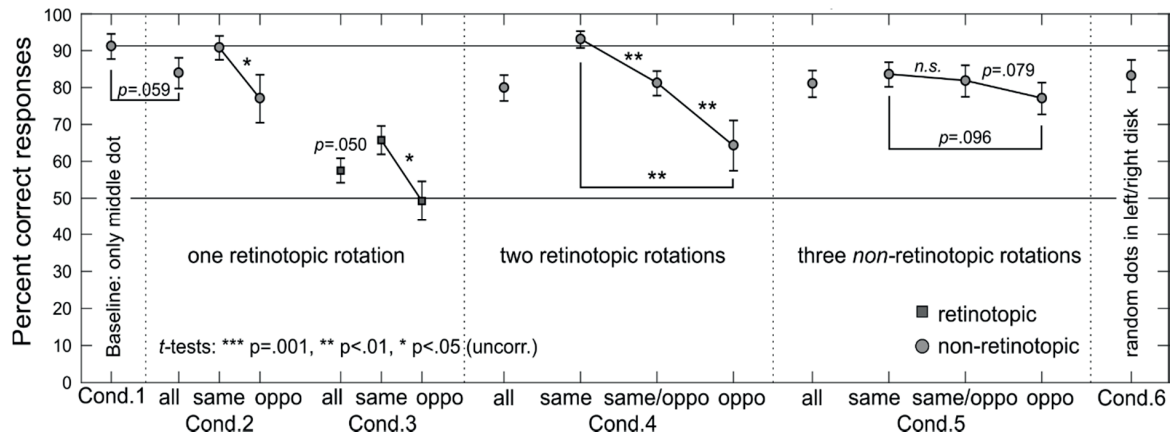
*Figure 5 Motion direction discrimination performance in Experiment 3. Conditions 1-3 replicate the findings of Experiment 1: Unconscious processing of an incongruent retinotopic rotation interferes with non-retinotopic rotation perception. In Condition 4, we used a stimulus with two retinotopic rotations. Performance decreased when one retinotopic rotation direction was incongruent with the non-retinotopic rotation, compared to when all were in the same direction. Performance decreased even more when both retinotopic rotations were incongruent with the non-retinotopic rotation. In Condition 5, we presented non-retinotopic rotations in all three disks and there was no retinotopic rotation. The participants reported the rotation in the middle disk. Performance was slightly lower than in the baseline condition (Condition 1), and differed only little depending on whether the left and right disks rotated in the same or opposite sense of the middle disk. In Condition 6, the dots in the left and right disk were placed randomly and the middle disk rotated. Performance was slightly lower than in the baseline condition. Error bars depict one standard error of the mean.*

## 5 General Discussion

Research on unconscious perception traditionally involves the presentation of a stimulus that is visible when presented alone, but invisible when presented together with a second stimulus. The invisible stimulus, even though it is not consciously perceived, can alter subsequent processes. For example, in masked priming, an invisible prime can speed up responses to subsequently presented targets.

Explanations of invisibility are ex- or implicitly based on retinotopic, inhibitory circuits. For example in masking, the processing of the first stimulus is inhibited by the second stimulus in early, retinotopic visual areas (e.g., Enns & DiLollo, 2000; Lamme, 2006; Fahrenfort, Scholte,

& Lamme, 2007; Breitmeyer & Öğmen, 2006). Effects of the invisible stimulus on subsequent stimuli are explained by the unconscious processing that took place up to that point, or by processing that takes place within a parallel system which does not underlie visibility.

Here, we found that, as in the classic retinotopic paradigms, unconscious retinotopic processing can influence conscious non-retinotopic processing to a substantial degree. As in ambiguous figures, all display elements were fully visible and only the retinotopic interpretation(s) of the display were suppressed. Nonetheless, the suppressed retinotopic motion interpretation strongly affected the conscious non-retinotopic motion percept. Similar effects have been reported with ambiguous figures: Conscious perception of the background was suppressed. Nonetheless, unconscious processing of the background caused responses to objects presented later to be slower when they were semantically related, compared to when they were unrelated (Peterson & Kim, 2001; Peterson & Skow, 2008).

Obviously, retinotopic inhibition cannot explain the invisibility of the retinotopic interpretations because the disks themselves are clearly visible. In addition, also the interpretation needs to be suppressed on an object, rather than a retinotopic level, because the Ternus-Pikler display can largely change. For example, one can use larger inter-disk spacings or different elements than disks, such as squares (Boi et al., 2009; Lauffs, Öğmen, & Herzog, 2017; Petersik & Rice, 2006). Finally, in masking the strong effects of the unconscious prime on conscious elements are usually explained by a pre-activation of the motor system. In congruent cases, reaction times speed up because the "correct" response is pre-activated, in incongruent cases they slow down for the same reason. Reaction times play no role in our paradigm because responses are not speeded and the displays goes on for more than 2.1 secs and at least two frames of 133 ms are needed to compute the rotation direction. Future

research needs to address whether longer durations of incongruent, retinotopic rotations lead to stronger effects on the non-retinotopic percept.

What mechanism can explain the influence of the invisible retinotopic rotation on the conscious non-retinotopic percept? By adding a second retinotopic rotation to the Ternus-Pikler display, we could investigate how multiple invisible retinotopic rotations interact during unconscious processing.

In the simplest model that comes to mind, unconscious retinotopic motion signals are summed up at a retinotopic integration stage, which then influences the non-retinotopic, conscious percept. Indeed, when two retinotopic rotations are incongruent with respect to the non-retinotopic rotation, performance changes much more strongly than when only one incongruent motion is presented (Figure 5). However, this model fails for two reasons. First, one incongruent and one congruent retinotopic rotation do not cancel each other. Performance decreases similarly to the condition with only one incongruent rotation (Figure 5). Second, it seems that congruent retinotopic rotations do not influence the non-retinotopic motion. This holds true even when two congruent rotations are presented (Figure 5). In this case the two retinotopic and the non-retinotopic rotations spin in the same direction. Also linear retinotopic motion had no influence on the non-retinotopic percept, further indicating that it is not the sheer retinotopic processing load, which changes performance. Hence, it seems that mainly incongruent signals inhibit each other. Our results may be explained by a model, in which incongruent retinotopic signals are summed up, which then inhibit the non-retinotopic percept. The effects are substantial. In the condition with two incongruent retinotopic rotations, performance dropped by about 20-30 % compared to a no motion condition and the congruent and linear motion conditions (Figure 5).

One caveat is that performance in the latter conditions was close to ceiling performance (≥ 84%). However, the negative effect of incongruent rotations was so substantial that we would have expected the congruent rotations to cause some form of improvement, in particular, when there were two of them.

The Ternus-Pikler display is a versatile tool to investigate long-lasting unconscious processing without suppressing the elements of the display itself. Traditional, retinotopic mechanisms are unlikely at work since what is perceived unconsciously and consciously depends on complex spatio-temporal processing, such as establishing group motion. The Ternus-Pikler display allows to investigate how unconscious, retinotopic processing interacts with each other and the conscious non-retinotopic processing, and how rivaling interpretations influence each other. The Ternus-Pikler display can be flexibly adjusted to the needs of the researcher, for example by adding additional disks and rotations, or using other visual features (see e.g., Boi et al., 2009). Thereby, it allows to ask different questions than cannot be answered with traditional paradigms.

## Acknowledgements

## Funding

## Conflicts of interest

All authors declare to have no conflict of interest.

# References

Ağaoğlu, M. N., Clarke, A. M., Herzog, M. H., & Öğmen, H. (2016). Motion-based nearest vector metric for reference frame selection in the perception of motion. *Journal of vision*, *16*(7), 14-14.

Bach, M. (1996). The Freiburg Visual Acuity Test-automatic measurement of visual acuity. *Optometry & Vision Science*, *73*(1), 49-53.

Bachmann, T., & Francis, G. (2013). *Visual masking: Studying perception, attention, and consciousness*. Oxford, UK: Academic Press.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, *96*(1), 145.

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and mind*, *2*(1), 5-38.

Boi, M., Öğmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino-vs. non-retinotopic processing. *Journal of Vision*, *9*(13), 5-5.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433-436.

Breitmeyer, B., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (No. 41). Oxford University Press.

Clarke, A. M., Öğmen, H., & Herzog, M. H. (2016). A computational model for reference-frame synthesis with applications to motion perception. *Vision Research*, *126*, 242-253.

Duncker, K. (1929). Über induzierte Bewegung. *Psychologische Forschung*, *12*(1), 180–259.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*(2), 181-192.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345-352.

Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, *19*(9), 1488-1497.

Herzog, M. H., Hermens, F., & Öğmen, H. (2014). Invisibility and interpretation. *Frontiers in Psychology*, *5*.

Johansson, G. (1950). Configurations in the perception of velocity. *Acta Psychologica*, *7*, 25-79.

Johansson, G. (1974). Vector analysis in visual perception of rolling motion. *Psychological Research*, *36*(4), 311-319.

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, *38*(4), 379-393.

Kovacs, I., Papathomas, T. V., Yang, M., & Fehér, Á. (1996). When the brain changes its mind: interocular grouping during binocular rivalry. *Proceedings of the National Academy of Sciences*, *93*(26), 15508-15511.

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494-501.

Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565), 549.

Lauffs, M. M., Öğmen, H., & Herzog, M. H. (2017). Unpredictability does not hamper nonretinotopic motion perception. *Journal of Vision*, *17*(9), 6-6.

Öğmen, H. & Herzog, M. H. (2010). The geometry of visual perception: Retinotopic and nonretinotopic representations in the human visual system. *Proceedings of the IEEE, 98*(3), 479-492.

Petersik, J. T., & Rice, C. M. (2006). The evolution of explanations of a perceptual phenomenon: A case history using the Ternus effect. *Perception*, *35*(6), 807-821.

Peterson, M. A., & Kim, J. H. (2001). On what is bound in figures and grounds. *Visual Cognition*, *8*(3-5), 329-348.

Peterson, M. A., & Skow, E. (2008). Inhibitory competition between shape properties in figure-ground perception. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 251.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437-442.

Pikler, J., (1917). *Sinnesphysiologische Untersuchungen*. Leipzig, Germany: Barth.

Ternus, J., (1926). Experimentelle Untersuchung über phänomenale Identität. *Psychologische Forschung, 7*, 81–136.

Tsuchiya, N., & Koch, C. (2004). Continuous flash suppression. *Journal of Vision*, *4*(8), 61-61.

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, *56*(2), 366-383.

Wheatstone, C. (1838). Contributions to the physiology of vision.--Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, 371-394.

Wolfe, J. M. (1984). Reversing ocular dominance and suppression in a single flash. *Vision Research*, *24*(5), 471-478.

World Medical Association (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Jama*, *310*(20), 2191-2194.

Zhaoping, L. (2005). Border ownership from intracortical interactions in visual area V2. *Neuron*, *47*(1), 143-153.

Zhou, H., Friedman, H. S., & Von Der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*(17), 6594-6611.

# Appendix 3

Lauffs, M. M., Shaqiri, A., Brand, A., Roinishvili, M., Chkonia, E., Öğmen, H., & Herzog, M. H. (2016).
Local versus global and retinotopic versus nonretinotopic motion processing in schizophrenia patients.
*Psychiatry Research*, *246*, 461-465.

Marc M. Lauffs[1], Albulena Shaqiri[1], Andreas Brand[2], Maya Roinishvili[3,4], Eka Chkonia[3,5], Haluk Öğmen [6,7], Michael H. Herzog[1]

[1] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] Institute for Psychology and Cognition Research, University Bremen, Bremen, Germany
[3] Institute of Cognitive Neurosciences, Agricultural University of Georgia, Tbilisi, Georgia
[4] Vision Research Laboratory, Beritashvili Centre of Experimental Biomedicine, Tbilisi, Georgia
[5] Department of Psychiatry, Tbilisi State Medical University, Tbilisi, Georgia
[6] Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA
[7] Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX, USA

CrossMark

# Local versus global and retinotopic versus non-retinotopic motion processing in schizophrenia patients

Marc M. Lauffs[a,*], Albulena Shaqiri[a], Andreas Brand[b], Maya Roinishvili[c,d], Eka Chkonia[c,e], Haluk Öğmen[f], Michael H. Herzog[a]

[a] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[b] Institute for Psychology and Cognition Research, University Bremen, Bremen, Germany
[c] Institute of Cognitive Neurosciences, Agricultural University of Georgia, Tbilisi, Georgia
[d] Vision Research Laboratory, Beritashvili Centre of Experimental Biomedicine, Tbilisi, Georgia
[e] Department of Psychiatry, Tbilisi State Medical University, Tbilisi, Georgia
[f] Department of Electrical & Computer Engineering, University of Denver, Denver, CO, USA

## ARTICLE INFO

## ABSTRACT

Schizophrenia impairs cognitive functions as much as perception. For example, patients perceive global motion in random dot kinematograms less strongly, because, as it is argued, the integration of the dots into a single Gestalt is complex and therefore deteriorated. Similarly, the perception of apparent motion is impaired, because filling-in of the illusory trajectory requires complex processing. Here, we investigated very complex motion processing using the Ternus-Pikler display. First, we tested whether the *perception* of global apparent motion is impaired in schizophrenia patients compared to healthy controls. The task requires both the grouping of multiple elements into a coherent Gestalt and the filling-in of its illusory motion trajectory. Second, we tested the perception of rotation in the same stimulus, which *in addition* requires the computation of non-retinotopic motion. Contrary to earlier studies, patients were not impaired in either task and even tended to perform *better* than controls. The results suggest that complex visual processing itself is not impaired in schizophrenia patients.

## 1. Introduction

Schizophrenia impairs perception to the same extent as cognition. While cognitive deficits are well investigated, less is known about visual deficits, even though they are amongst the most sensitive endophenotypes of schizophrenia (Chkonia et al., 2010).

Interestingly, simple local motion processing is not impaired in schizophrenia, whereas global and complex motion processing are disturbed. For example, discriminating the motion direction of simple drifting gratings is equally good in both patients and controls (Chen et al., 2003; Tibber et al., 2015). However, discriminating the coherent motion direction of a small subset of dots within a cloud of randomly moving dots, which requires both motion integration and segregation across dots, is deteriorated by a factor of about two (Chen et al., 2003; for a review, see Chen, 2011; Butler et al., 2008).

A special case of motion perception is apparent motion, where two elements are flashed at different locations separated by an inter-stimulus interval (ISI). Observers perceive motion from one element to the other, rather than two consecutively presented elements

(Wertheimer, 1912). Apparent motion processing has been suggested to be impaired in schizophrenia patients. However, experimental results are mixed (Saucer and Deabler, 1956; Saucer, 1958, 1959; Chambers and Wilson, 1986; Sanders et al., 2013; but see Sanders et al., 2012; Saucer, 1958; Tschacher et al.,2008). A deficit might be expected, because apparent motion needs complex processing, such as filling in the illusory motion trajectory between the elements.

Here, we investigated even more complex motion processing in schizophrenia by using the Ternus-Pikler display (TPD), combining apparent motion processing and non-retinotopic interpretation. Non-retinotopic processing is a fundamental aspect of visual perception. For example, if visual information would be processed strictly retinotopically, eye movements could not be differentiated from real motion in the world, which would arguably make it impossible to create a stable image of the world. In the TPD, three disks are briefly presented on a computer screen (Fig. 1). Following an ISI of variable duration, the disks reappear shifted by one position to the right. After another ISI, the disks reappear in their original position and the sequence starts over again (see animation on http://lpsy.epfl.ch/). The ISI determines
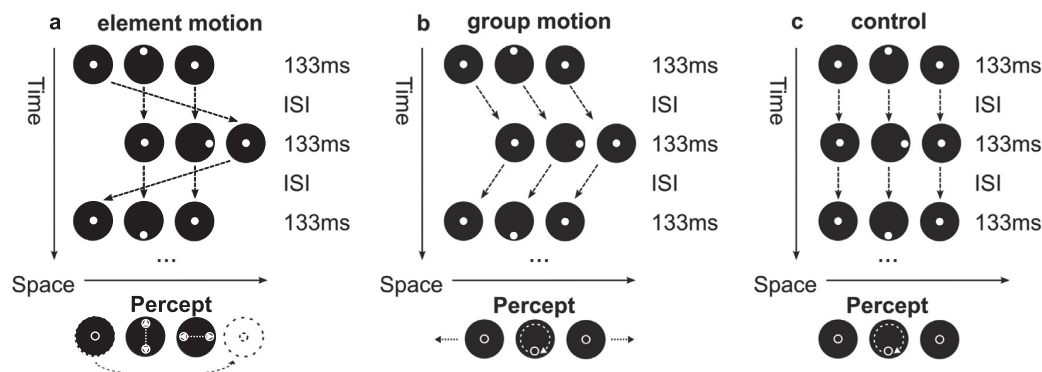
---

**Fig. 1.** Ternus-Pikler display. Three black disks with white dots are presented (in Cond. 1, the dots were not presented). After an ISI, the disks reappear shifted to the right by one position, so that the leftmost disk in frame 2 spatially overlaps with the middle disk in frame 1. Following another ISI the sequence starts over again. a) For short ISIs, the two central disks are perceived as stationary and the dots appear to move linearly up-down (left disk) and left-right (right disk). The outermost disk appears to "jump" back and forth. The arrows indicate the perceived object correspondence (arrows were not shown in the experiment). b) For long ISIs, the three disks are perceived to move left-right in tandem. In the middle disk, a dot appears to rotate. The dot rotation is a non-retinotopic combination of the linear up-down and left-right motions, which are themselves not visible. c) As a control condition, we presented the three disks without lateral displacement and with a long ISI of 200 ms. The dot in the middle disk rotates either clockwise or counterclockwise. This rotation is retinotopic. © Marc M. Lauffs.

the global motion percept: If the ISI is long (e.g., 100 ms), the three disks form a group that is perceived to move back and forth horizontally (*group motion*; Fig. 1b). If the ISI is short (e.g., 0 ms), two stationary disks are perceived, with the outermost disk appearing to jump from left to right and back (*element motion*; Fig. 1a).

First, we parametrically varied the ISI and participants reported whether they perceived element or group motion. Next, we added a dot to each disk. The motion of the dot is determined by the ISI: When element motion is perceived, the dots appear to move linearly up-down in the left central disk and left-right in the right central disk (Fig. 1a). In the "jumping" outermost disk, the dot is always perceived in the center. The perceived motion is *retinotopic*, i.e., the percept is in accordance with the stimulation on the retina. In the group motion condition, the dots in the two outer disks appear stationary, whereas the dot in the middle disk appears to rotate either clockwise or counter-clockwise (Fig. 1b). The rotation percept is *non-retinotopic*, because there is no rotation on the retina. The rotation is a combination of the retinotopic left-right and up-down motions (see arrows in Fig. 1), which are invisible themselves.

Global motion processing *precedes* local motion processing in the Ternus-Pikler display because the global disk motion must first be computed before the local dot motion can be computed relative to it (Clarke et al., 2013; but see Pooresmaeili et al., 2012). Hence, patients should be strongly deteriorated with the Ternus-Pikler display, if global motion processing is deficient.

## 2. Methods

### 2.1. Participants

Twenty-two adults diagnosed with schizophrenia and 20 healthy controls matched in age and education participated in the experiment. All participants had normal or corrected-to-normal vision, with a visual acuity of ≥0.8 (corresponding to 20/25) at least in one eye, as determined with the Freiburg Visual Acuity Test (Bach, 1996).

Schizophrenia patients were recruited from the Tbilisi Mental Health Hospital. Among the patients, 8 were inpatients and 14 outpatients. Healthy controls were recruited from the general population. General exclusion criteria were drug or alcohol abuse, neurological or other somatic illnesses. Participants were no older than 53 years.

Ethics approval was obtained in Tbilisi from the Georgian National Council on Bioethics. All participants signed informed consent and were informed that they could quit the experiments at any time.

Patients were diagnosed according to DSM-IV by means of an

interview based on the SCID, information of the staff, and the study of the records. Psychopathology of schizophrenia patients was assessed by an experienced psychiatrist (EC) by Scales for the Assessment of Negative Symptoms and Scales for the Assessment of Positive Symptoms (SANS, SAPS; Andreasen, 1983, 1984). All patients were receiving neuroleptic medication. Chlorpromazine equivalents and group characteristics are depicted in Table 1. Severity of positive and negative symptoms and CPZ dose were comparable to earlier studies with subjects from the same hospital (e.g., Plomp et al., 2013) and a hospital in Bremen, Germany (Herzog et al., 2004; Grimsen et al., 2013). In- and outpatients did not fundamentally differ in terms of age, illness duration, education, negative symptoms (SANS) and medication dose. Positive symptoms (SAPS) were higher in the inpatient group ($M$(SD) =11.0(3.6) vs. 8.1(3.0); $t$(20) =2.05, $p$=0.053), who on average also received higher medication doses (chlorpromazine $M$(SD) =697(429) vs. 556(366); $t$(19) =0.8, $p$=0.434). The medication data for one patient were not available.

### 2.2. Stimuli and task

Participants were tested in a dimly lit room. Stimuli were presented at a distance of 0.6 m on a Dell Latitude E5540 laptop (1920*1080 px, 60 Hz) and were programmed with Matlab and Psychtoolbox (Brainard, 1997; Pelli, 1997).

We used a variant of the Ternus-Pikler display (Fig. 1; Ternus, 1926; Pikler, 1917). In each stimulus frame, two horizontally aligned, equidistant black disks were presented for 133 ms, followed by a blank-screen ISI of variable duration. Participants were instructed to fixate on a fixation point presented between and below the two disks. A third disk was displayed either to the left or the right, alternating position

**Table 1**
Demographic data (mean ± S.D.) of schizophrenia patients and healthy controls.

|  | **Schizophrenia patients** | **Healthy controls** |
|---|---|---|
| ***N*** | 22 | 20 |
| **Age (years)** | 39.5 ± 9.3 | 40.4 ± 7.6 |
| **Gender (f/m)** | 5/17 | 6/14 |
| **Education level (years)** | 13.3 ± 2.4 | 14.85 ± 2.4 |
| **Duration of illness (years)** | 15 ± 9.1 | |
| **SANS** | 10.23 ± 5.5 | |
| **SAPS** | 9.4 ± 3.45 | |
| **CPZ** | 609.82 ± 387.1 | |
| **Handedness (R/L)** | 22/0 | 20/0 |

with every frame. The start side of the third disk was chosen randomly with equal probability. Depending on the ISI, there were two percepts: At short ISIs, the two central disks are perceived as stationary at one location, while the third disk is "jumping" from left to right and vice versa (element motion, EM, Fig. 1a). At long ISIs, the three disks appear to move in tandem from left to right and back (group motion, GM, Fig. 1b).

Each trial comprised seven stimulus frames, interleaved by ISIs, and followed by a blank screen of 0.75 s before the next trial started. In Condition 1, the ISI was chosen randomly with equal probability from 0, 17, 33, 50, 67, 200 ms for each trial. In the case of an ISI of 0 ms, the stimulus frames followed each other without intermittent blank screen and the stimulus onset asynchrony was equal to the duration of one stimulus frame (133 ms). Ten trials for each ISI were presented randomly (60 trials in total). The task was to indicate, via push-button presses, whether element motion or group motion prevailed.

In Condition 2, a dot was added to each disk (Fig. 1). Only four frames with dots were presented, preceded by two and followed by one frame without dots. We used only ISIs of 0 and 200 ms, which elicit strong element and group motion percepts, respectively. For each ISI, 24 trials were presented in random order. Trials with clockwise and counterclockwise rotation were presented in random order with equal probability. Participants indicated, via push-button presses, whether they perceived clockwise or counterclockwise rotation. In condition 2, we randomly interleaved 12 control trials in which the three disks were presented in the same position in each frame, resulting in a retinotopic rotation of the middle disk (Fig. 1c). For these trials, an ISI of 200 ms was used.

### 2.3. Stimulus specifics

The disks had a diameter of 1.6° and were presented 0.8° above a central fixation point (red square with 0.07° side lengths). The center-to-center inter-disk distance was 1.9°. The white dots had a diameter of 0.25° and were presented halfway between the disk's center and rim. Background color of the screen was midlevel gray. The start orientation of the rotating dot was chosen randomly to be at 0, 90, 180 or 270 degrees. When a participant failed to respond within 3 s, a beep indicated the omission and the trial was repeated at a random later moment. No other feedback was given.

### 2.4. Statistical analysis

All statistical tests were performed in the free and open-source JASP software (Love et al., 2015). Mixed ANOVAs were Greenhouse-Geisser corrected where appropriate. T-tests for unequal variances (Welch's test) were used where appropriate. Uncorrected tests led to the same conclusions in all cases.

## 3. Results

### 3.1. Condition 1: Group vs. element motion

In Condition 1, the proportion of trials in which group motion was reported increased monotonically with ISI, in both patients and healthy controls (Fig. 2). Overall, neither group reported significantly more group motion than the other ($F(1,40)=0.710$, $p=0.404$, $\eta^2=0.017$, 95% CI for $M_{control} - M_{patient} =[-13.5\%; 5.6\%]$). There was also no significant interaction between *group* and *ISI* in a mixed ANOVA design ($F(5, 200)=1.367$, $p=0.256$, $\eta^2=0.008$, $\varepsilon=0.601$). This shows that the increase in group motion percepts in patients was not significantly different from that in controls. Even without correcting for multiple comparisons, t-tests per ISI were all statistically non-significant.

As can be seen in Fig. 2, patients reported slightly more group motion percepts than controls for the trials with ISIs between 0 and
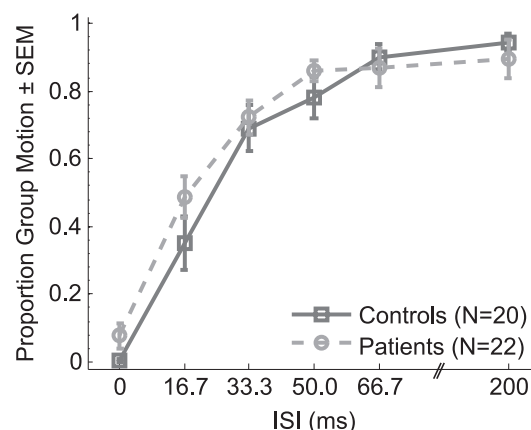


**Fig. 2.** Results Condition 1. Three disks (without the white dots) were shown and we varied the ISI. The proportion of group motion percepts increased with ISI for both groups. We did not find any significant differences between patients and controls..

50 ms and slightly less at the longest ISI of 200 ms. We suggest that these differences are mainly explained by patients' tendency to be more variable in their responses. For example, at the shortest ISI, control subjects reported group motion in only 0.5% (SD=2.24) of the trials. Since it can hardly be lower, the higher variance in patients (SD=17.44) necessarily leads to a higher mean. The inverse holds for the longest ISI, where controls reported group motion perception in 94.5% of the trials, and more variable responding can only lead to a lower mean. Supporting our suggestion, Levene's test for equality of variances was highly significant for ISIs 0 and 50 ms ($p\leq0.001$) and close to significance for ISI 200 ms ($p=0.055$).

### 3.2. Condition 2: Rotation discrimination

In Condition 2, white dots were added to the disks and participants indicated whether the central dot rotated clockwise or counterclockwise. We computed $d'$, a bias-free measure of sensitivity (Fig. 3a; Abdi, 2007; Macmillan and Creelman, 2005) but we also show results as the percentage of correct responses (Fig. 3b).

Patients' and controls' sensitivity did not differ significantly overall, nor was there an interaction between group and ISI ($p=0.616$ and $p=0.683$, respectively). For the 0 ms ISI, element motion was perceived and the rotation was invisible. Clockwise/counterclockwise discriminations were at chance level. Patients had a non-significantly *higher* sensitivity than healthy controls ($t(40)=-2.02$, $p=0.051$, $M_{control} - M_{patient} =-0.30 \pm 0.15$, 95% CI $[-0.59, 0.00]$, Cohen's d=0.6). For trials with 200 ms ISI and the control trials with retinotopic rotation, sensitivity did not differ significantly between the two groups ($p=0.962$ and $p=0.990$, respectively).

Likewise, the percentage of correct responses did not differ significantly between the groups and there was no interaction between group and ISI ($p=0.652$ and $p=0.858$, respectively). All planned comparisons were non-significant (all $ps > 0.24$).

### 3.3. Correlations

Overall, we did not find any coherent relation between performance and psychopathology. We correlated the positive and negative symptoms scores (SAPS/SANS) with group motion perception per ISI using Pearson's r. None of the correlations was significant (Cond. 1: all $ps > 0.567$ for positive symptoms, all $ps > 0.197$ for negative symptoms; Cond. 2: all $ps > 0.540$ for positive, all $ps > 0.202$ for negative symptoms). Illness duration did not correlate with the subjective ratings in Condition 1 (all $ps > 0.371$). In Condition 2, illness duration only correlated with sensitivity for the ISI of 0 ms ($r =-0.489$, $p=0.021$; all
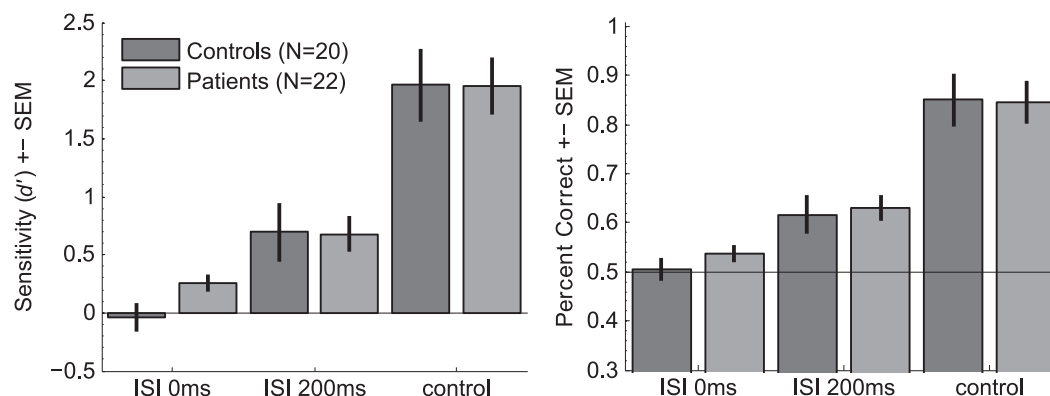
**Fig. 3.** Results Condition 2. Participants discriminated the rotation direction of the dot in the middle disk. Results are shown in terms of sensitivity d′ (left) and percent correct responses (right). Performance is at chance level for an ISI of 0 ms, because no group motion is perceived and the dots appear to move linearly up-down or left-right. Performance improves for an ISI of 200 ms and in the control condition.

other $ps > 0.690$). Chlorpromazine dose tended to correlate with group motion perception in Condition 1 for the ISI of 0 ms ($r = 0.416$, $p=0.06$), but not for the other ISIs (all $ps > 0.133$). In Condition 2, chlorpromazine dose only correlated with sensitivity for the 200 ms ISI ($r = -0.54$, $p=0.011$; all other $ps > 0.449$).

## 4. Discussion

Non-retinotopic processing is critical in most real world situations. For example, a reflector appears to move on an orbital trajectory on the wheel of a moving bicycle. However, the "true" trajectory is very different, namely, cycloidal (Fig. 4). We cannot perceive the true motion trajectory because we subtract the horizontal motion of the bike from the motion trajectory of the reflector. The bicycle motion serves as a reference for the reflector motion (Duncker, 1929; Johansson, 1974). Motion processing in the Ternus-Pikler display follows exactly the same principles as in the bike example. Clearly, non-retinotopic motion processing is as important as it is complex. For this reason, we expected schizophrenia patients to be highly disturbed in the Ternus-Pikler display, because it involves (a) global motion processing, (b) the computation of apparent motion trajectories, and (c) the computation of *non-retinotopic* motion. However, we found

patients to perform almost identical to controls.

First, we tested the *subjective* perception of the patients by changing the ISI. In line with previous studies, which found diminished apparent motion perception in patients (Sanders et al., 2013; Chambers and Wilson, 1968; Saucer and Deabler, 1956; Saucer, 1958), we expected that the change from element to group motion occurs at a much longer ISI in the patients, because the local, retinotopic motion dominates the processing of the global, non-retinotopic percept. However, if anything, patients in our study perceived group motion for *shorter* ISIs than controls. Hence, there is no evidence for a global motion deficit on the phenomenal level.

Next, we tested accuracy. We added dots to the disks and used two ISIs, eliciting either strong element motion or strong group motion percepts. In the case of element motion, we expected that all participants would be unable to see the non-retinotopic and illusory dot rotation because the retinotopic up-down, left-right dot motion percepts dominate. This was indeed the case. In the group motion condition, we expected patients to be much worse than controls, because complex global motion processing is required to establish group motion, which in turn is needed to compute the non-retinotopic dot rotation. However, performance in patients and controls was very similar. This is also true for the control condition, where perceiving the
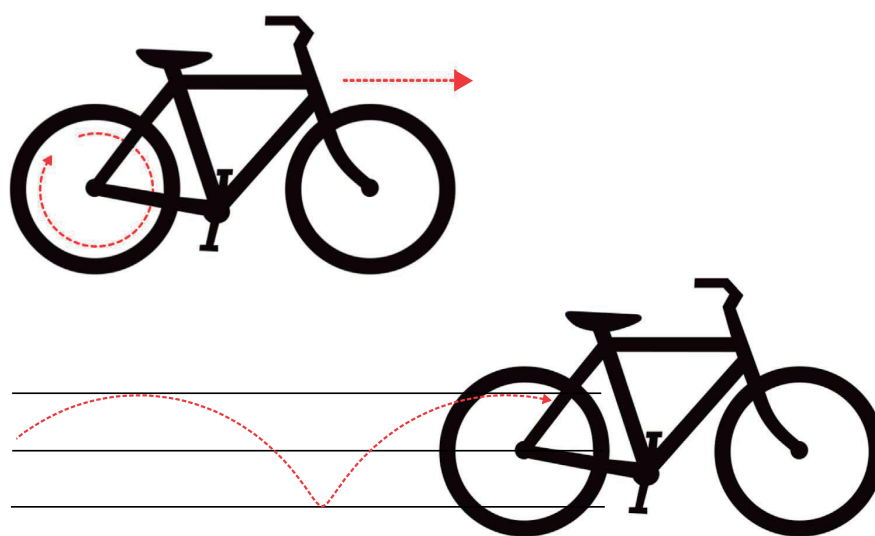


**Fig. 4.** The parts of a moving object are not perceived in retinal coordinates, but relative to the object. For example, a reflector on the wheel of a moving bicycle is perceived as circling, although its trajectory is cycloidal in retinal coordinates. It is perceived to circle, because the linear bicycle motion is subtracted from the cycloid motion. © Marc M. Lauffs.

dot rotation did not require non-retinotopic integration.

We like to mention that with an average $d'$ of around 0.6 (62–63% correct responses), dot rotation discrimination performance was much lower than in previous studies with healthy students (Boi et al., 2009). Experimental procedure, stimulus size, timing and other stimulus parameters were comparable across studies. Potentially, the difference is explained by the fact that the participants of the current study were older and had a more diverse educational background.

As with the Ternus-Pikler display, studies with other apparent motion paradigms did not find differences between patients and controls (Tschacher et al., 2008; Sanders et al., 2012; Saucer, 1958). Sanders and colleagues (2013) did not find a significant difference between the mean data of patients and controls. However, when they determined the alternation rate which yielded strongest apparent motion for each participant, significant differences between patients and controls were found. This analysis is not possible with the Ternus-Pikler display, because strongest group-motion percepts occur at the longest ISI for all observers alike.

We can only speculate about why the schizophrenia patients are not impaired in the Ternus-Pikler display. Potentially the Ternus-Pikler display is less attention demanding than other paradigms, such as the random dot kinematogram, where very close attention is needed. On the contrary, both group and dot motion are easily detected in the Ternus-Pikler display.

Whatever the exact mechanisms are, we can conclude that complexity of visual processing itself is not disturbed in the patients because the TPD requires the solution of many complex and interacting tasks.

## Conflicts of interest

## Funding

## Contributors

MML, AS, AB, MR, EC, HO, and MMH initiated and designed the experiment and wrote the paper. MR and EC collected the data. MML programmed the experiment and analyzed the data.

All authors contributed to and have approved of the final manuscript.

## References

Abdi, H., 2007. Signal detection theory. In: Salkind, N. (Ed.), Encyclopedia of Measurement and Statistics. Sage, Thousand Oaks, CA.

Andreasen, N.C., 1983. Scale for the Assessment of Negative Symptoms (SANS). University of Iowa, Iowa City.

Andreasen, N.C., 1984. The Scale for the Assessment of Positive Symptoms (SAPS). University of Iowa, Iowa City.

Bach, M., 1996. The Freiburg Visual Acuity Test-automatic measurement of visual acuity. Optom. Vis. Sci. 73 (1), 49–53.

Boi, M., Öğmen, H., Krummenacher, J., Otto, T.U., Herzog, M.H., 2009. A (fascinating) litmus test for human retino-vs. non-retinotopic processing. J. Vis. 9 (13), (5-5).

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Butler, P.D., Silverstein, S.M., Dakin, S.C., 2008. Visual perception and its impairment in schizophrenia. Biol. Psychiatry 64 (1), 40–47.

Chambers, J.L., Wilson, W.T., 1968. Perception of apparent motion and degree of mental pathology. Percept. Mot. Skills 26 (3), 855–861.

Chen, Y., Nakayama, K., Levy, D., Matthysse, S., Holzman, P., 2003. Processing of global, but not local, motion direction is deficient in schizophrenia. Schizophr. Res. 61 (2), 215–227.

Chen, Y., 2011. Abnormal visual motion processing in schizophrenia: a review of research progress. Schizophr. Bull. 37 (4), 709–715.

Chkonia, E., Roinishvili, M., Makhatadze, N., Tserava, L., Stroux, A., Neumann, K., Herzog, M.H., Brand, A., 2010. The shine-through masking paradigm is a potential endophenotype of schizophrenia. PLoS ONE 5 (12), e14268.

Clarke, A.M., Repnow, M., Öğmen, H., Herzog, M.H., 2013. Does spatio-temporal filtering account for nonretinotopic motion perception? Comment on Pooresmaeili, Cicchini, Morrone, and Burr (2012). J. Vis. 13 (10), 1–14.

Duncker, K., 1929. Uber induzierte Bewegung [On induced motion]. Psychol. Forsch. 12, 180–259.

Grimsen, C., Brand, A., Fahle, M., 2013. No evidence for prolonged visible persistence in patients with schizophrenia. PloS One 8 (3), e58940.

Herzog, M.H., Kopmann, S., Brand, A., 2004. Intact figure-ground segmentation in schizophrenia. Psychiatry Res. 129 (1), p55–p63.

Johansson, G., 1974. Vector analysis in visual perception of rolling motion. Psychol. Forsch. 36 (4), 311–319.

Love, J., Selker, R., Marsman, M., Jamil, T., Verhagen, A.J., Ly, A., Wagenmakers, E.J., 2015. JASP (Version 0.6. 5). Computer software

Macmillan, N.A., Creelman, C.D., 2005. Detection Theory: A User's Guide. Lawrence Erlbaum Associates, New York.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10 (4), 437–442.

Pikler, J., 1917. Sinnesphysiologische Untersuchungen [Physiological investigations of the senses]. J.A. Barth, Leipzig.

Plomp, G., Roinishvili, M., Chkonia, E., Kapanadze, G., Kereselidze, M., Brand, A., Herzog, M.H., 2013. Electrophysiological Evidence for Ventral Stream Deficits in Schizophrenia Patients. Schizophr. Bull. 39 (3), p547–p554.

Pooresmaeili, A., Cicchini, G.M., Morrone, M.C., Burr, D., 2012. "Non-retinotopic processing" in Ternus motion displays modeled by spatiotemporal filters. J. Vis. 12 (1), 1–15.

Sanders, L.L.O., Muckli, L., de Millas, W., Lautenschlager, M., Heinz, A., Kathmann, N., Sterzer, P., 2012. Detection of visual events along the apparent motion trace in patients with paranoid schizophrenia. Psychiatry Res. 198 (2), 216–223.

Sanders, L.L.O., de Millas, W., Heinz, A., Kathmann, N., Sterzer, P., 2013. Apparent motion perception in patients with paranoid schizophrenia. Eur. Arch. Psychiatry Clin. Neurosci. 263 (3), 233–239.

Saucer, R.T., Deabler, H.L., 1956. Perception of apparent motion in organics and schizophrenics. J. Consult. Psychol. 20 (5), 385.

Saucer, R.T., 1958. A further study of the perception of apparent motion by schizophrenics. J. Consult. Psychol. 22 (4), 256.

Saucer, R.T., 1959. Chlorpromazine and apparent motion perception by schizophrenics. J. Consult. Psychol. 23 (2), 134.

Ternus, J., 1926. Experimentelle Untersuchungen über phänomenale Identität [Experimental investigations of phenomenal identity]. Psychol. Forsch. 7, 81–136.

Tibber, M.S., Anderson, E.J., Bobin, T., Carlin, P., Shergill, S.S., Dakin, S.C., 2015. Local and global limits on visual processing in schizophrenia. PloS One 10 (2), e0117951.

Tschacher, W., Dubouloz, P., Meier, R., Junghan, U., 2008. Altered perception of apparent motion in schizophrenia spectrum disorder. Psychiatry Res. 159 (3), 290–299.

Wertheimer, M., 1912. Experimentelle Studien über das Sehen von Bewegung [Experimental studies on the seeing of movement]. JA Barth.

# Appendix 4

Hochmitz, I., Lauffs, M. M., Herzog, M. H., & Yeshurun, Y. (*in prep.*). Sustained spatial attention affects feature fusion through enhanced signal encoding

Ilanit Hochmitz[1], Marc M. Lauffs[2], Michael H. Herzog[2], Yaffa Yeshurun[1]

[1] Department of Psychology, University of Haifa, Israel
[2] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

**Sustained spatial attention affects feature fusion through enhanced signal encoding**

Ilanit Hochmitz[1], Marc M. Lauffs[2], Michael H. Herzog[2], Yaffa Yeshurun[1]

1. Psychology Department, University of Haifa, Israel

2. Laboratory of Psychophysics, Brain Mind Institute, EPFL, Lausanne Switzerland

# Abstract

When two verniers are presented in rapid succession at the same location feature fusion occurs. Instead of perceiving two separate verniers, participants typically report perceiving one fused vernier, whose offset is a combination of the two previous verniers, with the later one slightly dominating. Here, we examined the effects of sustained attention – the voluntary component of spatial attention – on feature fusion. Attention was manipulated via the degree of certainty regarding the stimulus location. In the attended condition, the stimulus appeared always in the same location, and in the unattended condition it could appear in one of two possible locations. Participants had to report the offset of the fused vernier. Experiments 1a and 1b measured attentional effects on feature fusion with and without eye-tracking. In both experiments, we found a higher rate of reports corresponding to the offset of the second vernier with attention than without attention, suggesting that attention strengthened the final fused percept. In Experiment 2, we manipulated the stimulus duration to encourage a final fused percept that is dominated by either the first or second vernier. We found that attention strengthened the already dominant percept, regardless of whether it corresponded to the offset of the first or second vernier. These results are consistent with an attentional mechanism of signal enhancement at the encoding stage.

# Introduction

Spatial covert attention allows us to grant priority in processing of visual information gathered from a specific location without making eye movements to that location (e.g., Posner, 1980). An extensive amount of evidence suggests that spatial attention has two components: a faster component, attracted to a location by sudden changes in the display – 'transient attention'; and a voluntary slower component, controlled by our goals – 'sustained attention' (e.g., Cheal & Lyon, 1991; Jonides, 1981; Muller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Posner, 1980; Remington, Johnston & Yantis, 1992). These two attentional components have been found to have different effects on behavior, suggesting separate attentional mechanisms (e.g., Briand, 1998; Hein, Rolke & Ulrich, 2006; Klein, 1994; Yeshurun & Carrasco 2008). In this study we focus on the sustained component of attention and examine its effect on feature fusion.

Feature fusion occurs when two stimuli, which differ in one feature, are presented briefly and in rapid succession at the same retinotopic location. Instead of two separate objects, observers typically report perceiving a single object, whose feature is a combination of the features of the two objects (e.g., Efron, 1967; 1973; Hermens, Scharnowski & Herzog, 2009). For example, if a red disk and a green disk are presented in rapid succession, the color of the two disks is fused and a single yellow disk is perceived (e.g., Efron, 1967; 1973; Yund, Morgan, & Efron, 1983). Similarly, if a vernier stimulus is immediately followed by another vernier stimulus with an offset opposite to the first one (i.e., an antivernier), the offsets of the two stimuli fuse and a single vernier is perceived. Since the offsets of the two verniers are in opposite directions they partly cancel each other, so that the perceived vernier offset is much smaller (Figure 1). The small perceived offset is in the direction of the second stimulus' offset (e.g., Herzog, Leseman, & Eurich, 2006; Herzog, Parish, Koch, & Fahle, 2003; Scharnowski, Hermens, & Herzog, 2007; Scharnowski, Hermens, Kammer, Ogmen, &

Herzog, 2007). For example, when a red disc is followed by a green disc, the fused disc

appears yellow with a slight greenish tone, whereas a green disc followed by a red disc looks

slightly reddish (e.g., Efron, 1967; 1973; Yund et al., 1983). Likewise, a vernier stimulus

followed by an antivernier stimulus results in a fused vernier with a slight offset in the

direction of the antivernier stimulus (e.g., Herzog et al., 2003; 2006; 2007; Scharnowski et

al., 2003).
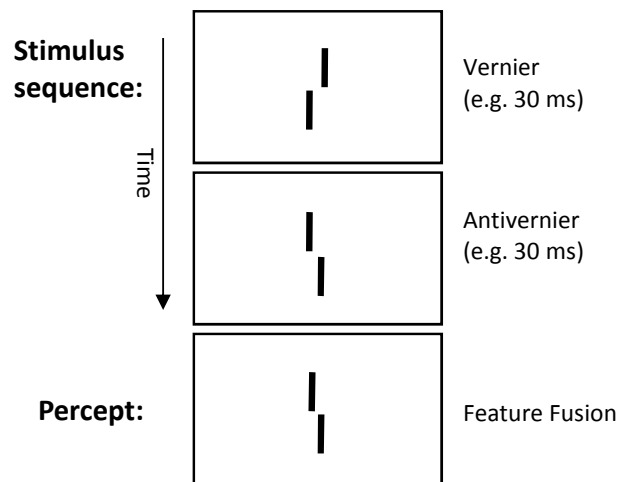


Figure. 1. Feature fusion. When a vernier and an antivernier (a vernier with opposite offset direction) are presented in rapid succession, observers do not perceive the individual elements but only one fused vernier with a smaller offset corresponding to the antivernier.

Feature fusion is considered to be a measure of temporal integration, because the features of

the two stimuli are integrated across time into a single, coherent, perceptual object.

Scharnowski, Hermens, Kammer et al. (2007) proposed a model that describes feature fusion

as a process that dynamically links the activity of neurons coding the offset of the individual

verniers. According to this model, feature detector neurons code the offset information (left

or right) of the individual verniers. Then, their activation is fed into an integration neuron.

Elements presented later influence perception more strongly than earlier ones because the

signal of the preceding vernier decays more than that of the trailing antivernier, causing the

antivernier to dominate the percept. This model can also account for findings demonstrating

that once a sequence of more than one vernier (v) and antivernier (av) is employed, the offset direction of the final percept is highly dependent on the verniers' temporal order and durations. For example, when presenting a v-av-v sequence, increasing the duration of the second vernier in the sequence (i.e., the third stimulus) increased vernier dominance to a larger degree than when the duration of the first vernier was increased, even though the energy of the verniers and antivernier was equal in both cases (e.g., Scharnowski, Hermens, & Herzog, 2007)

Although the model can account for a good deal of the available behavioral data, it can be argued that feature fusion is merely a case of backward masking, in which the antivernier masks the vernier. Yet, with backward masking the visibility of a target is impaired by the mask, whereas with feature fusion one fused vernier is clearly perceived. Critically, with backward masking a more prominent mask generates larger masking effects (i.e., results in poorer target visibility). In contrast, with feature fusion displays, when the single antivernier is replaced by a grating composed of 25 antiverniers (i.e., a considerably stronger 'mask'), the vernier dominates the final perception (i.e., it becomes more visible; e.g., Herzog, Fahle, & Koch, 2001; Herzog & Koch, 2001; Hermens et al., 2009; 2010). Furthermore, which of the two stimuli (the vernier or the antivernier) dominates perception and to what degree seems to depend strongly on the spatio-temporal parameters of the grating. For instance, previous studies (e.g., Hermens et al., 2009) have shown that when the antivernier is embedded in a grating of aligned verniers rather than in a grating of antiverniers as described above, the vernier dominance is slightly reduced. Similarly, increasing the spacing between the elements of the grating reduces vernier dominance. These findings have been interpreted in terms of spatial grouping: When the second stimulus was an antivernier embedded in an array of similar antivernier stimuli, it spatially grouped with the surrounding stimuli, and this prevented temporal fusion with the preceding vernier stimulus. Decreasing similarity between

the second stimulus and the antiverniers in which it was embedded, or increasing the distance between them prevented spatial grouping between the elements in the grating, allowing at least partial fusion to reappear. Overall, these results suggest that feature fusion is not a simple case of backward masking.

The current study examined the effect of spatial sustained attention on feature fusion. Feature fusion, as well as the spatiotemporal interaction that emerges when the antivernier is embedded in a grating, are considered to occur at an early stage of visual information processing (Hermens et al., 2009, 2010). Research on the effects of spatial attention on early visual processing provides some reasons to suppose that feature fusion might be affected by spatial attention. It is well established that attention improves early aspects of visual perception such as contrast sensitivity (e.g., Carrasco, Penpeci-Talgar, & Eckstein, 2000; Cameron, Tai, & Carrasco, 2002; Lu & Dosher, 1998, 2000; Dosher & Lu, 2000; Huang & Dobkins, 2005; Smith, Wolfgang, & Sinclair, 2004; Solomon, 2004) and spatial resolution (e.g., Carrasco, Williams, & Yeshurun, 2002; Golla, Ignashchenkova, Haarmeier, & Their, 2004; Yeshurun & Carrasco, 1998, 1999). For instance, regarding spatial resolution, it has been found that the advance allocation of spatial attention improves performance in acuity tasks such as the detection of a small spatial gap with 'Landolt-squares' (e.g., Carrasco et al., 2002; Golla et al., 2004; Yeshurun & Carrasco, 1999), as well as hyperacuity tasks like discrimination of offset-direction with vernier targets (e.g., Yeshurun & Carrasco, 1999). In addition, directing attention to the target location of a texture target enhanced the ability to segment it from the texture background when this target appeared in the periphery where the spatial resolution was too low for the scale of the texture, but impaired performance at more central locations where the spatial resolution was too high (Yeshurun & Carrasco, 1998, 2000).

Spatial attention has also been found to affect temporal aspects of visual processing. For

instance, it has been found that transient spatial attention degrades temporal resolution – the

ability to resolve rapid luminance changes in time. Using the two-flash fusion paradigm,

several studies demonstrated that observers' ability to detect a brief temporal gap occurring

between two successive light flashes was reduced when observers allocated their attention in

advance to the flashes location (e.g., Rolke, Dinkelbach, Hein, & Ulrich, 2008; Yeshurun,

2004; Yeshurun & Levy, 2003). Similarly, Hein, Rolke, and Ulrich (2006) demonstrated that

automatically oriented attention impairs temporal order discrimination of two spatially

adjacent dots, whereas voluntary oriented attention enhances it. Attentional effects were also

found for the opposing temporal process. Visser and Enns (2001) examined the effects of

attention on temporal integration, using the attentional blink paradigm. In this paradigm

participants are presented with two targets, separated by a temporal lag typically ranging

from 100 ms to 700 ms. When participants are asked to identify the two targets they usually

demonstrate near perfect identification of the first target, whereas the identification of the

second target varies as a function of the lag. Second target identification rates are poorest at

shorter lags and they increase as lag duration increases (Ramond, Shapiro, & Arnell, 1992). It

is suggested that the failure in the identification of the second target for shorter lags can be

attributed to impoverished attention resources due to the processing of the first target

(Shapiro, Arnell, & Raymond, 1997). To measure temporal integration, Visser and Enns used

as the second target a 5 x 5 dots matrix of which one dot was missing. The task was to

localize the missing dot. Critically, the dot matrix was divided into two consecutive frames,

each composed of 12 dots (25 dots minus the missing dot). Hence, dot localization required

integrating the two frames across time. Temporal integration was manipulated by varying the

inter-stimulus interval (ISI) between frames. As expected, missing dot localization improved

with shorter ISIs that allow the two matrix frames to integrate across time. However, this

performance enhancement was less evident when the matrix frames appeared within the time window of the attentional blink (i.e., when the matrix was presented in close temporal proximity to the first target). This finding suggests that under limited attentional resources temporal integration is shortened (Visser & Enns, 2001).

Given these demonstrations of attentional effects on early visual processing, it is reasonable to assume that spatial attention will also affect feature fusion. In the current study we tested this hypothesis. So far feature fusion was only examined in the fovea. However, in order to manipulate spatial attention the visual display should involve spatial uncertainty. To achieve that, we presented the feature fusion display in the periphery, and manipulated spatial sustained attention by varying the degree of uncertainty regarding the stimulus location. The attended condition was a full certainty condition, in which the stimuli – a vernier followed by an antivernier – appeared at the same peripheral location throughout the entire block. This allowed the observers to allocate spatial attention in advance to the stimulus location. In the unattended condition, spatial uncertainty was introduced: The stimuli could equally likely appear in one of two possible peripheral locations (to the right or left of fixation). Hence, in this condition the observers could not allocate attention to the stimulus location in advance. Experiments 1a and 1b examined whether spatial sustained attention affects feature fusion with and without measuring eye-movements. Experiment 2 explored the mechanism underlying the observed attentional effect. In all three experiments, the observers had to report their perceived offset direction, and we measured how often observers reported an offset that matches that of the first stimulus (the vernier offset). If the vernier offset report rate was above 50% it indicated that the vernier dominated the final percept (vernier dominance), and if it was below 50% it indicated that the antivernier (the second stimulus) dominated perception (antivernier dominance). If the vernier offset report was around 50% it indicated that the vernier and the antivernier had a similar contribution to the final percept

(Hermens et al., 2009). Here, we examined the effect of attention on feature fusion by comparing the reported vernier offset direction in the attended and unattended condition. If attention can affect feature fusion, the degree of vernier/antivernier dominance will be different in the two attentional conditions.

## Experiments 1a and 1b

Experiment 1a and 1b were designed to examine whether spatial sustained attention can affect feature fusion. To that end, we measured the extent to which feature fusion occurs with and without sustained attention. In both experiments, feature fusion was produced using a vernier and an antivernier, presented one after the other in the same location with no intermediate time interval. In the attended condition, the stimulus sequence appeared at the same peripheral location throughout the entire block. The participants were informed of this location before the block, stimulus location, which enabled them to attend this location in advance. In the unattended condition, the stimuli could appear in one of two possible locations, producing spatial uncertainty, which prevented the advance allocation of attention to the stimulus location.

Experiments 1a and 1b were similar in terms of stimuli and procedure, but in Experiment 1b eye-movements were tracked to rule out the possibility that any effect found in Experiment 1a might be, at least partially, due to eye movements. If spatial covert attention indeed affect feature fusion we should get similar results with and without controlling for eye-movements.

### Methods

*Participants.* Eight students from the University of Haifa participated in Experiment 1a, and ten students from the École Polytechnique Fédérale de Lausanne (EPFL) participated in Experiment 1b. Participants in both experiments were naive to the purpose of the study, and had normal or corrected to normal vision. The visual acuity of all participants was tested

using the Freiburg visual acuity test (Bach, 1996). Observers had to score at least 1.0 on this test (corresponding to 20/20) in order to participate. This study adhered to the Declaration of Helsinki.

*Stimuli and apparatus.* In Experiment 1a stimuli were presented on a 21" CRT monitor of a PowerMac G4 computer (1,024 × 768 resolution at refresh rate of 120 Hz), using MATLAB and the Psychophysics Toolbox extensions (Brainard, 1997). A dim background light illuminated the room. In Experiment 1b the stimuli were presented on a standard 24" LCD Asus VG248QE computer screen (1920x1080 resolution at refresh rate 120Hz). Also, unlike Experiment 1a, in this experiment eye-movements were recorded using an SMI iViewX Hi-Speed 1250 Eyetracker. Data were recorded binocularly at 500Hz but immediately averaged over both eyes to reduce noise.

The stimuli in both experiments were composed of a sequence of a white vernier and antivernier presented on a black background. The vernier consisted of two vertical lines, each measuring 10' (arc minutes) in height. The two lines were separated by a 1' vertical gap, and were slightly offset horizontally (randomly to the right or left with equal probability). The antivernier had the same spatial parameters as the vernier, except that its offset was in the opposite direction of the vernier (Figure 1). If the vernier was offset to the left, the antivernier was offset to the right and vice versa.

In both experiments the horizontal offset was adjusted for each participant, but using different methods. In Experiment 1a the offset adjustment took place in a preliminary session consisting of 6 blocks of 84 trials per each block. In this session, the participants performed the experimental task with several offset sizes, and the vernier-antivernier sequence always appeared to the right of fixation. The chosen offset had to meet two criteria: (1) It had to yield around 30% vernier dominance (i.e., the vernier offset direction was reported in 30% of the trials); (2) Participants had to perceive only one fused vernier with no apparent motion

percept. This was based on the finding that with large offset sizes some participants reported seeing the upper and lower segments move in opposite directions (Scharnowski, Hermens, Kammer et al., 2007). The chosen horizontal offsets ranged from 60" to 120" (*Mode* = 120"). In Experiment 1b the offsets were adjusted using a PEST staircase procedure, such that the determined offset yielded 25% vernier dominance. The staircase procedure was conducted in the same session but in separate blocks. Each of these blocks consisted of 80 trials identical to the experimental trials in the attended condition. The adjusted offset sizes ranged from 18" to 98" (*Mode*=35").

In the experimental session of both experiments there were two attentional conditions. In the attended condition, the vernier-antivernier sequence was always presented to the right of fixation at 2° of eccentricity. In the unattended condition, the stimuli could appear either to the right or to the left of fixation, with equal probability, at 2° of eccentricity.

*Procedure.* In both experiments participants viewed the stimuli from a distance of 2 m, and were asked to fix their gaze on a central fixation cross throughout the entire trial. Each trial began with a fixation cross. In Experiment 1a the fixation cross was presented for 750 ms, followed by the vernier-antivernier sequence, each presented for 30 ms with no ISI. In Experiment 1b the stimuli were only presented after a fixation on the central fixation cross was detected. Vernier and antivernier presentation time was identical to that of Experiment 1a, which was too brief to allow for eye movements to occur before the end of the presentation (Fischer & Ramsperger, 1984; Gezeck, Fischer, & Timmer, 1997; Rohrer & Sparks, 1993).

The remaining of the procedure was identical in both experiments. The offset direction of the vernier was equally often to the right or to the left, presented in a random order. Each attentional condition was tested in three separate blocks, each consisting of 80 trials. The order of the blocks was randomized across participants and the participants knew in advance

which block they were about to perform (one vs. two possible locations). The participants' task was to indicate the perceived offset direction (i.e., whether the upper line of the vernier was to the left or to the right of the lower line).

Before the experimental session, each participant performed a short practice session comprised of 20 trials per attentional condition.

**Results and discussion**

Overall, in both Experiment 1a and 1b, the percentage of trials in which the observers reported the vernier offset was lower than 50% (Figure 2). Thus, the results of this experiment replicate those of previous studies demonstrating that under such presentation conditions the offset of the antivernier dominates perception (e.g., Herzog et al., 2003; Herzog et al., 2006; Scharnowski, Hermens, Kammer et al., 2007). Most important for the goal of the current study, in both experiments, antivernier dominance differed in the two attentional conditions. Specifically, antivernier dominance was higher in the attended condition than in the unattended condition. A one-way repeated-measures ANOVA, performed on the vernier offset report rate, indicated that this effect was highly significant (Experiment 1a: $F(1,7) = 42.76$, $p < 0.0003$; Experiment 1b: $F(1,9) = 26.21$, $p < 0.0006$), though as evident in Figure 2 the effect in Experiment 1b was smaller ($t(15) = 4.07$, $p < 0.0006$). This difference in the effect size between the two experiments will be discussed in more detail in the *General discussion* section. These findings suggest that sustained spatial attention affects feature fusion. Specifically, with attention the final fused percept was clearer than without attention. Moreover, the fact that the effect of attention on feature fusion was found even when eye movements were precluded (i.e., Experiment 1b) suggests that the

effect observed in Experiment 1a was not merely due to eye movements, rather it is indeed
due to the allocation of covert attention to the stimulus location.

Because feature fusion is considered to occur at a relatively early stage of visual processing,
these results are consistent with previous studies demonstrating the effect of spatial attention
on early or basic aspects of visual processing, such as contrast sensitivity and spatial
resolution (reviewed in Carrasco & Yeshurun, 2009). Yet, the nature of this attentional effect
remains unclear. Experiment 2 aimed to provide a better understanding of the attentional
mechanism responsible for feature fusion enhancement.
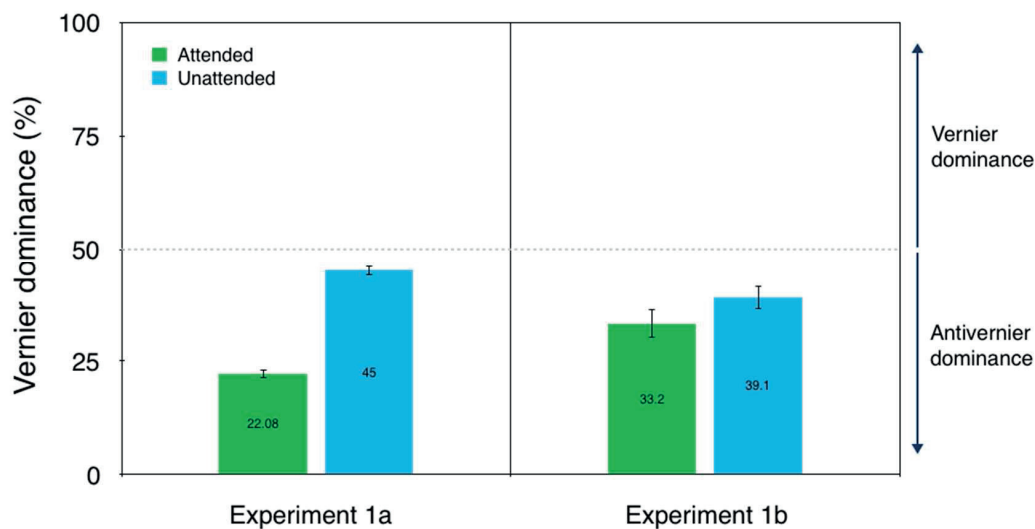


Figure 2. Vernier dominance (percentage of vernier offset report) as a function of the attentional
conditions in Experiments 1a and 1b. Error bars correspond to one standard error of the mean.

**Experiment 2**

Experiments 1a and 1b showed that antivernier dominance, which typically serves as a
marker for feature fusion, increased when participants allocated attention in advance to the
stimulus location. This finding suggests that sustained spatial attention can affect feature
fusion. In the current experiment we examined the mechanism underlying this effect of

attention on feature fusion. One possible explanation of this finding is that attention improves the quality of the representation of each of the stimuli in the sequence and therefore results in a better representation of the fused percept. One mechanism by which attention might exert this kind of effect is signal enhancement. According to the signal enhancement hypothesis, attention improves stimulus encoding at the attended location (e.g., Bashinski & Bacharach, 1980; Carrasco & Yeshurun, 1998; Downing, 1988; Lu & Dosher, 1998; Posner, 1980).

In the current study we tested the hypothesis that attention affects feature fusion by enhancing the representation of the final percept. We took advantage of the fact that the final percept of the fusion stimuli can be modified by manipulating the duration of the stimuli. Specifically, if the duration of the first stimulus in the sequence (the vernier) is longer than the duration of the second stimulus (the antivernier), its offset dominates the final fused percept (vernier dominance). In contrast, if the duration of both stimuli is equal, as was the case in Experiments 1a and 1b, the offset of the second stimulus, the antivernier, dominates perception (e.g., Scharnowski, Hermens & Herzog, 2007). Hence, if attending to the sequence location results in improved representation of the fused percept, this percept should be enhanced regardless of its direction (i.e., regardless of whether it reflects vernier dominance or antivernier dominance).

To test this hypothesis we included in this experiment the same attention conditions, but also two 'dominance' conditions: 25% and 75% vernier dominance. The 25% dominance condition was identical to Experiment 1b. That is, the duration of the two stimuli in the sequence was equal and the staircase procedure was designed to yield 25% vernier dominance. In the 75% dominance condition the vernier was presented twice as long as the antivernier and the staircase procedure was designed to yield 75% vernier dominance. Finally, for comparison, we also added a 'single vernier' condition in which only the vernier was presented (i.e., fusion was not involved).

If sustained attention affects feature fusion through signal enhancement then it should strengthen the already dominant percept. That is, in the 25% condition we expected to replicate the results of Experiments 1a and 1b: Antivernier dominance should be higher in the attended condition compared to the unattended condition. In the 75% condition, the more prominent percept is a fused vernier with an offset direction corresponding to that of the vernier, therefore in this condition, we expected to find higher vernier dominance in the attended than the unattended condition. As for the single vernier condition, Yeshurun and Carrasco (1999) previously found that spatial transient attention enhances vernier offset discrimination. This result was interpreted in terms of signal enhancement. However, while that study employed transient attention manipulated via peripheral cues, our study focuses on the sustained component of spatial attention. It has been previously argued that transient and sustained attention operate at different stages within the visual system, which may account for the fact that sometimes they show differential effects on performance (e.g., Briand, 1998; Chicaa, Bartolomeoa, & Lupiánez, 2013; Hein et al., 2006; Klein, 1994; Yeshurun & Carrasco, 2008). Thus, the effect of transient attention on vernier offset discrimination may or may not be replicated with sustained attention. If *sustained* spatial attention can also affect vernier offset discrimination via signal enhancement, we expected to replicate the results of Yeshurun and Carrasco (1999). In addition, because the single vernier condition does not involve fusion it also allows us to test whether sustained attention affects feature fusion in a unique manner, which is different from its effect on a single vernier offset discrimination.

**Method**

*Participants.*  Eleven students from the École Polytechnique Fédérale de Lausanne (EPFL) participated in this experiment; all were naive to the purpose of the study, and had normal or corrected to normal vision. Their visual acuity was assessed as in Experiment 1b.

*Stimuli, apparatus and procedure.* The stimuli, apparatus and procedure were identical to Experiment 1b except for the following: Two vernier dominance conditions were used, 25% and 75%. The 25% condition was identical to that of Experiment 1b. Both stimuli were presented for 30 ms and the staircase procedure performed before the experimental session was set to generate a final offset that yields 25% vernier dominance. In the 75% condition the vernier was presented for 30 ms and the antivernier was presented for 15 ms. Additionally, the staircase procedure was set to generate a final offset that yields 75% vernier dominance. In the single vernier condition a single vernier stimulus was presented for 30 ms, with an offset identical to that used in the 25% condition. Chosen offset sizes ranged from 50" to 60" (*Mode* = 60") in the 25% condition, and from 55" to 120" (*Mode* = 55") in the 75% condition.

Each dominance condition was run for four blocks of 80 trials, two blocks per attentional condition. The single vernier blocks were always run at the end of the session (i.e., after the 25% and 75% conditions). Whether the 75% or the 25% condition was run first was counterbalanced across participants.

**Results and discussion**

A two-way repeated-measures ANOVA, with the factors dominance (25%, 75%, single) and attention (attended, unattended), was conducted on the vernier offset report rate. As expected given previous studies who used a duration manipulation (e.g., Scharnowski, Hermens, & Herzog, 2007), there was a significant main effect for the dominance condition ($F(2,10) = 49.41, p < 0.0001$; Figure 3): In the 25% condition the antivernier dominated the percept, whereas the vernier dominated in the 75% and single conditions (Figure 3). There was no main effect of attention, but the interaction was highly significant ($F(2,10) = 39.89, p < 0.0005$). *t*-tests indicated that the effect of attention in all three dominance conditions was significant, but in opposite directions: In the 25% condition, vernier offset report was lower

in the attended than unattended condition ($t(10) = 5.95$, $p < 0.0001$), or in other words, antivernier dominance was higher in the attended than the unattended condition. This finding replicates the results of Experiments 1a and 1b. In contrast, in the 75% condition, the vernier offset report rate was significantly higher in the attended than unattended condition ($t(10) = 5.94$, $p < 0.0001$), or in other words, vernier dominance was higher in the attended than unattended condition. Finally, in the single vernier condition, the vernier offset report rate was also significantly higher in the attended than unattended condition ($t(10) = 4.073$, $p < 0.0011$), as expected if sustained attention improves vernier offset discrimination.

Thus, as predicted by the signal enhancement hypothesis, allocating sustained attention to the stimulus location strengthened the final fused percept that was dominant under divided diffused attention. That is, as detailed above, in the 25% condition the antivernier offset dominates, and this antivernier dominance is even higher with attention. In the 75% condition, the longer duration of the vernier compensates for its larger degree of decay and its offset dominates the fused percept. As with the 25% condition, the already dominant percept is even stronger with attention.

Interestingly, as can be seen in Figure 3, the effect of attention (i.e., the difference between

attended and unattended conditions) in the 25% condition ($M = 13.85$) is about twice as large

as the attentional effect in the 75% or single vernier conditions ($M = 6.14$ and $M = 7.11$,

respectively). To test whether this difference is reliable statistically, we performed a one-way

repeated-measures ANOVA on this difference between the attentional conditions, with

dominance (25%, 75%, single) as the within-subject factor. This analysis revealed a

significant effect of dominance ($F(2,10) = 7.28$, $p < 0.004$). Additional $t$-tests indicated that

the difference between the attentional conditions was significantly larger in the 25% than the

75% dominance condition and single-vernier condition ($t(10) = 7.72$, $p < 0.0016$, $t(10) =$

$6.74$, $p < 0.017$, respectively), but there was no significant difference between the 75% and

single vernier condition ($p = 0.288$). These various effects of attention allocation and

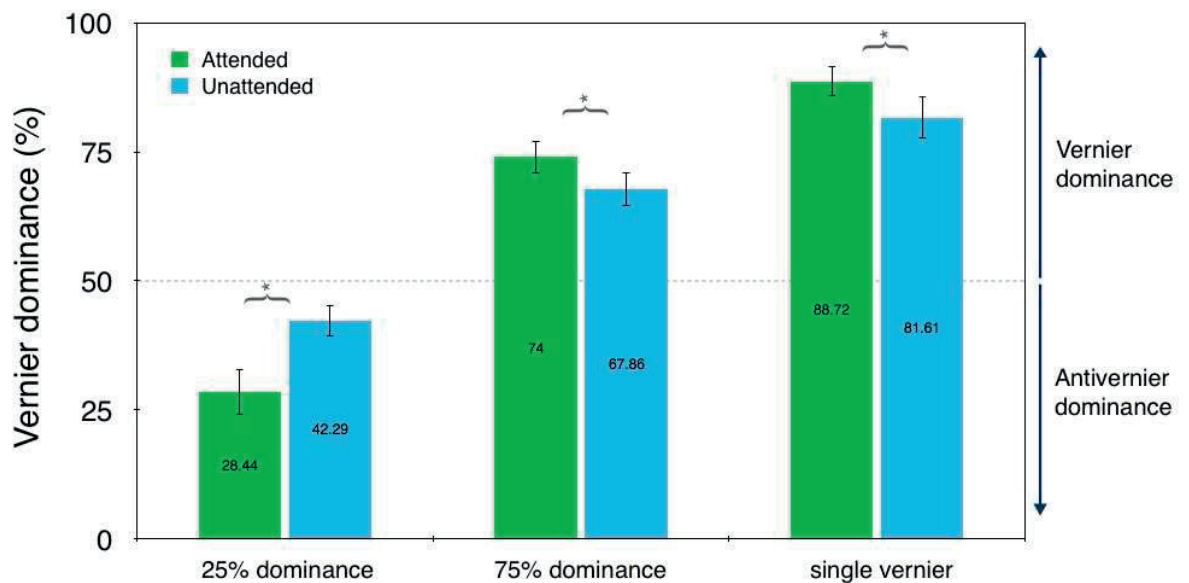dominance manipulations are discussed in detail in the following section.



Figure 3. Vernier dominance (percentage of trials in which the reported offset is in line with the first vernier) as a function of attentional and dominance conditions. Error bars correspond to one standard error of the mean.

**General Discussion**

This study examined the effects of covert spatial sustained attention on feature fusion. Experiments 1a and 1b measured the effect of attention on feature fusion with and without controlling for eye-movements. Experiment 2 examined whether the attentional effects found in Experiments 1a and 1b can be attributed to a signal enhancement mechanism. In all three experiments feature fusion was produced by presenting a sequence of vernier and antivernier stimuli, presented one after the other to the same location with no intervening temporal interval. Spatial sustained attention was manipulated by varying the degree of spatial uncertainty: In the attended condition the location of the stimulus sequence was fixed for the entire block, allowing attention allocation in advance to this location; in the unattended condition the location of the sequence varied randomly between two possible locations and therefore did not allow advance attention allocation to this location.

Attentional effects on feature fusion were found in all three experiments. In Experiments 1a and 1b we found a greater degree of antivernier dominance when participants could allocate attention in advance to the sequence location, and in Experiment 2 attention increased whatever dominance was already apparent in the unattended condition. These findings are the first demonstration that *sustained* spatial attention can affect the outcome of temporal feature fusion, and they are in line with a considerable number of studies demonstrating the effects of *covert* attention on early visual processing (e.g. Carrasco et al., 2002; Carrasco et al., 2004; Ling & Carrasco, 2006; Pestilli et al., 2007; Golla et al., 2004; Yeshurun & Carrasco, 1998;1999).

Different attentional mechanisms might account for the observed attentional modulation of feature fusion. One such attentional mechanism operates on decisional processes via modulation of the decision criteria the observers are adopting, or the decisional weights that are assigned to information gathered at the attended and unattended locations (e.g., Kinchla,

1980; Kinchla, Chen & Evert, 1995; Palmer, 1994; Shaw, 1984). This mechanism, however, cannot account for our current findings because it is only relevant when the attentional manipulation conveys useful information regarding the likelihood of the different possible behavioral responses (e.g., when one response is more likely in the attended than unattended condition). In all three experiments of this study the task was a two-alternative forced choice task, which included two equally likely responses regardless of the attentional manipulation, and hence it did not afford attentional effects on decision processes.

Another possible attentional mechanism operates through external or internal noise reduction (e.g., Dosher & Lu, 2000; Graham, Kramer & Haber, 1985; LaBerge, 1995; Lu & Dosher, 1998; Shiu & Pashler, 1994, 1995; Sperling & Dosher, 1986). The idea is that the advanced allocation of spatial attention to the relevant location allows observers to monitor only this location instead of monitoring all possible locations, thereby reducing the number of locations that needed to be monitored and accordingly the statistical noise introduced at these locations by irrelevant stimuli or internal noise. However, as with the previous attentional account, this noise-reduction mechanism cannot account for our findings. This is because in all three experiments of this study the task relevant stimuli (the vernier and antivernier) were suprathreshold and presented alone, without additional task-irrelevant items. When suprathreshold stimuli are employed, the internal noise generated at the other empty locations becomes negligible, and if there are no irrelevant items in the visual display there is also no external noise. Thus, given the displays employed in this study there was neither internal nor external noise to reduce and therefore attentional noise reduction does not appear to be involved in the attentional effects found here.

Ruling out these two possible mechanisms leaves us with a third attentional mechanism – signal enhancement (e.g., Bashinski & Bacharach, 1980; Carrasco & Yeshurun, 1998; Downing, 1988; Lu & Dosher, 1998; Posner, 1980). The hypothesis here is that allocating

attention to the relevant location enhances the quality of the sensory representation at the attended location, and evidence in support of this hypothesis was gathered in both behavioral studies (e.g., Bashinski & Bacharach, 1980; Downing, 1988; Lu & Dosher, 1998; Yeshurun & Carrasco, 1998; 1999) as well as neurophysiological studies (e.g., Brefczynski & DeYoe, 1999; Desimone & Ungerleider, 1989; Desimone, Wessinger, Thomas & Schneider, 1990; Martinez et al., 1999; Moran & Desimone, 1985; Motter, 1993; Spitzer, Desimone & Moran, 1988). For instance, single-cell recordings have demonstrated attentional modulations in neural responding of V1 and V4 cells (e.g., McAdams & Maunsell, 1999; Motter, 1993); neurons' responses to attended stimuli were stronger and more selective in both V4 (McAdams & Maunsell, 1999; Spitzer et al., 1988) and MT/MST (e.g., Treue & Maunsell, 1996). Similarly, fMRI studies have shown attentional facilitation in both striate and extrastriate visual cortex (e.g., Brefczynski & DeYoe, 1999; Martinez et al., 1999).

This hypothesis gains direct support in our study, particularly in Experiment 2. If attending the location at which the vernier and antivernier appear improves the sensory encoding of both stimuli, then the final, fused percept should also be strengthened, regardless of whether this final percept reflects vernier or antivernier dominance. Following this prediction, when the fused percept in the unattended condition was dominated by the antivernier offset (Experiments 1a, 1b, and 2 – 25% condition), attending the stimulus location resulted in an even stronger antivernier dominance. Likewise, when the fused percept in the unattended condition was dominated by the vernier offset (Experiment 2 – 75% condition), attending the sequence location resulted in an even stronger vernier dominance. These findings suggest that spatial sustained attention affects feature fusion via signal enhancement.

Also in line with the signal enhancement hypothesis is the finding that in the single vernier condition of Experiment 2, in which only a single stimulus was presented, vernier offset discrimination was better in the attended than unattended condition. This is consistent with

Yeshurun and Carrasco (1999) finding that directing transient spatial attention via peripheral cues to the vernier location improves vernier offset discrimination, and it suggests that sustained attention, like transient attention, can increase hyperacuity.

As mentioned above, Experiment 2 revealed another interesting finding. The effect of attention on feature fusion was more pronounced in the 25% dominance condition than the 75% and single vernier conditions. In fact, the effect size was almost twice as large in the 25% condition. This result might be due to an involvement of transient attention in addition to that of sustained attention. Transient attention operates in a stimulus-driven manner and it is typically triggered by a peripheral cue, abrupt onset or other rapid changes in the visual display. It is a fast form of attention with beneficial effects evident as early as 50 ms from the appearance of its trigger, though optimal effects require ~100 ms (e.g., Cheal & Lyon, 1991; Jonides, 1981; Muller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Posner, 1980; Remington, Johnston, & Yantis, 1992). It is possible, therefore, that the presentation of the first stimulus in our sequence – the vernier – attracted transient attention to its location. Indeed, it has been suggested that when a stream of stimuli is presented to the same location, like in a rapid serial visual presentation paradigm (RSVP), transient attention is triggered by the first stimulus and may assist the identification of the following stimulus (e.g., Bowman & Wyble, 2007). Although transient attention is attracted to the relevant location already by the onset of the first stimulus, the process of deploying attention takes time, and the second stimulus may disproportionally benefit from the deployment of transient attention (e.g., Bowman & Wyble, 2007; Wyble, Bowman, & Potter, 2009). Thus, due to the time required to deploy transient attention, if transient attention was indeed attracted to the sequence location by the onset of the vernier, the second stimulus in our sequence – the antivernier – might have been the one to benefit from the additional enhancement brought about by transient attention. Moreover, given the particularly brief presentation of each stimulus in the

sequence, and given the time required for the deployment of transient attention, this additional 'boost' of the second stimulus could only happen in the attended condition in which attention was already allocated to the sequence location allowing a faster deployment of transient attention. It is also probable that this disproportionate benefit in favor of the antivernier only took place in the 25% condition, since in the 75% condition the antivernier duration was halved leaving little room for transient attention to generate any effect, and the same logic holds for the single vernier condition. Taken together, we suggest that while in the 25% condition the antivernier processing was further facilitated by transient attention, in the 75% and single vernier conditions any attentional gain was restricted to sustained attention operating on both the vernier and antivernier representation.

Finally, another interesting finding that emerged in this study is that the effect of the spatial certainty manipulation was considerably smaller in Experiment 1b in comparison to Experiment 1a. Because the exclusion of eye-movements in Experiment 1b was the only difference between the two experiments, it is likely that the larger effect found in Experiment 1a was due to the fact that in the attended condition of this experiment, the participants could move their eyes to foveate the stimuli. This suggests that fusion is more prevalent in the fovea than in the periphery, which in turn strengthens the claim that feature fusion reflects temporal integration. This is because in comparison to foveal processing, the processing in the periphery exhibits shorter temporal integration (e.g., Swanson, Pan, & Lee, 2008). One reason for this difference between the central and peripheral regions is the distribution of receptive fields of parvocellular and magnocellular neurons. The fovea is characterized by a higher density of parvocellular neurons that decreases with increasing eccentricities, while magnocellular neurons are more prevalent at peripheral regions (e.g., Azzopardi, Jones & Cowey, 1999; Connolly & Van Essen, 1984; Merigan & Maunsell, 1993; Schiller & Logothetis, 1990). Many studies have demonstrated that parvocellular neurons have longer

response duration, slower decay, and longer temporal integration than magnocellular neurons (e.g., Derrington & Lennie, 1984; Maunsell, Nealey, & DePriest, 1990; Merigan & Maunsell, 1993; Schiller & Logothetis, 1990; Solomon, White, & Martin, 1999). Thus, the fact that temporal integration is indeed better in the fovea and that feature fusion also seems to be better in the fovea, supports the notion that feature fusion reflects temporal integration. However, in this study we did not compare directly the manifestation of feature fusion in the fovea and the periphery. A more direct comparison between the fovea and periphery is necessary to draw strong conclusions.

To summarize, this study is the first to examine the effects of spatial sustained attention on feature fusion, and it demonstrates that directing covert attention to the stimulus location affects feature fusion via signal enhancement. Specifically, we have found that attention strengthened the already dominating final fused percept, most likely by enhancing the representation of each of the fused stimuli.

# References

Azzopardi, P., Jones, K. E., & Cowey, A. (1999). Uneven mapping of magnocellular and parvocellular projections from the lateral geniculate nucleus to the striate cortex in macaque monkey. *Vision Research*, *39*, 2179–2189.

Bach, M. (1996). The "Freiburg visual acuity test." Automatic measure- ment of visual acuity. *Optometry and Vision Science, 73,* 49–53.

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception and Psychophysics*, *28*, 241–248.

Bowman H, Wyble B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review, 114*, 38–70

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience*, *2*, 370–374.

Briand, K. A. (1998). Feature integration and spatial attention: more evidence of a dissociation between endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1243-1256.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Cameron, E. L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, *42*, 949–967.

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research*, *40*, 1203–1215.

Carrasco, M., Williams, P. E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of Vision*, *2*, 467–479.

Carrasco, M., & Yeshurun, Y. (2009). Covert attention effects on spatial resolution. *Progress in Brain Research*, *176*, 65–86.

Chica, A.B., Bartolomeo, P., & Lupianez, J. (2013). Two cognitive and neural systems for endogenous and exogenous spatial attention. *Behavioural Brain Research*, *237*, 107-123.

Cheal, M., & Lyon, D. (1991). Central and peripheral precuing of forced-choice discrimination. *Quarterly Journal of Experimental Psychology*, *43A*, 859-880.

Connolly, M., & Van Essen, D. (1984). The representation of the visual field in parvocellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology*, *266*, 544–564.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 219- 240.

Desimone, R., & Ungerleider, L. G. (1989). Neural mechanism of visual processing in monkeys. In F. Boller, & J. Grafman, *Handbook of Neuropsychology* (pp. 267–299). Amsterdam: Elsevier BV.

Desimone, R., Wessinger, M., Thomas, L., & Schneider, W. (1990). Attentional control of visual perception: Cortical and subcortical mechanisms. *Cold Spring Harbor Symposia Quantitative Biology*, LV, 963–971.

Dosher, B. A., & Lu, L. (2000). Mechanisms of perceptual attention in precuing of location. *Vision Research, 40*(10–12), 1269–1292.

Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *Journal of Experimental Psychology*: *Human Perception Performance*, *14*, 188–202.

Efron, R. (1967). Duration of present. *Annals of the New York Academy of Sciences, 138*(A2), 713–729.

Efron, R. (1973). Conservation of temporal information by perceptual systems. *Perception & Psychophysics, 14,* 518–530.

Golla, H., Ignashchenkova, A., Haarmeier, T., & Their, P. (2004) Improvement of visual acuity by spatial cueing: a comparative study in human and non-human primates. *Vision Research.*, *44*(13): 1589–1600.

Graham, N., Kramer, P., & Haber, N. (1985). Attending to the spatial frequency and spatial position of near-threshold visual patterns. In M. I. Posner, & O. S. Marin (Eds.), *Attention and Performance XI* (pp. 269- 284). Hillsdale, NJ: Erlbaum.

Hein, E., Rolke, B., & Ulrich, R. (2006). Visual attention and temporal discrimination: Differential effects of automatic and voluntary cueing. *Visual Cognition*, 13, 29-50.

Hermens, F., Scharnowski, F., & Herzog, M. H. (2009). Spatial grouping determines temporal integration. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 595–610.

Hermens, F., Scharnowski, F., & Herzog, M. H. (2010). Automatic grouping of regular structures. *Journal of Vision*, *10*, 1–16.

Herzog, M. H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature, 415,* 433–436.

Herzog, M. H., Leseman, E., & Eurich, C. W. (2006). Spatial interactions determine temporal feature integration as revealed by unmasking. *Advances in Cognitive Psychology, 2,* 77–85.

Herzog, M. H., Fahle, M., & Koch, C. (2001). Spatial aspects of object formation revealed by a new illusion, shine-through. *Vision Research, 41,* 2325–2335.

Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences, USA, 98,* 4271–4275.

Herzog, M. H., Parish, L., Koch, C., & Fahle, M. (2003). Fusion of competing features is not serial. *Vision Research*, *43*, 1951–1960.

Herzog, M. H., Scharnowski, F., & Hermens, F. (2007). Long lasting effects of unmasking in a feature fusion paradigm. *Psychological Research*, *71*, 653–658.

Huang, L., & Dobkins, K.R. (2005) Attentional effects on contrast discrimination in humans: evidence for both contrast gain and response gain. *Vision Research*, *45*: 1201–1212.

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (pp. 187-204). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kinchla, R. A. (1980). The measurement of attention. In R. S. Nikerson (Eds.), *Attention and Performance IIX* (Vol. 8, pp. 213-238). Hillsdale, NJ: Erlbaum.

Kinchla, R. A., Chen, Z., & Evert, D. L. (1995). Precue effects in visual search: data or resource limited? *Perception and Psychophysics*, *57*(4), 441–450.

Klein, R. M. (1994). Perceptual-motor expectancies interact with covert visual orienting under conditions of endogenous but not exogenous control. *Canadian Journal of Experimental Psychology, 48*, 167-1 81.

LaBerge, D. (1995). *Attentional processing. The brain's art of mindfulness.* Cambridge, MA: Harvard University Press.

Lu, Z.-L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*(9), 1183–1198.

Lu, Z.-L., & Dosher, B. A. (2000). Spatial attention: Different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1534–1548.

Maunsell, J. H. R., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience, 10*, 3323-3334.

Martinez, A., Anllo-Vento, L., Sereno, M.I., Frank, L.R., Buxton, R.B., Dubowitz, D.J., Wong, E.C., Hinrichs, H., Heinze, H.J. & Hillyard, S.A. (1999) Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience, 2*(4), 364–369.

McAdams, C.J., & Maunsell, J.H. (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience, 19*(1), 431–441.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience, 16*, 369-402.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*, 782–784.

Motter, B.C. (1994) Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neurophysiology*, *14*(4), 2178–2189.

Muller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 315-330.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*, 1631-1646.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*, 1703–1721.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.

Raymond, J. E., Shapiro, K. L. & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849-860

Remington, R., Johnston, J. C., & Yantis, S. (1992). Attentional capture by abrupt onsets. *Perception and Psychophysics*, *51*, 279-290.

Rolke, B., Dinkelbach, A., Hein, E., & Ulrich, R. (2008). Does attention impair temporal discrimination? Examining non-attentional accounts. *Psychological Research*, *72*(1), 49-60.

Shaw, M. L. (1984). Division of attention among spatial locations: A fundamental difference between detection of letters and detection of luminance. In H. Bouma, & D. G. Bouwhuis (Eds), *Attention and Performance X* (pp. 109–120). Hillsdale, NJ: Erlbaum.

Scharnowski, F., Hermens, F., & Herzog, M. H. (2007). Bloch's law and the dynamics of feature fusion. *Vision Research*, *47*, 2444–2452.

Scharnowski, F., Hermens, F., Kammer, T., Ogmen, H., & Herzog, M. H. (2007). Feature Fusion Reveals Slow and Fast Visual Memories. *Journal of Cognitive Neuroscience*, *19*(4), 632–641.

Schiller, P. H., & Logothetis, N. K. (1990). The color-opponent and broad-band channels in the primate visual system. *Trends in Neuroscience*, *13*, 392-398.

Shiu, L., & Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology*: *Human Perception and Performance*, *20*(5), 1037–1054.

Shiu, L., & Pashler, H. (1995). Spatial attention and vernier acuity. *Vision Research*, *35*, 337–343.

Shapiro, K. L., Arnell, K. M. & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences*, *1*, 291-296.

Smith, P. L. (2000). Attention and luminance detection: Effects of cues, masks, and pedestals. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1401–1420.

Solomon, J.A. (2004) The effect of spatial cues on visual sensitivity. *Vision Research, 44*(12), 1209–1216.

Solomon, S. G., White, A. J., & Martin, P. R. (1999). Temporal contrast sensitivity in the lateral geniculate nucleus of a New World monkey, the marmoset Callithrix jacchus. *Journal of Physiology*, *517*(3), 907 917.

Sperling, G. & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of Perception and Human Performance* (Vol. 1, pp. 1–65). New York: Wiley.

Spitzer, H., Desimone, R., & Moran, J. (1988) Increased attention enhances both behavioural and neuronal performance. *Science*, *240*, 338–340.

Swanson, W. H., Pan, F., & Lee, B. B. (2008). Chromatic temporal integration and retinal eccentricity: Psychophysics, neurometric analysis and cortical pooling. *Vision Research*, *48*, 2657–2662.

Treue, S., & Maunsell, J. H. (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, *382*(6591), 539–541.

Wyble, B., Bowman, H., & Potter, M. C. (2009). Categorically defined targets trigger spatiotemporal attention. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 324–337.

Visser, T. A. W., & Enns, J. E. (2001). The role of attention in temporal integration. *Perception*, *30*, 135 145.

Yeshurun, Y. (2004). Isoluminant stimuli and red background attenuate the effects of transient spatial attention on temporal resolution. *Vision Research*, *44*, 1375-1387.

Yeshurun, Y., & Carrasco, M. (1998) Attention improves or impairs visual perception by enhancing spatial resolution. *Nature*, *396,* 72–75.

Yeshurun, Y., &Carrasco, M. (1999) Spatial attention im- proves performance in spatial resolution tasks. *Vision Research.*, *39*(2), 293–306.

Yeshurun, Y., & Carrasco, M. (2008). The effects of transient attention on spatial resolution and the size of the attentional cue. *Perception & Psychophysics*, *70*(1), 104–113.

Yeshurun, Y., & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience*, *3*, 622–627.

Yund, E. W., Morgan, H., & Efron, R. (1983). The micropattern effect and visible persistence. *Perception & Psychophysics*, *34*, 209–213.

# Appendix 5

Ballhausen*, N., Lauffs*, M. M., Herzog, M. H., & Kliegel, M. (*submitted*).
Investigating prospective memory via the eyes:
No evidence for a monitoring deficit in older adults.

Nicola Ballhausen[1*], Marc M. Lauffs[2*], Michael H. Herzog[2], Matthias Kliegel[1]

[1] Cognitive Aging Laboratory, Department of Psychology, University of Geneva, Switzerland
[2] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

* The authors contributed equally to this work.

# Investigating prospective memory via the eyes: No evidence for a monitoring deficit in older adults.

Ballhausen, N. [1,3,*,†], Lauffs, M. M. [2, *], Herzog, M. H. [2], & Kliegel, M. [1,3,4]

1 Department of Psychology, University of Geneva, Geneva, Switzerland.

2 Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

3 Center for the Interdisciplinary Study of Gerontology and Vulnerability, University of Geneva, Geneva, Switzerland.

4 Swiss National Center of Competences in Research LIVES–Overcoming vulnerability: life course perspectives, Lausanne and Geneva, Switzerland.

* These authors contributed equally to this work.

† Corresponding Author:

Nicola Ballhausen

University of Geneva

Boulevard du Pont d'Arve 40

CH- 1211 Genève 4

Phone: +41 22 379 89 31

Fax: +41 22 379 92 29

E-mail: nicola.ballhausen@unige.ch

## Abstract

Prospective memory (PM) refers to remembering to perform intended actions in the future. Older adults were found to show particularly reduced PM performance in tasks relying on strategic monitoring (i.e., non-focal tasks), and it was suggested that a deficit in monitoring would underlie these findings.

The present study, for the first time, set out to investigate focality effects on PM and monitoring behavior in younger and older adults using an eye tracking paradigm. While only older adults showed reduced PM performance in a non-focal task, no difference was found in the proportions of eye movements to the PM target region between age groups and focality conditions. Comparing the proportion of correctly remembered PM cues out of the PM target trials that were overtly monitored revealed that - comparably as for overall PM performance – in the non-focal condition fewer PM trials were correctly remembered than in the focal one in older adults.

The results suggest that focality effects in older adults cannot be explained by reduced overt monitoring behavior. Thus, difficulties in subsequent processes rather than a monitoring deficit may relate to reduced PM performance in non-focal tasks in older adults, as monitoring on its own did not lead to remembering.

**Investigating prospective memory via the eyes: No evidence for a monitoring deficit in**

**older adults.**

Prospective memory (PM) refers to remembering to do something at a particular

moment (or time period) in the future (Kliegel et al., 2016; McDaniel & Einstein, 2007). Typical

PM tasks are remembering to post a letter on the way home, attaching a file to an email

before sending it off, to take medication according to a schedule, or to withdraw money when

passing by an ATM. Those tasks can be remembered in a rather spontaneous way, i.e., the

intention pops into our mind without attentional resources being recruited to detect the right

moment to perform the intended action. For instance, when passing the supermarket, one

might just remember that one actually wanted to buy bread. On the contrary, these tasks can

be accomplished by strategically monitoring the environment for the right moment to

perform the intended action to arise (Einstein et al., 2005; McDaniel & Einstein, 2000). This

might for example be the case when one has to post ones job application, one makes sure to

not miss the letter box.  The latter processing style will be the focus of the present paper.

Monitoring processes, comprising the voluntarily deployment of attentional

resources, are a key element of the main theories for PM (e.g., multiprocess framework,

McDaniel & Einstein, 2000; preparatory attentional processes and memory processes theory,

PAM, Smith, 2003). The two-process model of strategic monitoring (Guynn, 2003), for

example, specifically identified two components of monitoring, namely the *retrieval mode*

and *target checking*. While retrieval mode conceptualizes a stable mental preparedness to

consider stimuli as target cues to retrieve intentions, target checking refers to active, periodic

checks of the environment for the PM target cue (Guynn, 2008). In line with the periodic

nature, Ball, Brewer, Loft, and Bowden (2015) proposed transient, rather than continuous

monitoring processes, to underlie cue detection. Similarly, the dynamic multiprocess

framework (Scullin, McDaniel, & Shelton, 2013) suggests that spontaneous retrieval and monitoring are utilized in a dynamic way, with individuals engaging in monitoring when they expect PM cues. This further is consistent with changes in attentional resource allocations when a certain context prompts the appearance of the PM target cue (Cohen, Gordon, Jaudas, Hefer, & Dreisbach, 2016; Kuhlmann & Rummel, 2014; Marsh, Cook, & Hicks, 2006).

Conceptually, models agree that a PM task should be particularly difficult if it requires strategic processing to detect the target cue. This typically is the case if the ongoing task does not require the processing of the features that define the PM target cue (e.g., while naming famous faces, responding to faces with eyeglasses, see Einstein & McDaniel, 2005). Those so-called non-focal PM tasks have been associated with lower PM performance (e.g., Einstein et al., 2005). Compared to focal PM tasks in which the defining features of the PM cue are processed during the ongoing task (in the aforementioned naming famous faces example, responding to faces of people named "John"), non-focal PM tasks show larger costs on the ongoing task. This could be expressed in a reduction in accuracy and/or an increase in response times (see for instance Cona, Bisiacchi, & Moscovitch, 2013). Particularly in experimental conditions that discouraged sustained monitoring, PM performance dropped significantly (Scullin, McDaniel, Shelton, & Lee, 2010). Consequently, lower PM performance for non-focal as compared to focal cues have been attributed to problems in monitoring, namely that subjects might forget to monitor or might not have enough attentional resources at their disposal to monitor for the target cue (Rose, Rendell, McDaniel, Aberle, & Kliegel, 2010; see also Smith, 2003).

This key role of monitoring processes in explaining PM performance has been transferred to the developmental perspective explaining age differences. Especially older adults have been shown to be negatively affected by non-focal tasks (for meta-analysis, see

Kliegel, Jäger, & Phillips, 2008). Attempts to explain these age differences in PM performance suggested that as older adults generally have less attentional resources available (Craik, 1986), they might have especial difficulties in recruiting attentional resources to monitor for the PM target cue. This monitoring impairment again would be notably detrimental for tasks relying on strategic monitoring such as non-focal tasks (McDaniel & Einstein, 2011).

Several studies have confirmed that older adults' performance suffers particularly in PM tasks that are thought to require high levels of strategic monitoring (e.g., non-focal tasks, Ihle, Hering, Mahy, Bisiacchi, & Kliegel, 2013; non salient cues, Altgassen, Phillips, Henry, Rendell, & Kliegel, 2010; high ongoing task absorption, Smith, Horn, & Bayen, 2012). However, only a few studies have so far *directly* compared monitoring processes between younger and older adults and have produced somewhat conflicting results. For example, it has been demonstrated that older adults show larger costs in the ongoing task than younger (Henry et al., 2015), what, when considering larger costs as a sign of greater recruitment of attentional resources, can be seen as more monitoring behavior in the older adults (note that these differences were not detectable when adjusting for general slowing). Furthermore, Reese and Cherry (2002) could show that younger adults have more task-irrelevant thoughts than older adults while performing a PM task. This difference was not related to actual PM performance, where younger and older adults performed equally. This finding is in line with the idea that younger adults have more attentional capacity than older adults. Reese-Melancon (2013) demonstrated that while performing a PM task, most individuals think about the ongoing task only, but that on the contrary thoughts about the PM task are linked to successful PM performance. This link was particularly strong in older adults performing a non-focal task, indicating that the self-initiation of monitoring processes signifies successfully coping with the main challenge of those PM tasks. The study further suggests that younger adults monitor

more often for the target cue than older adults do, consequently resulting in better PM performance. However, it was shown that both younger and older adults could similarly make use of the context (e.g., information provided about when the PM target cues would appear) to adjust their monitoring behavior (Kominsky & Reese-Melancon, 2017). Taken together it seems unclear as to whether there are quantitative differences between younger and older adults in their monitoring behavior. Results only partly support the idea that older adults' monitoring is impaired. Beyond that, the underlying mechanisms of possible (age-related) monitoring differences are also under debate. Currently, one of the key hypothesis in this regard is that age-related differences in overt PM cue monitoring relate to differences in working memory capacity (Bisiacchi, Tarantino, & Ciccola, 2008; Rose et al., 2010), with more working memory resources enabling individuals to monitor more often and more strategically while successfully processing the ongoing task.

Given that empirical evidence is scant, it is still an open question whether older adults just completely forget about the intention and thus do not monitor, or whether they still have the intended action in mind and monitor, but that this does not enable them to achieve a level of PM performance comparable to younger adults. If monitoring does not result in remembering, subsequent processes might be negatively affected by age. For example, although having monitored, older adults may less often realize the significance of a PM target cue compared to non-target ongoing task cues. In a recent study by Hering et al. (2016) it was shown that older adults, as compared to younger adults, indeed displayed problems realizing the relevance of the PM target cues. Specifically, their limited attentional sensitivity might be accountable for a lowered discrimination between the PM and ongoing task. Additionally, increasing response coordination demands at the post-retrieval phase was shown to relate to larger age effects (Ihle, Hering, Mahy, Bisiacchi, & Kliegel, 2013).

Besides the conceptual discussion on potential differences in monitoring between younger and older adults, present studies investigating monitoring face some methodological problems. While studies analyzing self-reported thoughts towards either the PM or the ongoing task (e.g., Cherry & LeCompte, 1999) are limited by a subjective reporting bias (e.g., post-test questionnaire with self-evaluation of how many times they thought about the PM task is affected by participant's own definition of 'a thought about the PM task', (meta-)memory, self-awareness), most studies consider costs on the ongoing task (regarding accuracy or response times) as an indicator of monitoring behavior (e.g., Einstein et al., 2005). Yet, costs in response times might not be the best indicator for monitoring behavior. A recent study of Ballhausen, Schnitzspahn, Horn, and Kliegel (2017) for example demonstrated that a high cue maintenance load results in enhanced ongoing task costs. Thus, even if costs are high, this does not necessarily mean that individuals monitor for the PM target cue. Particularly when comparing younger and older adults, general slowing in older age might be partly responsible for greater ongoing task response times of that group rather than monitoring differences. Henry et al. (2015), for instance, showed that when accounting for general slowing, ongoing task cost differences between younger and older adults did not remain significant anymore. Analyzing ongoing task cost differences thus might not necessarily capture differences in monitoring behavior.

To address these methodological challenges, we developed a paradigm in which PM target monitoring requires the observer to perform an overt, observable behavior: An eye movement. Here, we propose, for the first time, to use eye tracking to directly quantify and compare this monitoring behavior in both older and younger adults. Only very few studies have so far used eye tracking methods to capture PM monitoring behavior, and none of them investigated older adults. As opposed to indirect measures of monitoring (e.g., ongoing task

costs), in eye tracking paradigms we can study a behavior that only occurs when monitoring

takes place (i.e., a specific eye movement), which allows to quantify monitoring behavior. In

the following, all present studies of PM and eye tracking are presented, thereby highlighting

the results that are specific to eye tracking data.

Comparing the eye movement patterns in a PM task with free viewing and visual

search instructions, Hartwig, Schnitzspahn, Kliegel, Velichkovsky, and Helmert (2013)

demonstrated that the gaze behavior differed more from free viewing trials (shorter fixations)

and became more similar to visual search, the more successfully a PM task was solved. Dwell

time on PM targets that were missed was shorter than for those that were remembered.

Specifically, PM hits and misses did not differ regarding fixation duration and saccadic

amplitude, but misses were fixated less often (although still at 50% of all trials) than hits.

Combining eye tracking with virtual reality, it was demonstrated that before target detection,

saccades are longer than after target detection (Kalpouzos, Eriksson, Sjolie, Molin, & Nyberg,

2010), reflecting exploratory visual-search behavior as also shown by Hartwig et al. (2013). In

a different visual search task, participants had to look for specific letters and additionally had

to respond to predefined letters as PM target cue (West, Carlson, & Cohen, 2007). An analysis

of the fixation times revealed that PM target cues were fixated longer, both at the first fixation

and in total. In their study, the number of fixations did not differ between PM target and non-

target items, which is likely to relate to the nature of the visual search task. In line with

Hartwig et al. (2013), the authors demonstrated that correctly reported PM targets were

fixated more often and longer than those that were missed, although missed targets were

still fixated in about 50% of the trials. Together, these studies corroborate the idea that PM

intentions alter the eye movement patterns, reflecting engagement in strategic monitoring,

and failures in this engagement can account for lower PM performance.

Three further studies have used eye tracking to more directly investigate alterations of PM monitoring behavior in response to environment and context. Pictures that were either semantically related to the PM target picture or not were displayed within an ongoing visual search task (Shelton & Christopher, 2016). The PM target area was outside the observer's field of view on the upper-right corner of the screen and different pictures were presented at that location three times faster than the pictures in the ongoing task. It was shown that pictures that were semantically related to the target cue enhanced the number of fixations in the PM target cue region as a proxy of strategic monitoring. The total number of fixations correlated positively with PM performance. Further, to investigate context-sensitive PM monitoring, subjects were instructed that the PM target cue would appear at a certain trial number (Bowden, Smith, & Loft, 2017). Accordingly, subjects in that condition, as opposed to controls, fixated the trial number more often the more the relevant context was approached. Again, the number of checks related positively to PM accuracy.

While all five aforementioned eye tracking studies focused on younger healthy adults only, Chen, Zhou, Cui, and Chen (2013) compared healthy and depressed younger adults' performance in a PM task. Healthy controls showed no difference between ongoing and PM trials in both the number of fixations and their durations. However, participants with depression fixated both ongoing task and PM targets more often and longer than healthy controls, but still displayed lower PM performance. According to the authors this suggests that performing a PM task required a greater cognitive effort for the depressed than for the healthy adults.

All in all, it was shown that – at least in healthy younger adults – fewer (and shorter) fixations in the PM target area were linked to lower PM performance. This indicates that the PM target was more frequently reported when more attention was allocated to the target

area. Thus, the usage of eye tracking may provide novel insights into the debate on possible age differences in PM monitoring in younger versus older adults. We hypothesized that if an intention is presently kept active in working memory, subjects should monitor intentionally and thus show regular eye movements to the target area. This allowed us to test (a) whether older adults generally show lower levels of overt monitoring for PM cues, (b) whether age-related overt monitoring is related to PM cue features such as focality and (c) whether differences in overt PM monitoring are related to differences in PM performance.

## Methods

### Participants

Thirty younger and 46 older subjects were tested. The study was approved by the Human Research Ethics Committee of the University of Geneva. Subjects were recruited via flyers and, in case of the older adults, advertisement in local senior citizen organizations and community centers. A monetary reimbursement of CHF 15 was provided. Exclusion criteria were neurological and cognitive impairments, the use of psychotropic drugs, non-corrected vision impairments (cataract, strabismus, glass eye) and low visual acuity, as indicated by a score below 0.8 in the Freiburg Visual Acuity Test (Bach, 1996). Older adults were pre-screened for these criteria with the help of the French version of the modified telephone interview for cognitive status (F-TICS-m; Vercambre et al., 2010; Welsh, Breitner, & Magruderhabib, 1993). Only participants who did not meet the exclusion criteria and who scored at least 28 points in the F-TICS were included in the study.

Participants with inaccurate retrospective memory for task instructions (two older subjects) and inaccurate color perception (one older subject), as well as participants whose accuracy in the baseline ongoing task block was statistically indistinguishable (≥ 60%) from chance level (two older subjects) were excluded from the analyses. Furthermore, five older

subjects were excluded because we could not obtain a valid eye tracking calibration. For another five older subjects, the calibration allowed for measurement, but the resulting data quality was too low to qualify for analysis (more than one condition with less than 66% of trials or less than 4 PM trials retained). The remaining sample consisted of 31 older ($M_{age}$ = 69.9, $SD$ = 5.2, 22 female) and 30 younger ($M_{age}$ = 22.6, $SD$ = 2.7, 18 female) adults[1]. Comparing the two age groups, younger adults showed higher scores on the Balance subscale of the Wechsler intelligence scale (Wechsler, 2011), a proxy for fluid intelligence abilities ($t$(59) = 5.26), $p$ < .01), whereas older adults outperformed younger adults ($t$(59) = 2.75, $p$ < .01) on the Mill Hill vocabulary test (Deltour, 1993).

**Apparatus**

The experiment was programmed in Matlab with Psychtoolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997). Eye tracking data were recorded using an ET1000 eye tracker from TheEyeTribe (Dalmaijer, 2014). The eye tracker was recalibrated before every block and controlled from within Matlab using in-house software. Stimuli were presented on a BenQ XL2420z LCD monitor (1920 x 1080 px, 60 Hz). Participants were positioned in a chin rest and viewed the stimuli from a distance of 66cm, at which the monitor subtended 46° x 26° [degrees of visual angle].

**Stimuli**

Twenty-seven colored capital letters, grouped as three 3-by-3 matrices, were presented simultaneously in the center of the screen and 16.8° to the left and right periphery (see Figure 1; center-to-center distance). In each of these matrices, a central letter was surrounded by eight letters, therefore inducing *crowding* of that letter: The central letter and

---

[1] Note that this final sample was not different from the original one with regard to age, verbal knowledge, and fluid intelligence.

its color can only be identified when looked at directly, but not when seen peripherally (e.g., Levi, 2008; Whitney & Levi, 2011). Hence, when focusing on the foveated central letter, subjects were forced to first make an eye movement in order to monitor a peripheral central letter.

The stimulus specifics were randomly chosen with the same restrictions in all blocks and screen positions (left, center, right). In ongoing task trials (see below), central letters of the 3-by-3 matrices were chosen randomly from the letters 'A', 'E', 'I', 'O', 'L', 'N', 'S', 'T', with the restriction that half were consonants and half were vowels. The color of the central letters was randomly selected from eight colors (not black or white). In PM target trials, the PM target cue property (depending on condition, 'U', 'R', white or black letter in the central or right matrix) was presented at random at that central position, with the restriction that they did not appear during the first three trials of each block and that at least three ongoing task trials were presented between two occurrences.

The surrounding letters were chosen randomly from the aforementioned eight letters for each trial, with the restriction that in each trial half were consonants and half were vowels, two identical letters did not neighbor each other, and the central letter did not also appear in the surrounding letters. Similarly, for each trial the color of the surrounding letters was chosen randomly from the eight colors (not black or white) with the same restrictions, i.e., two letters of the same color did not neighbor each other and the color of the central letter did not also appear in the surrounding letters. Stimuli were presented on a midlevel gray background. Each letter was drawn centrally in an invisible 0.5-by-0.5° square, with a gap of 0.05° between squares.

Figure 1. Stimulus configuration for each trial. Note that the circles around the 3x3 letter matrices were not presented but are for illustrative purposes only.


**Task**

The *ongoing task* was to indicate whether the central letter in the screen center was a consonant or vowel (two-alternative forced choice) by using the left and right arrow keys of a generic computer keyboard. To ensure that this letter was foveated at trial onset, the trial only started after a steady fixation of at least 250ms was detected on a central fixation point (red, 0.1° diameter), in the same position where then the letter would appear. Both accuracy and speed were emphasized to the subjects. To ensure that the ongoing task was well understood, participants performed a training of 10 trials with auditory feedback. When performance was below 80%, the participant was reinstructed and the training repeated until performance was above 80%. Following training, each subject first performed a block of trials with only the ongoing task. The order of all remaining blocks was chosen randomly for each subject. Each block comprised 52 trials.

For the *PM task*, participants were instructed to complete an additional task, which consisted of remembering to press the space bar whenever one of two pre-defined cues

would appear. Participants were instructed to do so before responding to the ongoing task, but that the two tasks would be equally important. Targets were infrequent and occurred in 6 out of 52 trials (11.5%). Each participant performed four different PM task blocks that manipulated the focality of the target. In that context, we explored different non-focality manipulations: Local non-focality (i.e., PM target cues being spatially presented in the periphery of the ongoing task items) and content-wise non-focality (i.e., PM target cues not part of the information being treated during the ongoing task in terms of content). The resulting four focality conditions differed in the target cues the subjects had to respond to, thereby altering the monitoring requirements:

1) The initially foveated letter is a 'U' or 'R'. This condition can be seen as the most focal condition[2], as the PM target cue highly overlaps with the ongoing task with regard to location and information.

2) The central letter in the right periphery is a 'U' or 'R'. While the information treated in the course of the ongoing task overlaps with the PM target cue, the location is different.

3) The initially foveated letter is black or white. While the location treated in the course of the ongoing task overlaps with the PM target cue, the information is different.

4) The central letter in the right periphery is black or white. In this "highly non-focal" condition, both location and information of the PM target cue do not overlap with the ongoing task.

---

[2] Note that we did not aim at testing highly focal tasks that only rely on spontaneous retrieval.

While in condition 1 and 3 the PM target location was the same as in ongoing task trials, in condition 2 and 4 monitoring of the target location required the subject to make an eye movement.

A trial ended when a response for the *ongoing* task was registered and a blank-screen of 1 sec was presented as inter-trial interval before the next trial started with the presentation of a fixation point. If no response was registered during 5 secs after stimulus onset, the program moved on to the next trial. In case of PM task responses (space bar), the program waited another 5 secs for an ongoing task response[3]. The eye tracker was recalibrated before each block.

**Procedure**

After providing informed consent, subjects were screened on their visual acuity using the Freiburg Visual Acuity test (Bach, 1996) as well as their color vision using the Ishihara's test for color-blindness (Ishihara, 1992). Participants then received instructions for the ongoing task, followed by a training run. A test block of the ongoing task followed, in which eye movements were recorded. In the following, the four PM blocks were presented in randomized order. For each of them, instructions were provided, followed by a break in which one of the background tests was performed (first break: socio-demographic questionnaire, second break: Mill Hill vocabulary test, third and fourth break: Balances test). Then, the PM block started without repeating the PM instructions, again recording eye movements. After completing all test blocks, a questionnaire asking for retrospective memory for the PM task was administered.

---

[3] Note that the 5 secs response interval left enough time for all subjects to respond: Over the whole experiment only 2 response timeouts occurred and those were produced by a younger subject. By inference, this means that all other trials ended with a response to the ongoing task.

**Eye tracking data analysis**

First, samples with signal loss, i.e., for which the eye tracker reported no pupil size or no gaze position that lies within the screen bounds, were identified. All trials in which more than one third of the recorded samples were lost were discarded and subjects with excessive trial loss (more than one condition with less than 66% of trials or less than 4 PM trials retained) were excluded. Per included participant the average loss of trials was 2.75% per condition (*SD* = 6.5%; range 0-52%). On average, 4.49% (*SD* = 5.6%; range 0-34%) of samples were lost per subject and condition. On average, 5.8 of the 6 PM targets per block were retained for analysis (*SD* = 0.61; range 2-6).

Data quality was more variable and significantly worse in older than younger subjects. Older subjects lost more samples per trial on average ($F(1,59)$ = 6.09, *p* = .017; mean difference = 2.8%, *SEM* = 1.1%) and lost more trials due to excessive data loss ($F(1,59)$ = 6.12, *p* = .016; mean difference = 3.0%, *SEM* = 1.2%). As a result, fewer PM trials were available for analysis in older subjects ($F(1,59)$ = 5.46, *p* = .023; mean difference = 2.1, *SEM* = 0.089).

Further, the exact gaze position to regions-of-interest (ROI) and detected fixations were abstracted. We designated a 2.5° radius circle around the center of each group of letters as ROI, and all remaining parts of the screen as a fourth ROI. If the gaze was directed at one of the ROIs for more than 50 ms (3 consecutive samples) this was counted as a fixation. The spatial and temporal criterion for fixations (2.5°, 50 ms) were chosen after inspection of the raw data. The criteria can be quite liberal, because the a priori probability of the gaze falling into the left and right ROI is extremely small by design, due to the small ROI size and large distance from the screen center. Supporting this, fixations outside of the letter ROIs were extremely rare and inspection of the raw data indicates that the gaze position typically alternated between one central letter and another.

## Results

**PM Performance**

To analyze PM performance, the proportion of correctly remembered PM responses (hits) was calculated for each experimental condition (see Figure 2). As the data distribution did not follow the normality assumption (tested with Kolmogorov-Smirnov test, P-P plot as well as great absolute values of skewness and kurtosis), Friedman's ANOVA was used instead of the originally intended mixed ANOVA. Across participants, there was an effect of task condition, $\chi^2(3) = 8.19$, $p = .04$. To further investigate that effect, younger and older adults' data were analyzed separately. While in younger adults PM accuracy was the same in all four conditions, $\chi^2(3) = 1.63$, $p = .65$, older adults showed different accuracy in the four task conditions, $\chi^2(3) = 11.71$, $p < .01$. Post hoc tests (Dunn-Bonferroni-Tests) showed that PM performance was only significantly different between condition 1 and 4, $z = 2.66$, $p < .05$.
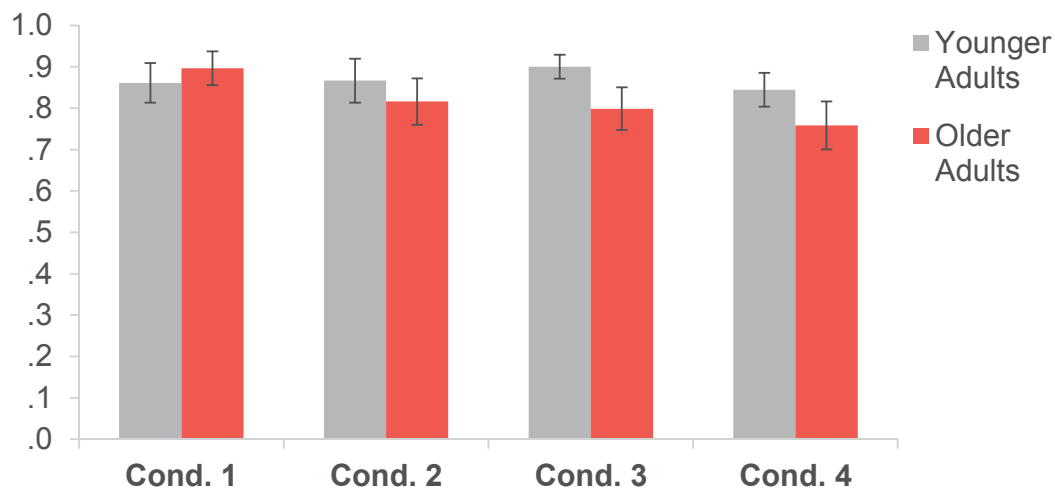


Figure 2. Average PM Accuracy (Proportions of Hits). Error bars represent +/- 1 SE.

**Ongoing Task Response Time Costs**

For each PM condition, every participant's mean response time in the block of only the ongoing task was subtracted from the mean response time of the ongoing task trials in the PM block. The resulting ongoing task response time costs reflect the cost of adding a PM task to an ongoing task (see Table 1). Following previous research, the first two trials of each block, PM target trials and the two trials following each PM target, as well as incorrect ongoing task trials were omitted from analysis (for similar procedure, see Boywitt & Rummel, 2012; Brewer, 2011; Cohen et al., 2016; Horn, Bayen, & Smith, 2013). A mixed-factorial ANOVA indicated a significant main effect of age, $F(1, 59) = 5.48$, $p = .02$, $\eta_p^2 = .09$, reflecting higher ongoing task response time costs for older compared with younger adults. Further, a main effect of task condition was observed, $F(3, 177) = 276.22$, $p < .001$, $\eta_p^2 = .82$. To investigate which conditions differed from each-other, pairwise comparisons between the specific conditions were used as post hoc tests. These were corrected for multiple testing using the Bonferroni correction by multiplying the p-values of each post hoc test by the number of conditions (Maxwell & Delaney, 2004). Pairwise comparisons revealed that while there was a significant difference between condition 1 and both 2 and 4, as well as 3 and both 2 and 4 ($p$s < .01), there was no difference between condition 1 and 3 ($p = .89$) as well as 2 and 4 ($p = .18$). The interaction of age and the task conditions was not significant either ($p = .39$).

Table 1

*Means and Standard Deviations of Eye Tracking Proportions of Ongoing and PM Trials, and*

*Proportion of Correct Responses to PM target Trials out of the Target Trials Monitored for*

*Younger (N = 30) and Older Adults (N = 31).*

| | Focal | Locally Non-Focal | Content-Wise Non-Focal | Highly Non-Focal (Location & Content) |
|---|---|---|---|---|
| | M (SD) | M (SD) | M (SD) | M (SD) |
| *OT Response Time Costs* | | | | |
| Younger Adults | 270 (170) | 1044 (272) | 254 (194) | 986 (325) |
| Older Adults | 376 (278) | 1231 (364) | 314 (225) | 1138 (355) |
| *Proportion of OT trials Monitored[a]* | | | | |
| Younger Adults | 1 (.00) | .97 (.07) | 1 (.00) | .99 (.02) |
| Older Adults | 1 (.00) | .95 (.18) | 1 (.00) | .97 (.08) |
| *Proportion of PM Target Trials Monitored[a]* | | | | |
| Younger Adults | 1 (.00) | .95 (.12) | 1 (.00) | .98 (.06) |
| Older Adults | 1 (.00) | .96 (.18) | 1 (.00) | .97 (.07) |
| *Correct PM Responses to Monitored Trials* | | | | |
| Younger Adults | .87 (.27) | .86 (.30) | .90 (.16) | .86 (.22) |
| Older Adults | .92 (.19) | .85 (.28) | .85 (.22) | .79 (.27) |

[a] Note that for the focal and the content-wise non-focal condition, monitoring is at 100% as PM target trials were

presented foveally.

**Monitoring data**

To explore whether older adults' lower performance in condition 4 was caused by reduced monitoring behavior, eye movements to the right ROI were examined. Descriptive data can be found in Table 1. More precisely, the proportion of trials in which at least one fixation in the right ROI was detected was analyzed using mixed-factorial ANOVAs. Analyzing monitoring in the ongoing task trials revealed neither an effect of age, nor of condition, nor an interaction ($ps \geq .29$). Similarly, conducting the analysis on only PM target trials did not show any difference with regard to age, condition or their interaction ($ps \geq .39$). Both findings suggest that differences in monitoring do not account for older adults' reduced PM performance in condition 4.

Further analyses investigated whether the reduced PM performance in condition 4 in the older adults were rather caused by processes subsequent to target monitoring. Firstly, the proportion of PM target trials that were monitored (note, however, that monitoring in condition 1 and 3 was at 100% since the targets were presented foveally) was compared with the actual performance in that condition. With only one exception (younger adults in condition 2, $p = .08$), actual PM performance was lower than the proportion of PM target trials that were monitored in all conditions in both younger and older adults ($ps \leq .01$). Consequently, the proportion of correctly reported PM cues out of only the trials that were monitored was analyzed using Friedman's ANOVA, as data again did not fulfil the normality assumption. In younger adults, no difference was found between conditions, $\chi^2(3) = 0.80$, $p = .85$. However, in older adults, the four conditions differed, $\chi^2(3) = 13.39$, $p < .01$. Post hoc Dunn-Bonferroni-Tests revealed that PM performance out of the trials that were monitored was only significantly different between condition 1 and 4, $z = 3.00$, $p = .02$, but not between any of the other pairwise comparisons ($ps \geq .46$).

An additional analysis aimed to quantify how many of the missed target cues had been monitored. Only the two conditions requiring eye movements to the right were taken into account. On average, older participants that showed misses had previously looked at 83.3% ($SD$ = 38.9%) of the cues they missed in condition 2 and 97.8% ($SD$ = 6.5%) in condition 4, compared to 89.6% ($SD$ = 19.8%) and 89.6% ($SD$ = 19.8%) in younger participants. Hence, in both group the large majority of misses occurred even though the participants had monitored for and looked at the targets. These values are lower-bound estimates, because our analysis might have missed fixations due to data-quality problems.

**Discussion**

The present study set out to investigate the effect of focality on PM performance in younger and older adults in one focal and three differently non-focal tasks. To further understand the effect of focality on PM, we studied whether a monitoring deficit underlies reduced PM performance in older adults. For the first time in older adults, we studied eye movements to the PM target region. We found an effect of focality on PM performance only in older adults and only in the "highly non-focal" condition, in which the PM target cue did not overlap with the ongoing task regarding both content and location. Yet, monitoring proportions (i.e., proportions of eye movements to the PM target region) did not differ between conditions in both age groups neither in all ongoing nor in only the PM trials. Comparing the proportion of correctly remembered PM cues out of the target trials that were monitored revealed that - comparable to overall PM performance – only in the highly non-focal condition fewer PM trials were correctly remembered than in the focal one.

First of all, the result of lower PM performance in the highly non-focal compared to the focal condition in older, but not younger adults, is in line with meta-analytical findings of greater age differences in non-focal compared to focal tasks (see Ihle et al., 2013; Kliegel et

al., 2008). However, no effect of focality was found when only one of the dimensions, either content or location, was manipulated in both age groups. Similar results, although in younger adults only, were obtained by Cona et al. (2013), who also could not show lower PM performance for non-focal compared to focal tasks. It seems that even though there is no overlap of the PM target cue with the processing of the ongoing task, which would be considered a non-focal task, PM performance is not necessarily reduced, not even in older adults.

The result might be related to our rather easy ongoing task which by deciding between vowels and consonants might not have absorbed many attentional resources. Consequently, enough resources might have remained available for performing the PM task. This is in line with Kliegel, Altgassen, Hering, and Rose (2011) who have argued that age or neuro-clinical effects may only emerge if there is a mismatch between the available attentional capacity in a given individual and the level and complexity required by the given task at hand. However, if neither content nor location of the PM target cue overlapped with the ongoing task, older adults were found to show reduced performance compared to the focal condition. This suggests that task demands might have exceeded the - compared to younger adults reduced - amount of resources available in older adults (Craik, 1986). Future studies are required to examine whether the pattern of results also holds true for e.g. more absorbing ongoing tasks, or whether also less demanding non-focality manipulations would affect PM performance when paired with more absorbing ongoing tasks. From a methodological point of view, these findings would also have implications for future non-focality manipulations. Additionally, at least in our paradigm no difference was obtained between non-focality manipulations that varied either location or content (see Kliegel et al., 2013 for another example varying cue

location and content in children). It would be of methodological importance to see if that would also be true in case of more demanding ongoing tasks.

The main goal of the present study was to examine overt monitoring in PM tasks in younger and older adults and to see how different pattern of monitoring might relate to PM performance differences caused by age and focality. Apart from the classical measuring of monitoring by ongoing task costs, we additionally measured monitoring by the proportion of eye movements to the respective target region (see also Shelton & Christopher, 2016). In general, monitoring proportions in the present study were very high and at a similar level in both younger and older adults. This was true for ongoing task trials and specifically for the PM target trials (both ≥ 95%). In contrast to Reese-Melancon (2013), our results argue against a general monitoring deficit in older adults. In none of the two age groups, monitoring proportions differed between conditions, suggesting that age-related lower PM performance in the highly non-focal condition could not be explained by age-related reduced monitoring in that condition (contrasting suggestions from McDaniel & Einstein, 2011).

To further corroborate this conceptually important conclusion, the proportion of the target trials that were monitored was compared with the proportion of PM target trials that were correctly remembered. Differences between these two measures appeared in almost all conditions. This suggests a conceptually important conclusion, namely that even successful overt monitoring for a target cue does not automatically lead to correctly remembering it. In fact, our analyses showed that at least 83% of the PM target cues that both young and old subjects failed to report had previously been looked at and hence monitored. Monitoring thus seems to be necessary but not sufficient in non-focal tasks. Indeed, when investigating the proportion of correctly remembered PM target cues out of only those trials that were monitored, only the highly non-focal condition differed from the focal condition in older

adults. Thus, although older adults monitored in that non-focal condition, they responded to less PM cues correctly than in the focal condition. How could one explain that finding?

As monitoring cannot account for the results, the lowered PM performance in the highly non-focal condition has to be caused by processes following monitoring in older adults. Several subsequent processes can be identified that might be negatively affected by age. Microstructure models of PM (Marsh, Hicks, & Watson, 2002; Rummel, Wesslein, & Meiser, 2017; Zuber, Kliegel, & Ihle, 2016) suggest that first the PM target event needs to be noticed, then the intention must be retrieved (and the context verified), before in a last step the intended action needs to be coordinated with the demands of the ongoing task: The due to the low frequency of PM trials prepotent ongoing task response needs to be inhibited, and in addition a switch from the ongoing task to the PM task set is required in order to finally perform the task.

Accordingly, in the present study older adults might have monitored the right target region, but once the PM target cues appeared they might not have realized their significance, and therefore not *noticed* the PM target cues. As shown in a recent EEG study by Hering et al. (2016), older adults may have difficulties realizing the relevance of the PM target cue compared to other ongoing task cues, and this difference seems to emerge relatively early in the cue processing phases. The authors suggest that a lowered discrimination between the PM and ongoing task might be caused by a reduced attentional sensitivity in older adults. Further, the *retrieval* of the intended task set might have been particularly problematic in the highly non-focal condition (see Ballhausen et al., 2017, for effects of maintenance load). This might have been due to the fact that the content of the target cues and the ongoing task were not overlapping. Finally, lower PM performance might have been a result of a *coordination problem* at the post-retrieval stage (see Bisiacchi, Schiff, Ciccola, & Kliegel, 2009). Both

inhibition of the ongoing task response as well as the switch from the ongoing to PM task might have been problematic in older adults, given their reduced executive functioning (e.g., Fisk & Sharp, 2004). Indeed, it was shown that increasing demands at response coordination relate to larger age effects (Ihle et al., 2013). However, present results cannot determine which of these post-monitoring processes led to the focality effect in older adults. Future studies need to identify which of these processes relate to reduced age-related performance in non-focal tasks.

The present study for the first time investigated PM monitoring behavior in both younger and older adults using eye tracking methods. This method allows to study allocation of attention, which is more functionally related to allocation of attentional resources to monitor for a target cue and thus more meaningful than the traditionally used ongoing task costs. Analyses of ongoing task response times was clearly sensitive to target location, as response time costs for the locally non-focal conditions were substantially greater than for the foveated conditions. However, large ongoing task costs do not necessarily relate to high monitoring, as intention maintenance (Ballhausen et al., 2017), general slowing in older adults (Henry et al., 2015), or target location could contribute to higher ongoing task costs. Thus, eye tracking seems to have advantages over cost measures and can be a good way of better understanding monitoring behavior (see also Bowden et al., 2017; Chen et al., 2013; Shelton & Christopher, 2016). However, due to age-related changes in the anatomy of the eye (e.g., Salvi, Akhtar, & Currie, 2006) as well as the greater need for glasses in the older sample, data quality was worse in the older adults compared to the younger ones. Lower quality of the data can enhance the chance that a fixation, although present, is not detected. For that reason, the proportion of trials that were monitored in the older adults must be interpreted as a „lower bound estimate". These limitations need to be kept in mind when comparing and

interpreting data of younger and older adults. In the current study, this potential limitation is of no concern, as we found no differences in the monitoring behavior of younger and older adults, who both monitored in the large majority of trials.

To sum up, older, but not younger adults, showed lower PM performance in a highly non-focal compared to a focal task. Present eye tracking data suggest that monitoring itself cannot account for focality effects in older adults. The results rather suggest that subsequent processes cause lower PM performance. Thus, monitoring may not automatically lead to correct remembering, particularly in older adults.

**References**

Bach, M. (1996). The Freiburg Visual Acuity test--automatic measurement of visual acuity.

*Optometry and Vision Science, 73*(1), 49-53.

Ball, B. H., Brewer, G. A., Loft, S., & Bowden, V. (2015). Uncovering continuous and transient

monitoring profiles in event-based prospective memory. *Psychonomic Bulletin &*

*Review, 22*(2), 492-499. doi:10.3758/s13423-014-0700-8

Ballhausen, N., Schnitzspahn, K. M., Horn, S. S., & Kliegel, M. (2017). The interplay of

intention maintenance and cue monitoring in younger and older adults' prospective

memory. *Memory & Cognition, 45*, 1113–1125. doi:10.3758/s13421-017-0720-5

Bisiacchi, P. S., Schiff, S., Ciccola, A., & Kliegel, M. (2009). The role of dual-task and task-

switch in prospective memory: Behavioural data and neural correlates.

*Neuropsychologia, 47*(5), 1362-1373. doi:10.1016/j.neuropsychologia.2009.01.034

Bisiacchi, P. S., Tarantino, V., & Ciccola, A. (2008). Aging and prospective memory: the role

of working memory and monitoring processes. *Aging Clinical and Experimental*

*Research, 20*(6), 569-577.

Bowden, V. K., Smith, R. E., & Loft, S. (2017). Eye movements provide insights into the

conscious use of context in prospective memory. *Consciousness and Cognition, 52*,

68-74. doi:doi.org/10.1016/j.concog.2017.04.003

Boywitt, C. D., & Rummel, J. (2012). A diffusion model analysis of task interference effects in

prospective memory. *Memory & Cognition, 40*(1), 70-82. doi:10.3758/s13421-011-

0128-6

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436.

Brewer, G. A. (2011). Analyzing Response Time Distributions Methodological and Theoretical Suggestions for Prospective Memory Researchers. *Zeitschrift Für Psychologie-Journal of Psychology, 219*(2), 117-124. doi:10.1027/2151-2604/a000056

Chen, S., Zhou, R., Cui, H., & Chen, X. (2013). Deficits in cue detection underlie event-based prospective memory impairment in major depression: an eye tracking study. *Psychiatry Research, 209*(3), 453-458. doi:10.1016/j.psychres.2013.01.015

Cherry, K. E., & LeCompte, D. C. (1999). Age and individual differences influence prospective memory. *Psychology and Aging, 14*(1), 60-76. doi:Doi 10.1037/0882-7974.14.1.60

Cohen, A.-L., Gordon, A., Jaudas, A., Hefer, C., & Dreisbach, G. (2016). Let it go: the flexible engagement and disengagement of monitoring processes in a non-focal prospective memory task. *Psychological Research*. doi:10.1007/s00426-016-0744-7

Cona, G., Bisiacchi, P. S., & Moscovitch, M. (2013). The Effects of Focal and Nonfocal Cues on the Neural Correlates of Prospective Memory: Insights From ERPs. *Cerebral Cortex*. doi:10.1093/cercor/bht116

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409-422). Amsterdam: Elsevier.

Dalmaijer, E. (2014). Is the low-cost EyeTribe eye tracker any good for research? *PeerJ PrePrints, 2*, e585v581. doi:10.7287/peerj.preprints.585v1

Deltour, J. J. (1993). *Echelle de vocabulaire Mill Hill de J. C. Raven: Adaptation française et normes comparées du Mill Hill et du Standard Progressive Matrices (PM38). Manuel*. Braine-le-Château: L'application des techniques modernes.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science, 14*(6), 286-290. doi:DOI 10.1111/j.0963-7214.2005.00382.x

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology-General, 134*(3), 327-342. doi:Doi 10.1037/0096-3445.134.3.327

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26*(7), 874-890. doi:10.1080/13803390490510680

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. *International Journal of Psychology, 38*(4), 245-256. doi:Doi 10.1080/00207590244000205

Guynn, M. J. (2008). Theory of monitoring in prospective memory: Instantiating a retrieval mode and periodic target checking. In M. Kliegel, A. M. Mark, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 53-76). New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates.

Hartwig, J., Schnitzspahn, K. M., Kliegel, M., Velichkovsky, B. M., & Helmert, J. R. (2013). I see you remembering: what eye movements can reveal about process characteristics of prospective memory. *International Journal of Psychophysiology, 88*(2), 193-199. doi:10.1016/j.ijpsycho.2013.03.020

Henry, J. D., Joeffry, S., Terrett, G., Ballhausen, N., Kliegel, M., & Rendell, P. G. (2015).

Prospective Memory Function in Late Adulthood: Affect at Encoding and Resource

Allocation Costs. *Plos One, 10*(4), e0125124. doi:10.1371/journal.pone.0125124

Hering, A., Wild-Wall, N., Gajewski, P. D., Falkenstein, M., Kliegel, M., & Zinke, K. (2016). The

role of cue detection for prospective memory development across the lifespan.

*Neuropsychologia, 93*(Pt A), 289-300. doi:10.1016/j.neuropsychologia.2016.11.008

Horn, S. S., Bayen, U. J., & Smith, R. E. (2013). Adult age differences in interference from a

prospective-memory task: a diffusion model analysis. *Psychonomic Bulletin &*

*Review, 20*(6), 1266-1273. doi:10.3758/s13423-013-0451-y

Ihle, A., Hering, A., Mahy, C. E., Bisiacchi, P. S., & Kliegel, M. (2013). Adult age differences,

response management, and cue focality in event-based prospective memory: a

meta-analysis on the role of task order specificity. *Psychology and Aging, 28*(3), 714-

720. doi:10.1037/a0033653

Ishihara, S. (1992). *Ishihara's tests for colour-blindness*. Tokyo: Kanehara & Co. Ltd.

Kalpouzos, G., Eriksson, J., Sjolie, D., Molin, J., & Nyberg, L. (2010). Neurocognitive systems

related to real-world prospective memory. *Plos One, 5*(10), e13304.

doi:10.1371/journal.pone.0013304

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception,*

*36*(SUPPL.).

Kliegel, M., Altgassen, M., Hering, A., & Rose, N. S. (2011). A process-model based approach

to prospective memory impairment in Parkinson's disease. *Neuropsychologia, 49*(8),

2166-2177. doi:10.1016/j.neuropsychologia.2011.01.024

Kliegel, M., Ballhausen, N., Hering, A., Ihle, A., Schnitzspahn, K. M., & Zuber, S. (2016).

Prospective Memory in Older Adults: Where We Are Now and What Is Next.

*Gerontology, 62*(4), 459-466. doi:10.1159/000443698

Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based

prospective memory: a meta-analysis on the role of focal versus nonfocal cues.

*Psychology and Aging, 23*(1), 203-208. doi:10.1037/0882-7974.23.1.203

Kliegel, M., Mahy, C. E. V., Voigt, B., Henry, J. D., Rendell, P. G., & Aberle, I. (2013). The

development of prospective memory in young schoolchildren: The impact of ongoing

task absorption, cue salience, and cue centrality. *Journal of Experimental Child

Psychology, 116*(4), 792-810. doi:10.1016/j.jecp.2013.07.012

Kominsky, T. K., & Reese-Melancon, C. (2017). Effects of context expectation on prospective

memory performance among older and younger adults. *Memory, 25*(1), 122-131.

doi:10.1080/09658211.2015.1131300

Kuhlmann, B. G., & Rummel, J. (2014). Context-specific prospective-memory processing:

evidence for flexible attention allocation adjustments after intention encoding.

*Memory & Cognition, 42*(6), 943-949. doi:10.3758/s13421-014-0405-2

Levi, D. M. (2008). Crowding--an essential bottleneck for object recognition: a mini-review.

*Vision Research, 48*(5), 635-654. doi:10.1016/j.visres.2007.12.009

Marsh, R. L., Cook, G. I., & Hicks, J. L. (2006). Task interference from event-based intentions

can be material specific. *Memory & Cognition, 34*(8), 1636-1643.

Marsh, R. L., Hicks, J. L., & Watson, V. (2002). The dynamics of intention retrieval and

coordination of action in event-based prospective memory. *Journal of Experimental

Psychology: Learning, Memory, and Cognition, 28*(4), 652-659.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective

memory retrieval: A multiprocess framework. *Applied Cognitive Psychology, 14*,

S127-S144. doi:Doi 10.1002/Acp.775

McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of

an emerging field*. Thousand Oaks, CA: Sage.

McDaniel, M. A., & Einstein, G. O. (2011). The neuropsychology of prospective memory in

normal aging: A componential approach. *Neuropsychologia, 49*(8), 2147-2155.

doi:DOI 10.1016/j.neuropsychologia.2010.12.029

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming

numbers into movies. *Spatial Vision, 10*(4), 437-442.

Reese-Melancon, C. (2013). Age, Focal Processing, and Monitoring in Event-Based

Prospective Memory. *Journal of Adult Development, 20*(3), 151-157.

doi:10.1007/s10804-013-9166-6

Reese, C. M., & Cherry, K. E. (2002). The effects of age, ability, and memory monitoring on

prospective memory task performance. *Aging Neuropsychology and Cognition, 9*(2),

98-113. doi:DOI 10.1076/anec.9.2.98.9546

Rose, N. S., Rendell, P. G., McDaniel, M. A., Aberle, I., & Kliegel, M. (2010). Age and

individual differences in prospective memory during a "Virtual Week": the roles of

working memory, vigilance, task regularity, and cue focality. *Psychology and Aging,

25*(3), 595-605. doi:10.1037/a0019771

Rummel, J., Wesslein, A. K., & Meiser, T. (2017). The role of action coordination for

prospective memory: Task-interruption demands affect intention realization. *Journal

of Experimental Psychology: Learning, Memory, and Cognition, 43*(5), 717-735.

doi:10.1037/xlm0000334

Salvi, S. M., Akhtar, S., & Currie, Z. (2006). Ageing changes in the eye. *Postgraduate Medical Journal, 82*(971), 581-587. doi:10.1136/pgmj.2005.040857

Scullin, M. K., McDaniel, M. A., & Shelton, J. T. (2013). The Dynamic Multiprocess Framework: evidence from prospective memory with contextual variability. *Cognitive Psychology, 67*(1-2), 55-71. doi:10.1016/j.cogpsych.2013.07.001

Scullin, M. K., McDaniel, M. A., Shelton, J. T., & Lee, J. H. (2010). Focal/nonfocal cue effects in prospective memory: monitoring difficulty or different retrieval processes? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(3), 736-749. doi:10.1037/a0018971

Shelton, J. T., & Christopher, E. A. (2016). A fresh pair of eyes on prospective memory monitoring. *Memory & Cognition*. doi:10.3758/s13421-016-0601-3

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(3), 347-361.

Smith, R. E., Horn, S. S., & Bayen, U. J. (2012). Prospective memory in young and older adults: The effects of ongoing-task load. *Aging, Neuropsychology, and Cognition, 19*(4), 495-514. doi:Doi 10.1080/13825585.2011.633161

Vercambre, M. N., Cuvelier, H., Gayon, Y. A., Hardy-Leger, I., Berr, C., Trivalle, C., . . . Clavel-Chapelon, F. (2010). Validation study of a French version of the modified telephone interview for cognitive status (F-TICS-m) in elderly women. *International Journal of Geriatric Psychiatry, 25*(11), 1142-1149. doi:Doi 10.1002/Gps.2447

Wechsler, D. (2011). *Echelle d'intelligence de Wechsler pour adultes (WAIS-IV) - Quatrième édition*.

Welsh, K. A., Breitner, J. C. S., & Magruderhabib, K. M. (1993). Detection of Dementia in the

Elderly Using Telephone Screening of Cognitive Status. *Neuropsychiatry*

*Neuropsychology and Behavioral Neurology, 6*(2), 103-110.

West, R., Carlson, L., & Cohen, A. L. (2007). Eye movements and prospective memory: What

the eyes can tell us about prospective memory. *International Journal of*

*Psychophysiology, 64*(3), 269-277. doi:10.1016/j.ijpsycho.2006.09.006

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious

perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160-168.

doi:10.1016/j.tics.2011.02.005

Zuber, S., Kliegel, M., & Ihle, A. (2016). An individual difference perspective on focal versus

nonfocal prospective memory. *Memory & Cognition, 44*(8), 1192-1203.

doi:10.3758/s13421-016-0628-5

# Appendix 6

Lauffs, M. M., Karacsony, D., & Herzog, M. H. (*ongoing*)
What makes a masterpiece? Surprisingly low inter-observer gaze-path similarity when viewing Renaissance art.

Authors: Marc M. Lauffs[1], Darius Karacsony[2], Michael H. Herzog[1],

[1] Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] Department of Architecture, École Polytechnique Fédérale de Lausanne (EPFL), University of Geneva, Switzerland

# What makes a masterpiece? Surprisingly low inter-observer gaze-path similarity when viewing Renaissance art

[…]

When inspecting/viewing a stationary image the gaze primarily consists of two elements: Extended periods of time during which the eyes' position is relatively stable (called *fixations*) and rapid, ballistic eye movements from one fixation to the next (called *saccades*). A sequence of fixations and saccades is also referred to as a *gaze path*. Virtually all information intake takes place during fixations, while vision is largely suppressed during saccades (e.g., Raymond, 1990; Holt, 1903; Burr, Morrone, & Ross, 1994; Ross, Burr, & Morrone, 1996; Castet & Masson, 2000). Smooth eye-movements only occur when the eyes pursue a moving object, but not when viewing a stationary image, and are therefore not discussed here. If the painter indeed uses the image to tell a story, the successive discovery of the story's elements should be reflected in the viewer's gaze. More precisely, which parts of the image are fixated, in what order, and for how long, should at least partly be driven by the underlying historia and therefore be similar across different observers viewing the same image.

To test this assumption, we recorded the gaze of human observers with an eye tracker, while they viewed images of paintings on a computer screen. We selected 10 masterpiece paintings of the Renaissance era. We only selected paintings depicting scenes that could also be told as a verbatim story, such as Masolino's *Healing of the Cripple and Raising of Tabitha* (1424) or Lippi's *Disputation with Simon Magus and Crucifixion of St Peter* (1482). In addition, we included Picasso's *Guernica* (1937) - a modern, abstract work, painted in black and white, that is evidently very different from the other paintings.

We then compared the similarity of our observers' gaze paths using ScanMatch (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010) for Matlab (TheMathWorks, Natick, MA, USA). ScanMatch computes the similarity of two gaze paths as a function of the changes needed to match one to the other.

We tested two concrete hypotheses. First, we assumed that if the artist achieved to guide the gaze of the observer, this should overrule idiosyncrasies in the viewing behavior of the individual. Therefore, the gaze path should be influenced more by which image somebody looks at than who is looking at it. To test this hypothesis, we computed the average gaze path similarity within images and across observers and compared it to the average similarity within observers, but across images. Second, we tested whether the similarity of our observers' gaze paths is *higher* than would be expected by chance. To determine the baseline similarity, we generated a large number of quasi-random gaze paths, in which the direction of the saccades was determined purely by chance. To keep the simulated data comparable to the human data, we constrained the simulated gaze paths to have the same number of fixations, the same fixation duration, the same saccade amplitudes, and the same starting point as the human data for the respective image. For a second set of simulations, we sampled from the human fixation positions and durations with replacement (bootstrapping), closely reproducing the human fixations but in random order.

By computing the gaze path similarity of all unique pairs of *human* observers and all unique pairs of *simulated* observers, we could then compare the average similarity of the human gaze paths against two different chance models. If the gaze of the observers is indeed driven by some stylistic element that the author embedded in the painting, the gaze paths of different

observers should be more similar than gaze paths that are random and, hence, not driven by the image at all.

## Methods

### Observers

Eight observers (6 female, age range 23-27 years) participated in the experiment as part of a Master's class in the EPFL architecture program. All but one observer wore glasses. Six subjects had normal or corrected-to-normal visual acuity, as indicated by a score above 1 in at least one eye in the Freiburg Visual Acuity Test (Bach, 1996). One subject had values above 0.7 in both eyes and for one subject the acuity test was not performed. The right eye was dominant for all observers. All participants gave written informed consent prior to the experiment, which was approved by the local ethics committee and performed in accordance with the Declaration of Helsinki (World Medical Association, 2013).

### Apparatus

The observers were comfortably seated and positioned in the headrest of a SMI iViewX HiSpeed 1250 eyetracker. Eyetracking data were recorded binocularly at 500 Hz and averaged over both eyes to reduce noise. The room was darkened, except for artificial lighting above the participant. Images were presented on a 24inch Asus VG248QE LCD monitor (1920 x 1080 px, 120 Hz, http://display-corner.epfl.ch/) using Matlab and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997). At the viewing distance of 1 m the monitor subtended 30 x 17 ° [degrees of visual angle].

**Procedure**

The images were presented in the center of the screen and with the maximum possible size that preserved the original aspect ratio. Parts of the screen not covered by the image were black. The images were presented in random order for 30 secs each. Between images a 1.5 secs blank screen was displayed, followed by a central fixation point presented for 3 secs. The observers were instructed to fixate the central fixation point, so that the gaze of all subjects would lie on the screen center when the image appears.

We presented images of 10 paintings of the Renaissance era and 1 modern abstract work by Picasso:

- Third painting of Story of Nastagio degli Onesti by Sandro Botticelli (1483)

- Primavera by Sandro Botticelli (1482)

- Flagellation of Christ by Piero della Francesca (1460)

- Annunciazione by Leonardo da Vinci (1472)

- Raising of the son of Theophilus and Saint Peter enthroned by Masaccio (1427)

- The hunt in the forest by Paolo Uccello (1470)

- Disputation with Simon Magus and Crucifixion of St Peter by Filippino Lippi (1482)

- The tribute money by Masaccio (1425)

- Healing of the Cripple and Raising of Tabitha by Masolino (1424)

- The school of Athens by Raphael (1511)

- Guernica by Picasso (1937)

## Analysis

The analysis is schematically depicted in Figure 1. First, we detected fixations in the eyetracking recordings (Figure 1, top), using an in-house algorithm implemented in Matlab. We then transformed the fixation-sequence into strings. We subdivided the screen into a 20 by 11 cell grid that covered the entire screen area. The grid size was chosen so that it would yield roughly square cells of 1.5 by 1.5°, which is well above the measured accuracy of our eye-tracker (Nyström et al. 2013) and below the functional field of view in picture viewing (Nelson and Loftus, 1980). Each cell of the grid was labeled with a unique letter combination of a lower-case letter indexing the row and an upper-case letter indexing the column (e.g., cell aB = row a, column B; Figure 1, middle). We then coded the fixation positions with the letters of the cell it fell in. The duration of the fixation was encoded by repeating the letters once for every 50 ms that the fixation lasted (Figure 1, bottom). Multiple fixations in different parts of the same cell were coded separately, but with the same
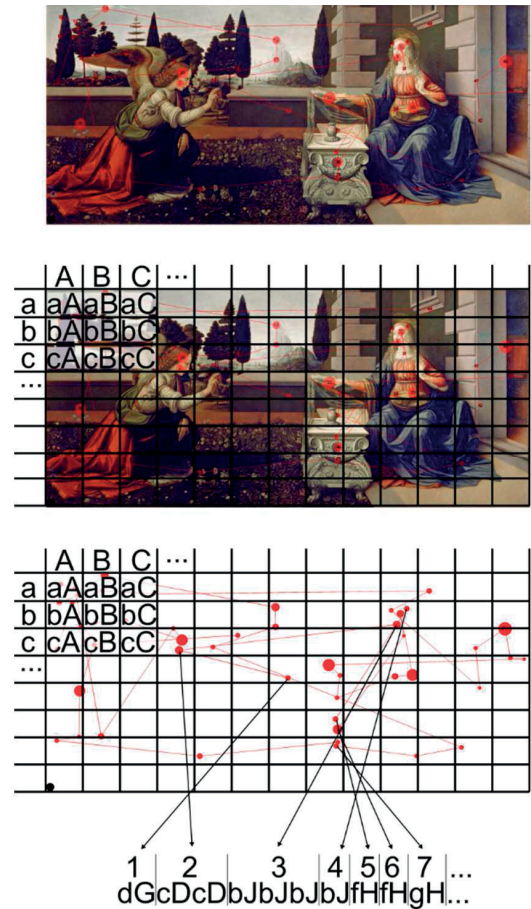


*Figure 6* Schematic of the analysis. Top: Fixations were detected in the raw eyetracking data. Middle: The image was subdivided in a grid of 20 x 11 cells (1.5 x 1.5 ° per cell). Each cell was named with a combination of an upper-case letter indexing the column and a lower-case indexing the row of its position. Bottom: The fixation positions were then coded with the letter indices of the grid cells. For example, the first fixation usually fell in cell 'dG' (row d, column G), the center of the screen, because a central fixation cross was presented before the image. Fixations 3 and 4 are separate fixations in different positions (here, lips and halo of the same person), and are hence coded separately. However, since both positions fall in the same cell they are coded with the same letters. See main text for further details.

letters. The resulting string representation of the gaze path encodes the position, order, and duration of its fixations. For example, the string "dG cD cD bJ bJ bJ bJ fH fH gH" represents a gaze path with a first fixation in the center of the image ("dG"), a slightly longer second fixation in the upper left-hand part of the image ("cD cD"), an even longer third fixation in the upper right-hand part of the image ("bJ bJ bJ"), and so forth (cf. Figure 1, bottom).

We then used the ScanMatch (Cristino et al., 2010) toolbox for Matlab (TheMathWorks) to compare these strings pairwise. ScanMatch matches two strings by inserting, deleting, or substituting letters or gaps in either string. Each change is associated with a cost. Substitutions are more costly for distant than proximate cells (e.g., it is less costly to substitute cell aB by aC than dF). By leveraging an algorithm originally developed for efficient sequencing of DNA data (Needleman & Wunsch, 1970), ScanMatch finds the most cost-efficient combination of changes to align both strings and computes am alignment score that expresses the similarity. The alignment score is normalized to a number between 0 and 1 by taking into account the maximum length of the original strings. Fewer and less costly changes lead to a high alignment score, more and more costly changes lead to a low alignment score. Hence, the higher the alignment score, the more similar are the strings and underlying gaze paths.

The ScanMatch toolbox has a number of free parameters that we varied to ensure the robustness of our conclusions. Most importantly, we ran the analyses with no cost for the insertion of gaps or with a gap cost of -2. We also ran the analyses neglecting the duration of fixations or taking fixation durations into account (i.e., temporal bins of 0 or 50 ms). For all analyses, we defined substitution costs based on the Euclidian distance between cells. We used the recommended substitution threshold of two times the standard deviation of the saccade amplitudes (Cristino et al., 2010). The substitution threshold determines at what

distance between ROIs substitutions are penalized by a negative score. The chosen substitution threshold causes the algorithm to align cells only when they lie roughly within saccade distance from each other. We ran the analyses with the substitution threshold computed individually for each image or with a single global substitution threshold based on all measured saccades. We used the same full-screen 20x11 grid for all images, although some images did not fill the entire screen due to an incompatible aspect-ratio. This might artificially increase the similarity of gaze paths on smaller compared to larger images, e.g., by decreasing the possible variance of the fixation positions. This is only a concern for across-image comparisons and is irrelevant for within-image comparisons.

## Simulations I – Sampling from normal distributions

Per image, we computed 1000 quasi-random gaze paths. To warrant comparability to the human observers, we restricted the simulations using the descriptive statistics of the human data for the same image. The number of simulated fixations, fixation durations, and saccade amplitudes were randomly drawn from a normal distribution with the same mean and variance as the human data for the same image. The direction of the simulated saccades was determined randomly. All simulated gaze paths started within one mean saccade amplitude from the screen center, because the human observers, too, were asked to focus on a central fixation point before they were shown the image. Only simulated fixation positions within the bounds of the image were accepted. Only simulated fixation durations above 50 ms were accepted, as this was the minimum fixation duration our algorithm could detect in the human data.

## Simulations II - Bootstrapping

Above mentioned simulations suffer from the disadvantage that they are randomly drawn from normal distributions although the human data were not necessarily normally distributed (for example, fixation durations were heavily right-skewed). Therefore we created a second set of 1000 simulated gaze paths by bootstrapping the human data. For each image we listed all human fixations irrespective of order and which observer they were recorded from. We then drew randomly with replacement from this list until the desired number of fixations was reached. Fixation durations were sampled in the same way, but independently of the fixation positions. To guarantee that the simulations started close to the fixation point like the human data, the first simulated fixation position was drawn randomly with replacement from the human observers' first fixations on the given image. The number of fixations per simulation was determined by drawing randomly with replacement from the human fixation counts for the given image. Hence, only fixation positions and durations present in the human data were used for the bootstrapped simulations. Accordingly, the number of fixations and fixation durations were very similarly distributed in the bootstrapped and human data, but saccade amplitudes were much larger and more variable in the simulations.

**Statistics**

We performed two types of comparisons: 1) We compared the average within-subject similarity and within-image similarity in a two-sample *t*-test. 2) We calculated the average within-image similarity across human observers, simulations, and bootstrapped simulations for each image. We then performed a mixed ANOVA with factors group (3; human, simulated, bootstrapped) and image (11) to test for a main effect of the group factor and eventual differences depending on the image. All statistical tests were performed in JASP (http://jasp-stats.org).

# Curriculum Vitae

CONTACT INFORMATION

Marc M. Lauffs

Ecole Polytechnique Fédérale de Lausanne (EPFL)

Brain Mind Institute, Laboratory of Psychophysics

EPFL SV BMI LPSY, AI3102, Station 19, 1015 Lausanne, Switzerland

marc.lauffs@epfl.ch, marc.lauffs@gmail.com

PERSONAL INFORMATION

Born: May 12th 1985 (Krefeld, D)

Nationality: German

EDUCATION

04.2013 - 12.2017 PhD Neurosciences, Ecole Doctorale en Neurosciences, Ecole Polytechnique Fédérale de Lausanne EPFL (Lausanne, CH)

09.2005 - 06.2012 MSc & BSc Psychology, Universiteit Twente (Enschede, NL)

09.1995 - 06.2004 Abitur, Maria-Sibylla-Merian Gymnasium (Krefeld, D)

EMPLOYMENT HISTORY

04.2013 - 12.2017 Doctoral Researcher, Laboratory of Psychophysics, Brain-Mind Institute, Ecole Polytechnique Fédérale de Lausanne EPFL (Lausanne, CH)

05.2016 - 06.2016 Research Consultant, Airbus Defence & Space GmbH (Immenstaad, D)

06.2010 - 06.2012 MSc & BSc Researcher, Department of Cognitive Psychology and Ergonomics, Universiteit Twente (Enschede, NL)

OTHER PROFESSIONAL EXPERIENCES

07.2009 - 01.2013 Travel Guide & Destination Manager, ruf Reisen GmbH (Bielefeld, D)

PUBLICATIONS

Lauffs, M. M., Öğmen, H., & Herzog, M. M. (2017). Uncertainty does not hamper nonretinotopic motion perception. *Journal of Vision, 17*(9):6, 1-10.

Lauffs, M. M., Shaqiri, A., Brand, A., Roinishvili, M., Chkonia, E., Öğmen, H., & Herzog, M. H. (2016). Local versus global and retinotopic versus non-retinotopic motion processing in schizophrenia patients. *Psychiatry Research*, *246*, 461-465.

Lauffs, M. M., Choung, O. H., Öğmen, H., & Herzog, M. H. (*submitted*)

Ballhausen*, N., Lauffs*, M. M., Herzog, M. H., & Kliegel, M. (*submitted*)
(* equal contributions)

CONFERENCE CONTRIBUTIONS

09.2017 Talk: Swiss Psychological Society, bi-annual meeting, Lausanne, CH.

08.2016 Poster: Lemanic Neuroscience Annual Meeting, Les Diablerets, CH.

08.2016 Talk: European Conference on Visual Perception, Barcelona, ESP.

05.2016 Poster: World Psychiatric Association Regional Meeting, Tbilisi, GE.

08.2015 Poster: Lemanic Neuroscience Annual Meeting, Les Diablerets, CH.

08.2015 Talk: European Conference on Visual Perception ECVP, Liverpool, UK.

05.2015 Poster: Vision Science Society annual meeting, St. Petersburg, FL, USA.