

Reference-frames in vision: Contributions of attentional tracking to nonretinotopic perception in the Ternus-Pikler display

Marc M. Lauffs*

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



Oh-Hyeon Choung*

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



Haluk Öğmen

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



Michael H. Herzog

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



Dirk Kerzel

Faculté de Psychologie et des Sciences de l'Éducation,
Université de Genève, Geneva, Switzerland



Perception depends on reference frames. For example, the “true” cycloidal motion trajectory of a reflector on a bike’s wheel is invisible because we perceive the reflector motion relative to the bike’s motion trajectory, which serves as a reference frame. To understand such an object-based motion perception, we suggested a “two-stage” model in which first reference frames are computed based on perceptual grouping (bike) and then features are attributed (reflector motion) based on group membership. The overarching goal of this study was to investigate how multiple features (i.e., motion, shape, and color) interact with attention to determine retinotopic or nonretinotopic reference frames. We found that, whereas tracking by focal attention can generate nonretinotopic reference-frames, the effect is rather small compared with motion-based grouping. Combined, our results support the two-stage model and clarify how various features and cues can work in conjunction or in competition to determine prevailing groups. These groups in turn establish reference frames according to which features are processed and bound together.

Introduction

According to behavioristic theories, stimuli generate *sensations*, which are stored in memory. Correlations between stimuli (classical conditioning) or stimuli and behavioral responses (operant conditioning) cause these sensations to be *associated* with each other to generate complex representations. However, many predictions of this approach failed when subjects responded in disagreement with the “association strengths” of stimuli. Several theoreticians then resorted to the concept of *attention* as pointed out by Koffka (1922, p. 535): “*wherever there is an effect that cannot be explained by sensation or association, there attention appears upon the stage.*” A key shortcoming of behavioristic theories is their inability to define the stimulus independent of the observer, a problem known as the “stimulus definition” problem. To deal with this problem, Tolman introduced “intervening variables” to account for how internal states of the observer can modify the stimulus (Tolman, 1938). In contrast to behaviorists, Gestalt psychologists proposed that organized structures, i.e., Gestalts, form the fundamental units of visual processing (Wagemans, 2015). Hence, to define a stimulus, one needs to refer to the internal

Citation: Lauffs, M. M., Choung, O. H., Öğmen, H., Herzog, M. H., & Kerzel, D. (2019). Reference-frames in vision: Contributions of attentional tracking to nonretinotopic perception in the Ternus-Pikler display. *Journal of Vision*, 19(12):7, 1–15, <https://doi.org/10.1167/19.12.7>.

<https://doi.org/10.1167/19.12.7>

Received April 30, 2019; published October 17, 2019

ISSN 1534-7362 Copyright 2019 The Authors



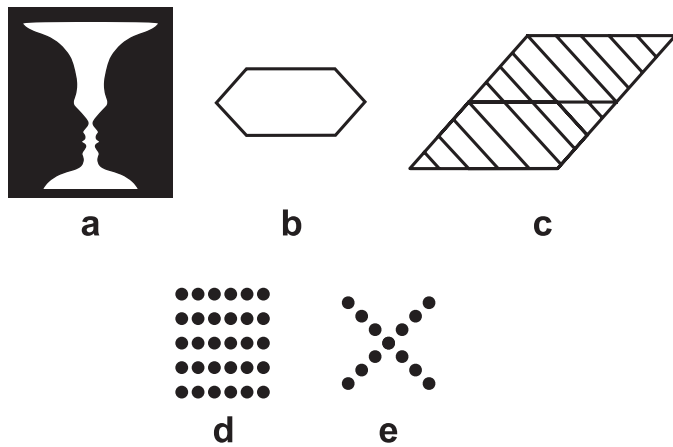


Figure 1. Examples of different grouping organizations. Reprinted with permission from Aydın, Herzog, & Ögmen (2011). Copyright 2011, Elsevier.

cognitive structures of the subject. Gestaltists proposed that grouping operations can generate a variety of possible groups, starting from simple ones and proceeding toward more complex ones, until the structure meets the goals of the observer. They introduced the concept of *attitude* to express this active role of internal structures in giving a directness to perception, such as the expectation of a particular organization or outcome. Attention, in turn, was defined as a special case of attitude, which is *unspecific* toward a particular organization or outcome. These ideas found support in experiments that employed decoy tasks to minimize the involvement of attention (Krechevsky, 1938; Köhler & Adams, 1958).

The relationship between perceptual grouping and attention was investigated by Beck (1966a, 1966b; 1967) who proposed that grouping occurs before the deployment of attention, a view supported by others (Caelli & Julesz, 1979; Julesz, 1991). On the other hand, Treisman and colleagues suggested that grouping and binding of stimuli occur pre-attentively *within* feature dimensions, but they require attention to occur *across* feature dimensions (Treisman, 1982; Treisman & Gelade, 1980). One pitfall of controlling attention by decoy/cover tasks is that it relies on memory for report. Hence, these studies may conflate the role of attention in memory with its role in perceptual grouping. Moore and Egeth (1997) used an implicit-measure approach to circumvent this problem and provided evidence that grouping occurs pre-attentively.

Although these results may appear contradictory, the view proposed by Gestalt psychologists, that attention and grouping are different but interacting processes, are consistent with these findings. Figure 1 shows some simple examples to illustrate this concept. In Figure 1a, the figure-ground organization is bistable in that both organizations (faces and vase) are of

similar strength. Focusing attention can alter our percept. The hexagon in Figure 1b is contained in Figure 1c; however, no matter how much attention is directed into elements in Figure 1c that make up the hexagon, one cannot perceive the hexagon as a figure in itself in Figure 1c.

If attention and grouping are different but interacting processes, it remains to determine how they interact. According to Gestalt psychology, elements are grouped into wholes by a variety of laws such as common fate, similarity, proximity, good continuation, etc. But what holds elements of a group together? We suggest that each group possesses a reference frame according to which elements are synthesized into groups. A simple example of this is the motion of a reflector on a bicycle wheel. In isolation, it is perceived to move on a cycloidal trajectory, which corresponds to its retinotopic motion when the eyes do not follow the reflector, but are directed at a stationary fixation mark. However, when a second reflector is added at the center of the wheel, the two reflectors become part of the same group and the trajectory of the reflector appears circular (e.g., Proffitt & Cutting, 1980; Proffitt, Cutting, & Stier, 1979). We perceive circular motion because the two reflectors on the wheel are grouped into the same Gestalt and the motion of the reflector at the center of the wheel serves as a *reference frame* for the group. At the same time, it is impossible to perceive the cycloidal motion because this “non-retinotopic” reference frame dominates (Boi, Ögmen, Krummenacher, Otto, & Herzog, 2009; Lauffs, Choung, Ögmen, & Herzog, 2018; Lauffs, Ögmen, & Herzog, 2017). Hence, reference frames determine how stimuli are grouped and what we perceive consciously.

To capture these ideas in a simple model, we proposed a “two-stage” spatiotemporal processing architecture (Figure 2), in which the first stage consists of grouping stimuli by Gestalt principles with modulation from attentional processes to establish reference frames for each group (Ögmen & Herzog, 2010, see Figure 18). These reference frames are then used to attribute features according to group membership and synthesize “objects” or “Gestalts” of perception.

The importance of grouping is evident when considering that vision is retinotopic in the first stages of vision. For example, neighboring elements in the visual field activate neighboring photoreceptors in the retina. This retinotopic encoding principle is preserved in the LGN and early visual areas. However, as shown by the bicycle example previously mentioned, perception is usually nonretinotopic. Hence, the grouping operations at the first stage establish not only retinotopic reference frames, but also nonretinotopic reference frames.

The Ternus-Pikler display is a suitable way to investigate the transition from retinotopic to non-

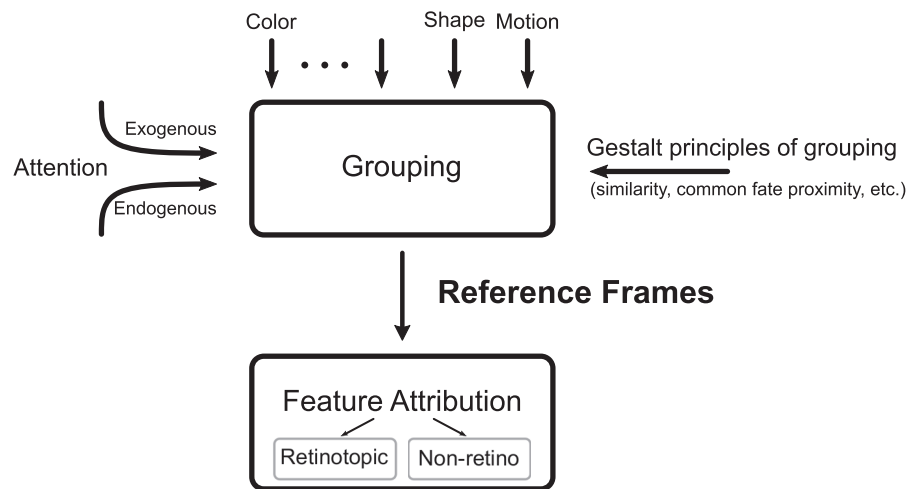


Figure 2. The two stage model: Features such as shape, motion, and color are used along with Gestalt principles of grouping, such as similarity, common fate, and proximity, to form perceptual groups. A reference frame is determined for each group and stimulus features are attributed to each group according to these reference frames. The retinotopic or the nonretinotopic percept results from differences in the attribution of features. Attention can play a role in this process for example by modulating the perceptual groups, in particular when they are ambiguous, or by directly establishing a reference frame that follows the spatiotemporal trajectory of focal attention.

retinotopic motion perception. When two disks are briefly presented on a computer screen and reappear at the same position after an interstimulus interval (ISI), observers perceive two flickering disks. In the Ternus-Pikler display, a third disk is added alternately to the left or right to induce apparent motion. When the ISI is very brief (e.g., 0 ms) the third disk appears to jump from the left to right of the two stationary disks, which is referred to as element motion. When the ISI is long (e.g., 200 ms), the three disks form a perceptual group and all three disks appear to shift left and right in concert, which is referred to as group motion. In past experiments (e.g., Boi et al., 2009), we used this effect to study nonretinotopic motion perception. We added a white dot to each disk, which we repositioned from frame to frame (Figure 3). The observer's gaze was focused on a central fixation point and eye fixation was monitored by an eye tracker. Hence, the stimulus positions were the same in screen-based and retinotopic coordinates. We chose the dot positions so that the dots in the two stationary disks appeared to move up-down or left-right when no group or element motion was perceived. In the following, this linear dot motion is referred to as the retinotopic dot motion because it describes the dot's motion in retinotopic coordinates (Figure 3a, b). When the ISI was long enough for group motion, the dot in the middle disk was perceived to rotate. Because the dot did not rotate in retinotopic coordinates, dot rotation had to be computed non-retinotopically after group motion was established. The perceived dot rotation is therefore referred to as nonretinotopic dot motion (Figure 3c). In other words, the grouping of elements across the two frames of the

Ternus-Pikler display determines element correspondence across the two frames (see the arrows in Figure 3), which in turn serves as reference frames according to which the dot rotation is perceived. In retinotopic coordinates, the dots still move linearly up-down and left-right. Here, we used a modified version of the Ternus-Pikler display (Lauffs et al., 2018) where the dots rotated in retinotopic coordinates (Figure 3d), instead of linear movement. Therefore, both retinotopic and nonretinotopic dot motion included either clockwise or counterclockwise rotation.

In previous studies (Boi et al., 2009; Lauffs et al., 2017; Lauffs et al., 2018), it was hypothesized that the grouping of disks across the two frames determines the reference frame for evaluating the motion of the dots. It was also shown that attention can modulate spatiotemporal grouping in Ternus-Pikler displays, with group motion requiring more attentional resources to prevail against element motion (Aydin, Herzog, & Ögmen, 2011). Hence, an alternative account is possible. As shown in Figure 2, attention can play a role in determining the reference frame either indirectly by modulating spatiotemporal groupings, especially when they are ambiguous, or directly by setting a reference frame that follows its tracking trajectory. The latter idea is supported by studies showing a strong link between attentional tracking and the perception of apparent motion.

Two seminal studies showed that attentional tracking of salient texture elements (Lu & Sperling, 1995) or isoluminant color (Cavanagh, 1992) results in perceived motion of the tracked feature. Verstraten, Cavanagh, and Labianca (2000) measured the maximal speed of

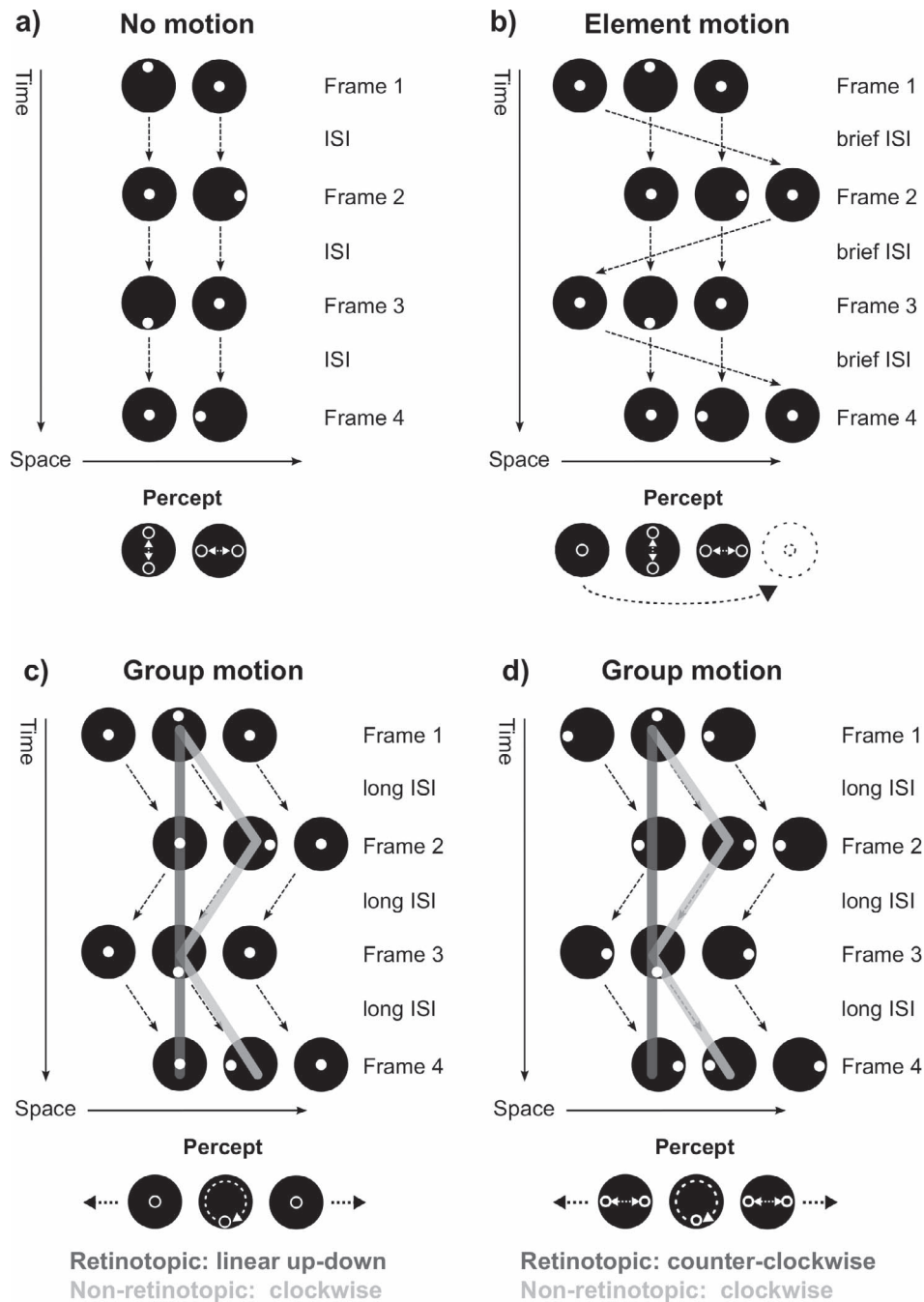


Figure 3. Ternus-Pikler display with linear retinotopic and circular non-retinotopic motion (Boi et al., 2009). (a) No motion condition. When two disks are presented on the screen, the dot in one disk is perceived to move up-and-down in one and left-and-right in the other disk. (b) When a third disk is added alternately to the left and right with a short ISI (e.g., 0 ms), the disks are perceived as two flanking disks in the middle and one disk jumping left-and-right (element motion). Dots in two middle disks are perceived to move as in the two disks condition, while the dot in the third disk stays in the center (retinotopic percept). (c) When the ISI is prolonged (e.g., 200 ms), the three disks are perceived to move left-and-right in concert (group motion), and the dot motion percept changes: The dot in the middle disk is perceived to rotate (nonretinotopic percept), and the dots in the left and right disks are perceived to stay in the center. (d) In the current experiment, circular retinotopic motion, instead of linear retinotopic motion, is used as in Lauffs et al. (2018). Therefore, both the retinotopic and nonretinotopic interpretation included either clockwise or counterclockwise rotation.

attentional tracking using bistable apparent motion stimuli. When confronted with bistable apparent motion, observers may decide to track a designated part of the stimulus, which disambiguates the apparent motion and results in a fixed perceived direction of motion. For instance, when a cross (×) and a plus sign (+) are presented in rapid alternation, observers perceive randomly reversing clockwise or counter-clockwise rotation. However, when they track a spoke of the cross, a consistent direction of motion is perceived. In this case, the experimenter may evaluate the precision of the attentional tracking by asking the observer to compare the position of a probe to the position of the tracked stimulus after various intervals. It was observed that the maximal rate at which observers could track a rotating stimulus with an accuracy of 75% correct was 4–8 Hz. That is, tracking was not so much limited by the angular velocity of the spoke, but rather by the flicker rate of the apparent motion stimulus.

The flicker rate in studies on nonretinotopic motion perception in the Ternus-Pikler display is well within the limits of attentional tracking. For instance, group motion occurs in the Ternus-Pikler display with a stimulus-onset-asynchrony (SOA) of 333 ms, which corresponds to a flicker rate of 3 Hz (Lauffs et al., 2018). Thus, the rate of apparent motion is slow enough to allow for attentional tracking. It is therefore possible that attentional tracking of the central position in the Ternus-Pikler display during group motion contributed to the perception of nonretinotopic dot motion. The overarching goal of the present study was to investigate how multiple features, such as motion, shape, and color, interact with attention to determine retinotopic or nonretinotopic reference frames. To evaluate the role of attentional tracking, we compared the perception of rotational dot motion in conditions where the relevant dot was highlighted by relative position (group motion), absolute position, color, or shape. All cues are expected to facilitate the tracking of the relevant motion stimulus, allowing us to evaluate the contribution of attentional tracking to the motion percept.

Experiment 1

In Experiment 1, we asked whether the perception of the retinotopic and nonretinotopic dot motion could be enhanced by showing the relevant disk in a distinct color. Through its salience, color may attract and direct attention whose spatiotemporal trajectory provides a reference frame for motion perception (see left side of model in Figure 2). On the other hand, Hein and Moore (2012) demonstrated that color may also affect

grouping by establishing spatiotemporal correspondence across stimulus frames (see right side of model in Figure 2). Thus, effects of color could result from attentional tracking or spatiotemporal correspondence. Experiment 1 deliberately accepts the ambiguity to establish effects of color, whereas Experiment 3 investigates conditions where the disks in the critical conditions were equally salient so that effects of attentional tracking could be isolated.

We expected beneficial effects of the color cue in conditions where performance was poor (see conditions C4 and C6 in Figure 4). Perception of retinotopic motion was poor when group motion prevailed with three disks, but excellent when only two stationary disks were shown (see conditions C4 and C2 in Figure 4). Possibly, group motion replaced the retinotopic reference frame by a nonretinotopic one thereby making retinotopic motion invisible (Lauffs et al., 2018). For the perception of nonretinotopic motion, the situation was opposite: Performance was poor when the two disks were stationary and excellent when a third disk was added and group motion was perceived (see conditions C6 and C8 in Figure 4).

With group motion, a colored disk at a fixed retinal position (condition C3 vs. C4) may improve perception of retinotopic dot motion either because perceived group motion is reduced (Hein & Moore, 2012) or because attentional tracking of the relevant disk is facilitated. Conversely, a colored disk at the center position in the group (condition C7 vs. C8) may improve perception of nonretinotopic dot motion either because perceived group motion is increased (Hein & Moore, 2012; Proffitt & Cutting, 1980) or because attentional tracking of the central disk is facilitated. With two stationary disks, perception of nonretinotopic motion may improve because color establishes spatiotemporal correspondence of the relevant disk or because the attentional cue facilitates attentional tracking (condition C5 vs. C6). Importantly, color allows establishing spatiotemporal correspondence of the response-relevant disk by linking the most salient stimulus across frames. This process requires very little attention, but may be accomplished in a bottom-up manner (Itti & Koch, 2001).

Methods

Participants

Twelve participants took part in the experiment after giving written informed consent. Two participants were excluded from analysis, because their performance was lower than 60% correct in the retinotopic condition with two disks or the nonretinotopic condition with three disks (where excellent performance is expected based on previous experiments; Lauffs et al., 2018). We retained the data of 10 participants (mean age: 22.6 ±

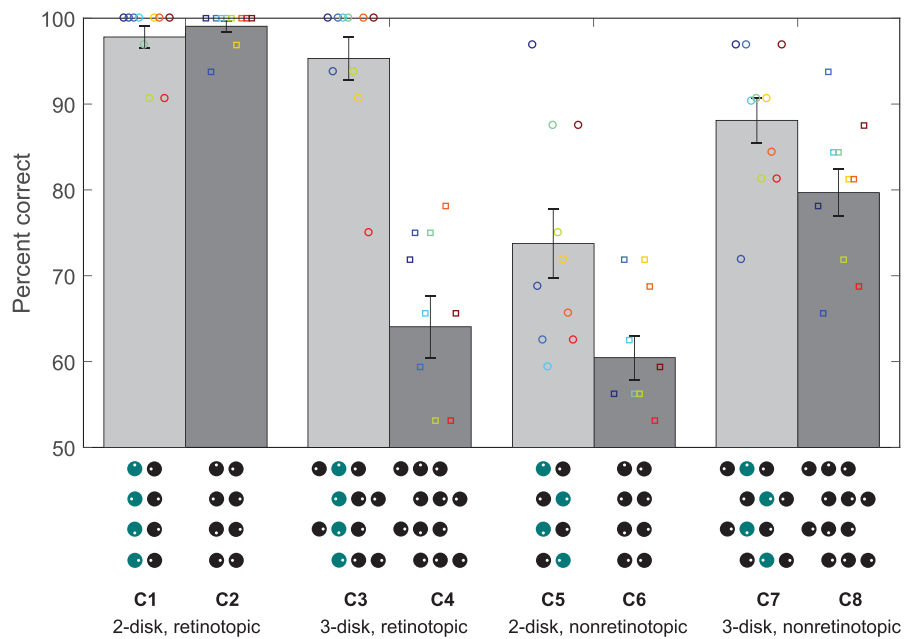


Figure 4. Dot rotation direction discrimination in Experiment 1. (C1, C2) Retinotopic tracking with 2 disks: Two disks are presented and participants were asked to track the retinotopic rotation. (C3, C4) Retinotopic tracking with three disks: Three disks are presented and participants were asked to track the retinotopic rotation. (C5, C6) Nonretinotopic tracking with two disks: Two disks were presented and participants were asked to track the nonretinotopic rotation. (C6, C8) Nonretinotopic tracking with three disks: Three disks were presented and participants were asked to track the nonretinotopic rotation. Error bars are ± 1 SEM. Colored data points show the mean performance of the individual observers.

2.7 years; half were female, all were right-handed, and eight had right eye dominance). Seven participants had taken part in experiments with the Ternus-Pikler display in the past, but all were naïve to the purpose of the current experiment. All participants had normal or corrected-to-normal visual acuity, as indicated by a binocular score greater 1.0 in the Freiburg Visual Acuity Test (Bach, 1996). All experiments were conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013) and were approved by the local ethics committee.

Apparatus

Stimuli were displayed on a gamma-calibrated 24.5-in. BenQ XL 2540B LCD monitor (1,920 × 1,080 pixels, 60 Hz; <http://display-corner.epfl.ch>; BenQ Corporation, Taipei, Taiwan). Viewing distance was 66 cm. The participants' chin and forehead were positioned in a SMI iViewX Hi-Speed 1250 eye tracker (SensoMotoric Instruments, Teltow, Germany), which was used to monitor eye fixation. Sample rate was 500 Hz and binocular samples were averaged to reduce noise. Trials containing eye movements larger than 1.5° (degrees of visual angle) or periods of data loss longer than 250 ms were discarded and repeated at a randomly chosen moment in the same block. Responses were collected using hand-held push buttons. When no response was

registered within 3 s or an eye movement was detected, the trial was repeated at a randomly chosen moment in the same block. A feedback tone sounded when no response was registered. In the case of eye movements, a feedback tone was played and a text message was displayed, reminding the participant to keep their gaze on the fixation point. No error feedback was provided.

Task, procedure, and design

The eight experimental conditions (C1–C8) are illustrated in Figure 4. We presented one block of 32 trials for each of the eight conditions. The order of blocks with color cue was random, but blocks with a colored disk were always followed by a block of the corresponding condition without color cue (i.e., C1 was followed by C2). Observers were instructed before each block of trials, using a video animation of the upcoming stimulus (provided in the supplementary materials, Supplementary Movies S1–S45). In a block with a color cue, participants were instructed to report the rotation of the dot (clockwise vs. counterclockwise) of the colored disk. In blocks without color cue, the position of the task-relevant disk was the same as in the preceding block *with* color cue. That is, observers were instructed to track the disk that appeared on the left-hand side in the first frame, and whether it would change position every other frame or not. Thus, there

was no ambiguity regarding the response-relevant disk when all disks were black. In each trial, either two or three disks with white dots were presented (see Figure 4 and Supplementary Movies S1–S8). In the conditions with the color cue, one disk was turquoise and the other disk(s) were black. Otherwise, all disks were black. The turquoise disk was either presented at the same position on the screen in all frames of the trial (retinotopic conditions) or switched position with a neighboring black disk in every other frame (nonretinotopic conditions). In each block of trials, the retinotopic and nonretinotopic rotation directions (clockwise vs. counterclockwise) and the initial orientation of the relevant rotation (3, 6, 9, 12 o'clock) were counterbalanced in a fully factorial fashion. The retinotopic and nonretinotopic rotation could hence be in the same (e.g., both clockwise) or in opposite directions (i.e., one clockwise and the other counterclockwise).

Before the experiment, participants performed one training block of 16 trials with auditory error feedback, and one warm-up block of 48 trials without error feedback. In these blocks, we only used the stimuli where one of the disks was turquoise and presented the different conditions in random order. The participants were instructed to report the rotation of this disk after presentation of the last frame.

Stimulus

Either two or three disks with a diameter of 2° were presented 4° above a central fixation point (diameter = 0.05° , red, 20 cd/m^2). The disks were horizontally aligned and separated by a gap of 0.5° . Each disk contained a white dot (100 cd/m^2) with a diameter of 0.25° . A white dot was positioned either in the center of the disk or halfway between the center and the rim in different angular positions (3, 6, 9, or 12 o'clock). The disks were either black (0.4 cd/m^2) or turquoise (33 cd/m^2) depending on the condition. The background was midlevel gray (50 cd/m^2).

The disks were presented for 100 ms and reappeared after an interstimulus interval (ISI) of 200 ms. Per trial, only four stimulus frames with disks and dots were presented, preceded by two frames and followed by one frame, in which the disks were presented without dots. In trials with two disks, the disks were presented in the same positions in all frames. In trials with three disks, a third disk with a white dot was added alternately to the left and right. The disks were then perceived to move as a coherent group, alternately to the left and right, by exactly one interstimulus distance (2.5°). When only two disks were presented, the disks appeared to flicker at the same position.

From frame to frame, the dots were displaced within the disks to induce apparent motion. The perceived dot motion depended on the number of disks and the

motion type (retinotopic, nonretinotopic) that was presented. When three disks were presented, the dot in the middle disk was perceived to rotate and the dots in the flanking disks were perceived to jump up-down or left-right in every second frame (*nonretinotopic* percept). When two disks were presented, the dot in the left disk was perceived to rotate and the dot in the right disk was perceived to jump up-down or left-right in every second frame (*retinotopic* percept). Importantly, the dot positions were identical in the two and three disk conditions (as introduced in Lauffs et al., 2018). The addition or omission of the third disk changed whether the dot motion was perceived in retinotopic or nonretinotopic coordinates, by changing the perceptual organization (i.e., whether the disks are perceived as moving left-right or stationary).

Results

We calculated the mean percentage of correct responses for each of the eight conditions. Group and individual means are shown in Figure 4. We subjected individual means to a 2 (dot motion: retinotopic, nonretinotopic) \times 2 (number of disks: 2, 3) \times 2 (color cue: present, absent) analysis of variance (ANOVA) and found a significant three-way interaction, $F(1, 9) = 61.23$, $p < 0.001$, which justified separate two-way ANOVAs for conditions with retinotopic and nonretinotopic dot motion. Other effects were also significant in the three-way ANOVA, but are not reported for brevity.

A 2 (number of disks: 2, 3) \times 2 (color cue: present, absent) ANOVA on percent correct for judgments of retinotopic dot motion found a main effect of the number of disks, $F(1, 9) = 61.71$, $p < 0.001$, and presence of a color cue, $F(1, 9) = 79.45$, $p < 0.001$. These main effects were modulated by a significant interaction, $F(1, 9) = 89.8$, $p < 0.001$. Inspection of conditions C1–C4 in Figure 4 suggests that judgments of retinotopic dot motion were poor when group motion was perceived (64% in C4), but highly precise in the remaining conditions (>95% in C1, C2, and C3). Possibly, the adoption of a reference frame based on group motion made the judgment of response-relevant retinotopic position difficult (condition C4), but when the color cue indicated the response-relevant retinotopic position (condition C3) this retinotopic cue was sufficient to alter the reference frame into a retinotopic one. Paired t tests confirmed that performance was better with than without color cue (condition C3 vs. C4, 95% vs. 64%), $t(9) = 10.35$, $p < 0.001$. The perception of retinotopic dot motion with color cue was indistinguishable from the perception of retinotopic dot motion without color cue (condition C1 vs. C2, 95% vs. 99%), $p = 0.168$.

Inspection of individual means in Figure 4 suggests that there was a ceiling effect in conditions C1, C2, and C3, which may compromise the normality of the data. By the Kolmogorov-Smirnov test, we confirmed that the data were not normally distributed in these conditions, $p_s < 0.004$. Therefore, we replaced the aforementioned t tests with a nonparametric test (related-samples Wilcoxon signed-rank test), but found the results to be unchanged.

The same 2×2 ANOVA was also performed on individual percent correct in the conditions with nonretinotopic dot motion (conditions C5–C8 in Figure 4). Performance was better with three than with two disks (84% vs. 67%), $F(1, 9) = 72.99$, $p < 0.001$, suggesting that group motion and its attendant non-retinotopic reference frame helped perceive the non-retinotopic dot motion. Further, the main effect of color cue, $F(1, 9) = 11.31$, $p = 0.008$, showed that the color cue improved performance (80% vs. 70%). There was no interaction, $p = 0.313$, suggesting that the differences in the effect of color, which are apparent in Figure 4, were not reliable. Separate paired t tests confirmed that the color cue improved performance in the two-disk condition where performance was initially poor (60% vs. 74%), $t(9) = 2.47$, $p = 0.036$, but also in the three-disk condition where performance was initially good because of group motion (80% vs. 88%), $t(9) = 5.67$, $p < 0.001$.

Discussion

We measured the effect of a color cue on the perception of retinotopic and nonretinotopic dot motion. Whereas the retinotopic disk motion was masked by group motion in the same color condition, adding the color cue led to a strong increase in performance (see conditions C3 and C4 in Figure 4), suggesting that color either reduced the conflicting group motion (Hein & Moore, 2012) or facilitated the attentional tracking of retinotopic motion. For *non-retinotopic* dot motion, color improved performance in the condition benefitting from the intuitive perception of group motion (condition C7 vs. C8). The improvement may result either from improved perception of group motion (Hein & Moore, 2012) or from enhanced attentional tracking. Finally, color also improved perception of nonretinotopic motion with two stationary disks (condition C5 vs. C6). The condition with two stationary disks relies exclusively on attentional tracking because no other cue is available. With a colored disk, it is likely that attentional tracking of dot motion was facilitated. However, it may also be that the colored disk helped to establish spatio-temporal correspondence across stimulus frames.

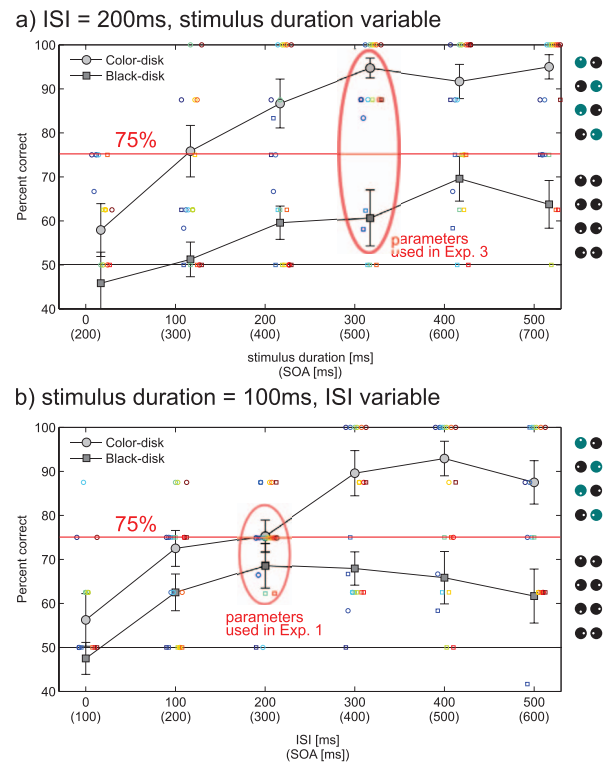


Figure 5. Performance with nonretinotopic dot rotation in Experiment 2. We varied either the stimulus duration (a) or the ISI (b). When the tracked disk was distinctly colored, performance improved from chance-level to near-perfect performance. When both disks were black, performance did not exceed 70% correct. Colored data points represent individual participants' performance. Error bars represent one SEM.

Experiment 2

In Experiment 1, we showed that perception of both the retinotopic and nonretinotopic rotation improved with color cues. In Experiment 2, we evaluated the timing parameters that lead to the maximal benefit of color before we ran the critical comparisons in Experiment 3. In one condition, we parametrically varied the stimulus duration and kept the ISI duration constant. In another condition, we varied the ISI duration and kept the stimulus duration constant.

Methods

The methods for Experiment 2 were identical to Experiment 1 with the following exceptions. We only used the 2-disk stimulus and instructed observers to track the nonretinotopic dot rotation, starting with the left disk in the first frame (see Figure 5 and Supplementary Movies S8–S32). In the condition with fixed ISI, the stimulus duration was randomly varied from trial to trial (17–517 ms in 100 ms steps) and the

ISI was fixed at 200 ms. In the condition with fixed stimulus duration, the ISI duration was randomly varied from trial to trial (0–500 ms in 100 ms steps) and the stimulus duration was fixed at 100 ms. For comparison with research by Verstraten et al. (2000), we also calculated the SOA (= stimulus duration + ISI). The SOA varied from 217 ms to 617 ms with fixed ISI and from 100 ms to 600 ms with fixed stimulus duration. The different stimulus/ISI durations and the directions of the retinotopic and nonretinotopic rotations (clockwise, counterclockwise) were used in a balanced $6 \times 2 \times 2$ factorial design and presented in random order. The rotation started randomly in a 3, 6, 9, and 12 o'clock orientation with equal probability. Per condition, 48 trials were presented (i.e., eight trials per stimulus duration/ISI). Each condition was first run with a color cue, followed by an identical block without color cue. The order of conditions was counterbalanced across participants.

Results and discussion

The ability to correctly indicate the nonretinotopic dot rotation increased with ISI and stimulus duration (Figure 5). Because the SOAs were not the same in conditions with fixed ISI and fixed stimulus duration, and SOA is key to attentional tracking of apparent motion (Verstraten et al., 2000), separate ANOVAs were carried out.

A 2 (tracking cue: color, none) \times 6 (stimulus duration: 17, 117, 217, 317, 417, 517 ms) on individual means from the condition with fixed ISI of 200 ms showed that performance was better with tracking of the color cue compared to tracking without the color cue (84% vs. 58%), $F(1, 9) = 70.56$, $p < 0.001$. The effect of stimulus duration, $F(5, 45) = 13.86$, $p < 0.001$, showed that performance increased from 52% at the shortest to 79% at the longest duration. The interaction was not significant, $p = 0.33$, suggesting that the rate at which performance increased with increasing stimulus duration (the slope of the curves in Figure 5a) was not reliably different between the two tracking conditions.

Another 2 (cue: color, none) \times 6 (ISI: 0, 100, 200, 300, 400, 500 ms) on individual means from the condition with fixed stimulus duration of 100 ms (see Figure 5b) showed better performance with one colored disk than with two black disks (79% vs. 62%), $F(1, 9) = 14.92$, $p = 0.004$, and increasing performance with ISI (from 52% at the shortest to 75% at the longest ISI), $F(5, 45) = 12.51$, $p < 0.001$. Additionally, there was an interaction of cue and ISI, $F(5, 45) = 2.97$, $p = 0.021$, confirming that the difference between color and no cue condition increased from 8% at the shortest ISI to 26% at the longest ISI.

These results indicate that a color cue improved the perception of nonretinotopic dot motion. Without color cue, however, performance did not exceed 70%, indicating that perception of nonretinotopic motion was severely limited, even with long periods of time between stimulus frames.

In Figure 5, we added a horizontal line to mark 75% correct responses, which were used as threshold in Verstraten et al. (2000). For the condition with color, the stimulus duration corresponding to 75% correct responses was around 100 ms (i.e., SOA of 300 ms; see Figure 5a). The ISI corresponding to 75% correct responses was around 200 ms (i.e., SOA of 300 ms, Figure 5b). Thus, we estimate the rate of apparent motion resulting in 75% correct to be roughly 3 Hz, which is at the lower end of the tracking limits reported in Verstraten et al. (2000).

Concerning the selection of an SOA that maximizes differences between attentional tracking with and without color cue, the experiment did not provide a clear answer. In particular, there was no statistical evidence for an interaction between time interval and color cue when the ISI was fixed (see Figure 5a). For lack of a better criterion, we selected the interval with the descriptively largest difference between conditions, the stimulus duration of 300 ms and the ISI of 200 ms (i.e., SOA of 500 ms). Finally, it is noteworthy that tracking of nonretinotopic motion without any external cue (i.e., with the black dots) was never better than 70%. Thus, the attentional selection of alternating horizontal stimulus positions was poor, but could be improved when color established spatiotemporal correspondence.

Experiment 3

In Experiment 3, we investigated the perception of nonretinotopic dot motion in more detail. Experiment 1 showed that making the response-relevant disk salient by means of a color cue improved the perception of nonretinotopic dot motion (see conditions C5-C8 in Figure 4). Because the salient color cue also changed spatiotemporal correspondence, improved performance could not be attributed to attentional tracking alone. In Experiment 3, the conditions of interest featured two colored disks of equal saliency (with physically isoluminant colors). Thus, attentional selection of one of the two colors was necessary. In contrast, it was no longer possible to establish spatiotemporal correspondence based on saliency. Rather, the response-relevant, but inconspicuous color had to be attentionally tracked. In another condition, we presented two equally salient shapes (with equal surface area) instead of two colors. The central question was whether attentional tracking of color or shape would improve

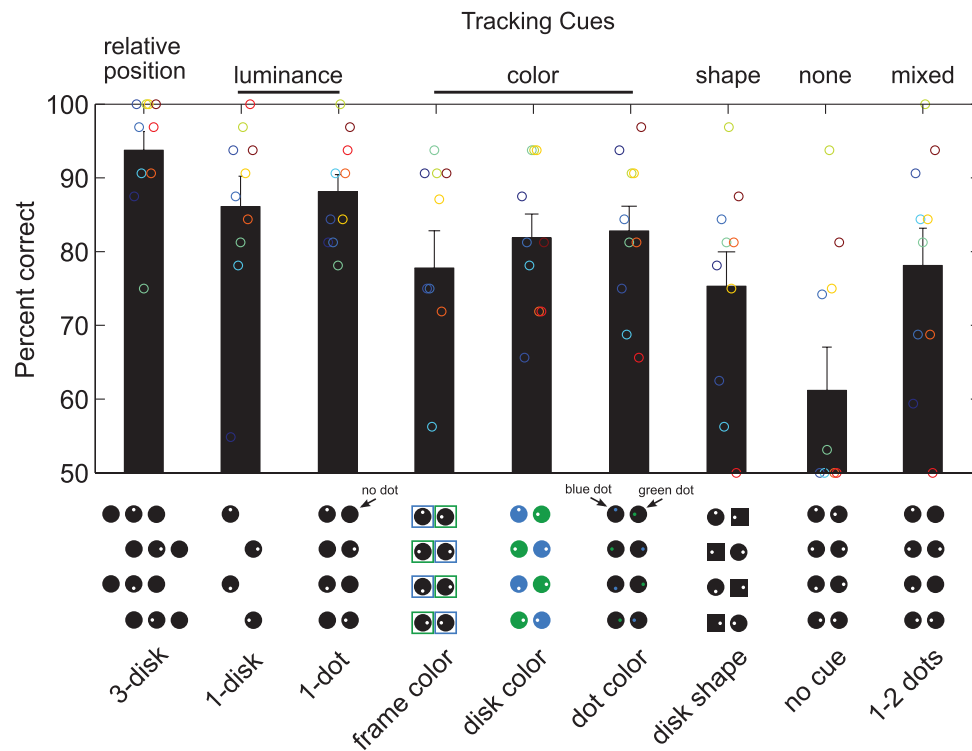


Figure 6. Discrimination of nonretinotopic dot rotation direction in Experiment 3. We varied the tracking cues to investigate whether attentional tracking of color or shape was as efficient as luminance-defined cues or group motion. Participants were asked to report the nonretinotopic rotation in all conditions. Error bars indicate one *SEM*. Colored circles depict the mean performance of the individual observers.

perception of nonretinotopic motion relative to the condition without external cues. Further, we asked whether nonretinotopic motion perception would reach similar levels for salient stimuli, requiring little attentional tracking, as for inconspicuous elements that depended on attentional tracking. A single dot on two stationary disks and a single dot on a single disk were used to induce spatiotemporal correspondence by saliency. We refer to these cues as luminance-defined cues. Finally, we repeated the condition with group motion to evaluate whether attentional tracking or luminance-defined cues allow for similar levels of nonretinotopic motion perception.

Methods

The methods for Experiment 3 were identical to Experiment 1, unless noted otherwise. We used a stimulus duration of 300 ms and an ISI of 200 ms. Ten new, naïve observers participated in the experiment (mean age: 21.9 years; $SD = 2$; range: 19–25; six female; nine right-handed; seven right eye dominance). All observers had normal color vision as tested with the Ishihara test for color deficiency (Ishihara, 1987). Observers were instructed to report the direction of the nonretinotopic dot rotation (clockwise vs. counter-

clockwise), starting with the left disk in the first stimulus presentation of each trial. We again used the movies provided in the supplementary materials to instruct the participants (see Figure 6 and Supplementary Movies S33–S45). The conditions were presented in random order, with one block of 32 trials each. Retinotopic and nonretinotopic rotation directions and the initial orientation of the rotating dot (3, 6, 9, 12 o'clock) were counterbalanced in a randomized full-factorial design for each observer and block. Before the experiment, each observer performed one practice block of 16 trials using the same stimulus as in Experiment 1. Auditory feedback for incorrect responses was provided only during the practice block.

Group motion

Three black disks moved left and right in concert. Only the middle disk contained a white dot that performed a nonretinotopic rotation. This condition is the baseline condition and referred to as “three-disk” condition.

Luminance

In one condition, a single black disk moved left and right and contained a white dot that performed a

nonretinotopic rotation (one-disk). In another condition, one of two disks had a rotating dot (one-dot). The dot was presented in the left disk in uneven numbered frames and in the right disk in even numbered frames.

Color

Participants tracked the color of the disk that was shown on the left position when the dot appeared. The color of the tracked dot remained the same in a block of trials. In one condition, each disk was enclosed by a square-shaped frame, which was blue for one and green for the other disk (physically isoluminant, both 45 cd/m²). In another condition, the disks were not black, but one was blue and the other green. In a third condition, both disks were black, but one dot was blue and the other green.

Shape

Instead of two disks, we used one disk and one square. Both were black with white dots and had the same surface area. Participants tracked the shape on the left when the dots appeared. The shape stayed the same in a block of trials.

No cue

Two stationary black disks were presented. Both disks had white dots in all frames. In addition to the nonretinotopic rotation, a retinotopic rotation could be perceived in the left disk, and a retinotopic up-down or left-right dot motion in the right disk.

Mixed

This condition was similar to the luminance condition with a single disk, except that the left disk had an additional dot in the even numbered frames. These additional dots were positioned so that a retinotopic rotation of the left disk could be perceived.

Results

The mean percentages of correct responses for the perception of nonretinotopic dot rotation are presented in Figure 6. We had created several versions of the luminance and color cue condition because we were unsure which would be most effective. To test for eventual differences, we performed a one-way ANOVA on the three-color cue conditions, but found no significant effect, $F(2, 18) = 1.48$, $p = 0.254$. We therefore collapsed the color conditions. Similarly, we found no difference between the two luminance conditions, $t(9) = 0.58$, $p = 0.576$, and, therefore,

collapsed these conditions as well. To evaluate effects of attentional tracking, we compared the color cue condition to the group motion and luminance cue conditions. Results from t tests were confirmed by nonparametric Wilcoxon tests. Performance with color cues was worse than with group motion (81% vs. 94%), $t(9) = 2.81$, $p = 0.021$, but better than without cues (81% vs. 61%), $t(9) = 3.71$, $p = 0.005$. Importantly, there was no significant difference between color and luminance cues (81% vs. 87%), $t(9) = 1.25$, $p = 0.244$, showing that attentional tracking of color was as efficient as spatiotemporal correspondence established by luminance cues. Performance with attentional tracking of shape was nonsignificantly worse than tracking of color, $t(9) = 2.13$, $p = 0.062$. Possibly, it was more difficult to discriminate between the two shapes in peripheral vision than between the two colors. Further, we explored performance in a condition in which the presence and absence of external cues alternated (one to two dots) and found performance (78%) to be in the range of performance with external luminance, color, or shape cues (75%-87%).

Discussion

Experiment 3 showed that the perception of nonretinotopic dot motion is best in the Ternus-Pikler display with group motion. Perception of nonretinotopic motion with other cues, such as luminance, color, or shape was worse, suggesting that group motion is a powerful cue to nonretinotopic motion perception. We were interested in isolating attentional tracking in nonretinotopic motion perception. As in Experiments 1 and 2, performance in the no-cue condition was poor, showing that attentional tracking of position alone was difficult. However, the drop in performance from group motion to the no-cue condition may underestimate the efficiency of attentional tracking because external cues beyond stimulus position were missing. After all, group motion is established by the Gestalt principle of “common fate”, which is a powerful cue beyond disk position, whereas only disk position was available in the no-cue condition. Therefore, we created conditions where observers had to attentionally track a color or a shape when two equally salient colors were available. Although spatiotemporal correspondence could be established by the Gestalt principle of similarity, it was nonetheless necessary to select one of the two colors to accomplish the task. Thus, the color conditions isolate attentional tracking, but provide an external cue beyond disk position. We found that attentional tracking worked as well as luminance-defined cues, showing that attentional tracking may contribute to nonretinotopic motion perception. However, performance with color (or shape) cues was worse than with

group motion, suggesting that nonretinotopic motion perception cannot be entirely accounted for by attentional tracking. Further, performance with luminance-defined cues was also worse than with group motion. In the one-disk condition, only a single disk was shown so that establishing spatiotemporal correspondence was easy. However, performance was actually worse than with group motion, suggesting that the reference frame created by the outer disks facilitated perception compared to a single object.

General discussion

Elements are grouped into wholes by a variety of Gestalt laws such as common fate, similarity, proximity, good continuation, etc. The two-stage model that we consider here (Figure 2) proposes that each group is endowed by a reference frame that guides the attribution of stimulus features according to group identities. For example, consider the phenomenon of crowding, which refers to the failure to recognize visual stimuli when flanked by other stimuli (Andriessen & Bouma, 1976). The target stimulus itself is visible but features of the target and adjacent stimuli (flankers) appear mixed up (Pelli, Palomares, & Majaj, 2004). However, we showed that when the flankers do not belong to the same group as the target, the target itself becomes easily identifiable (Herzog & Manassi, 2015; Herzog, Sayim, Chicherov, & Manassi, 2015; Saarela & Herzog, 2008). One possible explanation for this effect is that segregating the target and flankers into two distinct groups creates two different reference frames so that features of the flankers are not attributed to the target and vice versa.

For dynamic stimuli, perceptual grouping occurs in space *and* time and the resulting spatiotemporal reference frames determine how features are attributed within perceptual groups. In sequential metacontrast, stimuli are grouped into different motion streams and the processing of features of individual elements depend on group membership of the elements (Herzog, Otto, & Ögmen, 2012; Otto, Ögmen, & Herzog, 2006; Otto, Ögmen, & Herzog, 2009). Vernier offsets of elements within the same group (motion stream) are integrated, whereas Vernier offsets of elements in different groups are not, regardless of spatiotemporal proximity (Otto et al., 2006). Similarly, as shown here and in previous work (Boi et al., 2009; Lauffs et al., 2018), in the Ternus-Pikler display, which pits retinotopic reference frames against grouping-based nonretinotopic reference frames, feature processing depends on spatiotemporal grouping and the attendant reference frame.

As shown in Figure 2, the choice of reference frames can also be influenced by attentional mechanisms. Attention can modulate and alter perceptual groups, especially if they are ambiguous. Previous research suggests that a reference-frame can also be established based on the spatiotemporal trajectory of focal attention (Cavanagh, 1992; Lu & Sperling, 1995; Verstraten et al., 2000). Here, we investigated whether cues such as color, shape, or luminance facilitate motion perception on the to-be-tracked object. In Experiment 1, we focused on two conditions that typically result in poor performance. First, it is difficult to perceive the retinotopic dot when there is group motion (see condition C4 in Experiment 1; e.g., Boi et al., 2009; Clarke, Ögmen, & Herzog, 2016; Lauffs et al., 2017; Lauffs et al., 2018). However, despite being invisible, retinotopic motion may strongly interfere with perception of nonretinotopic motion (Lauffs et al., 2018). Second, it is hard to perceive nonretinotopic dot motion with two stationary disks (see condition C6 in Experiment 1). We found that color cues improved performance in both cases, suggesting that attentional tracking may contribute to the perception of retinotopic and nonretinotopic motion perception. However, an alternative account in terms of changes in the perception of group or element motion (Hein & Moore, 2012) cannot be ruled out.

One argument against a strong role of attentional tracking is that perception of nonretinotopic motion with two stationary disks, which relies exclusively on attentional tracking, is poor (condition C6 in Experiment 1, no-cue conditions in Experiments 2 and 3). It may be that the perception of retinotopic motion inside the two stationary disks interfered with tracking the nonretinotopic motion between the left and right position. When a color cue was added (condition C5 vs. C6 in Experiment 1, color cue conditions in Experiments 2 and 3), performance improved moderately, but was far from ceiling. Similarly, performance improved moderately when we added a color cue with nonretinotopic motion perception (condition C7 vs. C8 in Experiment 1). The latter improvement may arise from attentional tracking or an enhancement of the group motion. Whatever the exact mechanism(s) at work, the effects of the color cue are limited. As a further case in point, Experiment 3 showed that nonretinotopic motion perception was worse with any of the tracking cues compared to group motion. Thus, our experiments point to a limited role for attentional tracking, but cannot decide whether nonretinotopic motion perception with group motion arises from a motion processor after group motion is established or results directly from an attentional tracking. It was just recently found that even smooth motion percepts can result from tracking (Allard & Arleo, 2016).

The flicker rate in our paradigm was well within the limits of attentional tracking (Verstraten et al., 2000). Our results are surprising when put into the context of the existent literature on multiple object tracking (Cavanagh & Alvarez, 2005; Meyerhoff, Papenmeier, Jahn, & Huff, 2013; Pylyshyn & Storm, 1988; Vater, Kredel, & Hossner, 2016), where it has been shown that observers can track up to four disks over several seconds even when the disks cross the trajectories of a large number of distractor disks. In the light of powerful attentional tracking of objects undergoing smooth motion, it is surprising that observers can hardly track one disk in the two-disk condition without the color cue (e.g., condition C6 in Experiment 1, no-cue condition in Experiments 2 and 3). This suggests that the spatiotemporal trajectory of focal attention in isolation plays a relatively weak role in the first stage of the two-stage model (Figure 2). Instead, interactions between attention and other grouping cues, seem to provide a much stronger basis for determining the reference frame underlying stimulus processing.

Surprisingly, Experiment 2 showed that increasing the ISI or duration in this condition did not change the results substantially. Even when the ISI was 500 ms, attentional tracking was almost impossible when external cues were absent. In this case, the entire sequence lasted for 2.2 s; retinotopic motion still prevailed. With the addition of the color cue, ceiling performance was reached for SOAs of 400 to 500 ms. We suggest that built-in motion routines have a privileged role in establishing *spatiotemporal* groups. This leaves a limited role for attentional tracking for substantial times of processing. The primacy of motion in transforming retinotopic coordinates into nonretinotopic ones makes ecological sense. Due to our own movements and those of external objects, we receive highly dynamic stimuli according to retinotopic coordinates. Hence, motion is a relevant, abundant, and readily available feature. However, color and spatial cues also play an important role in perceptual grouping (Figure 2). They can override motion cues and allow tracking, especially when they are congruent with the task-dependent trajectory of attentional tracking. Whereas the role of attentional tracking is limited, we showed that attention operates after the reference frames are established, i.e. in nonretinotopic coordinates (Experiment 3; Boi et al., 2009; Boi, Vergeer, Ögmen, & Herzog, 2011; Scharnowski, Hermens, Kammer, Ögmen, & Herzog, 2007).

Taken together, our results add evidence that attention and grouping are different but interacting processes. Our study highlights the importance of reference frames and supports the two-stage model (Figure 2). The model clarifies how various features and cues can work in conjunction or in competition to determine prevailing groups. These groups in turn

establish reference frames according to which features are processed and bound together. Attention allows the selection from a variety of possible groupings, the grouping that fits best on the observer's internal state and goals, especially when Gestalt laws produce groups of similar strengths.

Keywords: nonretinotopic processing, reference frame, attentional tracking, motion perception

Acknowledgments

M. M. Lauffs, O. H. Choung, and M. H. Herzog were supported by the Swiss National Science Foundation 320030_176153 “Basics of visual processing: from elements to figures.” D. Kerzel was supported by the Swiss National Science Foundation 100019_182146.

*Marc M. Lauffs and Oh-Hyeon Choung contributed equally to this manuscript.

Commercial relationships: none.

Corresponding author: Oh-Hyeon Choung.

Email: oh-hyeon.choung@epfl.ch.

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

References

- Allard, R., & Arleo, A. (2016). Position-based vs energy-based motion processing. *Journal of Vision*, *16*(12): 670, <https://doi.org/10.1167/16.12.670>. [Abstract]
- Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, *16*(1), 71–78, [https://doi.org/10.1016/0042-6989\(76\)90078-X](https://doi.org/10.1016/0042-6989(76)90078-X).
- Aydin, M., Herzog, M. H., & Ögmen, H. (2011). Attention modulates spatio-temporal grouping. *Vision Research*, *51*(4), 435–446, <https://doi.org/10.1016/j.visre.2010.12.013>.
- Bach, M. (1996). The Freiburg Visual Acuity Test—Automatic measurement of visual acuity. *Optometry and Vision Science*, *73*(1), 49–53.
- Beck, J. (1966a). Slant and shape variables in perceptual grouping. *Science*, *154*, 538–540.
- Beck, J. (1966b). Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, *1*(5), 300–302.

- Beck, J. (1967). Perceptual grouping produced by line figures. *Perception & Psychophysics*, 2(11), 491–495.
- Boi, M., Ögmen, H., Krümmenacher, 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, 1–11, <https://doi.org/10.1167/9.13.5>.
- Boi, M., Vergeer, M., Ögmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology*, 21(20), 1732–1737, <https://doi.org/10.1016/j.cub.2011.08.059>.
- Caelli, T., & Julesz, B. (1979). Psychophysical evidence for global feature processing in visual texture discrimination. *Journal of the Optical Society of America*, 69(5), 675–678, <https://doi.org/10.1364/JOSA.69.000675>.
- Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257(5076), 1563–1565.
- Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Science*, 9(7), 349–354, <https://doi.org/10.1016/j.tics.2005.05.009>.
- Clarke, A. M., Ögmen, H., & Herzog, M. H. (2016). A computational model for reference-frame synthesis with applications to motion perception. *Vision Research*, 126, 242–253, <https://doi.org/10.1016/j.visres.2015.08.018>.
- 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), 975–988, <https://doi.org/10.1037/a0028197>.
- Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: a fresh look at object recognition. *Current Opinion in Behavioral Sciences*, 1, 86–93, <https://doi.org/10.1016/j.cobeha.2014.10.006>.
- Herzog, M. H., Otto, T. U., & Ögmen, H. (2012). The fate of visible features of invisible elements. *Frontiers in Psychology*, 3, 119, <https://doi.org/10.3389/fpsyg.2012.00119>.
- Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *Journal of Vision*, 15(6), 5, <https://doi.org/10.1167/15.6.5>.
- Ishihara, S. (1987). *Test for colour-blindness*. Tokyo, Japan: Kanehara.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews: Neuroscience*, 2(3), 194–203, <https://doi.org/10.1038/35058500>.
- Julesz, B. (1991). Early vision and focal attention. *Reviews of Modern Physics*, 63(3), 735, <https://doi.org/10.1103/RevModPhys.63.735>.
- Koffka, K. (1922). Perception: an introduction to the Gestalt-Theorie. *Psychological Bulletin*, 19(10), 535.
- Köhler, W., & Adams, P. A. (1958). Perception and attention. *The American Journal of Psychology*, 71(3), 489–503.
- Krechevsky, I. (1938). An experimental investigation of the principle of proximity in the visual perception of the rat. *Journal of Experimental Psychology*, 22(6), 497, <https://doi.org/10.1037/h0058982>.
- Lauffs, M. M., Choung, O.-H., Ögmen, H., & Herzog, M. H. (2018). Unconscious retinotopic motion processing affects non-retinotopic motion perception. *Consciousness and Cognition*, 62, 135–147, <https://doi.org/10.1016/j.concog.2018.03.007>.
- Lauffs, M. M., Ögmen, H., & Herzog, M. H. (2017). Unpredictability does not hamper nonretinotopic motion perception. *Journal of Vision*, 17(9):6, 1–10, <https://doi.org/10.1167/17.9.6>. [PubMed] [Article]
- Lu, Z.-L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature*, 377, 237, <https://doi.org/10.1038/377237a0>.
- Meyerhoff, H. S., Papenmeier, F., Jahn, G., & Huff, M. (2013). A single unexpected change in target-but not distractor motion impairs multiple object tracking. *i-Perception*, 4(1), 81–83, <https://doi.org/10.1068/i0567sas>.
- Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance*, 23(2), 339, <https://doi.org/10.1037/0096-1523.23.2.339>.
- Ögmen, 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, <https://doi.org/10.1109/JPROC.2009.2039028>.
- Otto, T. U., Ögmen, 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), 1079–1086, <https://doi.org/10.1167/6.10.7> [PubMed] [Article].
- Otto, T. U., Ögmen, H., & Herzog, M. H. (2009). Feature integration across space, time, and orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 35(6), 1670, <https://doi.org/10.1037/a0015798>.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal*

- of Vision, 4(12):12, 1136–1169, <https://doi.org/10.1167/4.12.12>. [PubMed] [Article]
- Proffitt, D. R., & Cutting, J. E. (1980). An invariant for wheel-generated motions and the logic of its determination. *Perception*, 9, 435–449.
- Proffitt, D. R., Cutting, J. E., & Stier, D. M. (1979). Perception of wheel-generated motions. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 289–302.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179–197.
- Saarela, T. P., & Herzog, M. H. (2008). Time-course and surround modulation of contrast masking in human vision. *Journal of Vision*, 8(3):23, 1–10, <https://doi.org/10.1167/8.3.23>. [PubMed] [Article]
- Scharnowski, F., Hermens, F., Kammer, T., Ögmen, H., & Herzog, M. H. (2007). Feature fusion reveals slow and fast visual memories. *Journal of Cognitive Neuroscience*, 19(4), 632–641, <https://doi.org/10.1162/jocn.2007.19.4.632>.
- Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychological Review*, 45, 1–41, <https://doi.org/10.1037/h0062733>.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 194, <https://doi.org/10.1037/0096-1523.8.2.194>.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136, [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5).
- Vater, C., Kredel, R., & Hossner, E.-J. (2016). Detecting single-target changes in multiple object tracking: The case of peripheral vision. *Attention, Perception, & Psychophysics*, 1–16, <https://doi.org/10.3758/s13414-016-1078-7>.
- Verstraten, F. A. J., Cavanagh, P., & Labianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research*, 40(26), 3651–3664, [https://doi.org/10.1016/S0042-6989\(00\)00213-3](https://doi.org/10.1016/S0042-6989(00)00213-3).
- Wagemans, J. (2015). *The Oxford handbook of perceptual organization*. Oxford, UK: Oxford University Press, <https://doi.org/10.1093/oxfordhb/9780199686858.001.0001>.
- World Medical Association (2013). “Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects”. *JAMA*, 310(20), 2191–2194, <https://doi.org/10.1001/jama.2013.281053>.