

Object identity determines trans-saccadic integration

Leila Drissi-Daoudi

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



Guido Marco Cicchini

Consiglio Nazionale delle Ricerche (CNR), Institute of Neuroscience, Pisa, Italy



Humans make two to four rapid eye movements (saccades) per second, which, surprisingly, does not lead to abrupt changes in vision. To the contrary, we perceive a stable world. Hence, an important question is how information is integrated across saccades. To investigate this question, we used the sequential metacontrast paradigm (SQM), where two expanding streams of lines are presented. When one line is spatially offset, the other lines are perceived as being offset, too. When more lines are offset, all offsets integrate mandatorily; that is, observers cannot report the individual offsets but perceive one integrated offset. Here, we asked observers to make a saccade during the SQM. Even though the saccades caused a highly disrupted motion trajectory on the retina, offsets presented before and after the saccade integrated mandatorily. When observers made no saccade and the streams were displaced on the screen so that a similarly disrupted retinal image occurred as in the previous condition, no integration occurred. We suggest that trans-saccadic integration and perception are determined by object identity in spatiotopic coordinates and not by the retinal image.

(saccades) per second to bring the most interesting aspects of a visual scene onto the fovea, where resolution is highest. A straightforward question is how information from several saccades is integrated into a stable representation and how the effects of the saccades are discounted (for reviews, see Bridgeman, 2011; Melcher, 2011; Melcher & Morrone, 2015; Wurtz, 2008).

According to early theories of trans-saccadic integration, the nervous system *compensates* for the effects of eye movements by relocating retinal inputs using an efference copy (von Helmholtz, 1896). In stark contrast to these theories, it was proposed that no integration across saccades takes place at all and that vision starts anew with each saccade (Bridgeman & Mayer, 1983; Rayner & Pollatsek, 1983). For example, even large changes within an image go unnoticed when they occur during a saccade (Bridgeman & Mayer, 1983). Likewise, participants experience difficulties fusing pre-saccadic and post-saccadic items even when they belong to the same meaningful configuration (Irwin, 1991; Jonides, Irwin, & Yantis, 1983); however, more recent findings have shown evidence for trans-saccadic transfer of information (Demeyer, De Graef, Wagemans, & Verfaillie, 2010; Fornaciai, Binda, & Cicchini, 2018; Harrison & Bex, 2014; Melcher, 2005; Melcher, 2007; Melcher & Morrone, 2003; Paeye, Collins, & Cavanagh, 2017; Stewart & Schütz, 2018; Wittenberg, Bremmer, & Wachtler, 2008; Wolfe & Whitney, 2015; Zimmermann, Burr, & Morrone, 2011). It has even been shown that trans-saccadic integration is optimal or near-optimal; that is, pre- and post-saccadic information is weighted according to

Introduction

The early visual system is organized retinotopically; that is, neighboring points in the environment are projected to neighboring points on the retina, and this topological organization is preserved through early visual areas (e.g., Engel, Glover, & Wandell, 1997). Humans make two to four rapid eye movements

Citation: Drissi-Daoudi, L., Öğmen, H., Herzog, M. H., & Cicchini, G. M. (2020). Object identity determines trans-saccadic integration. *Journal of Vision*, 20(7):33, 1–13, <https://doi.org/10.1167/jov.20.7.33>.



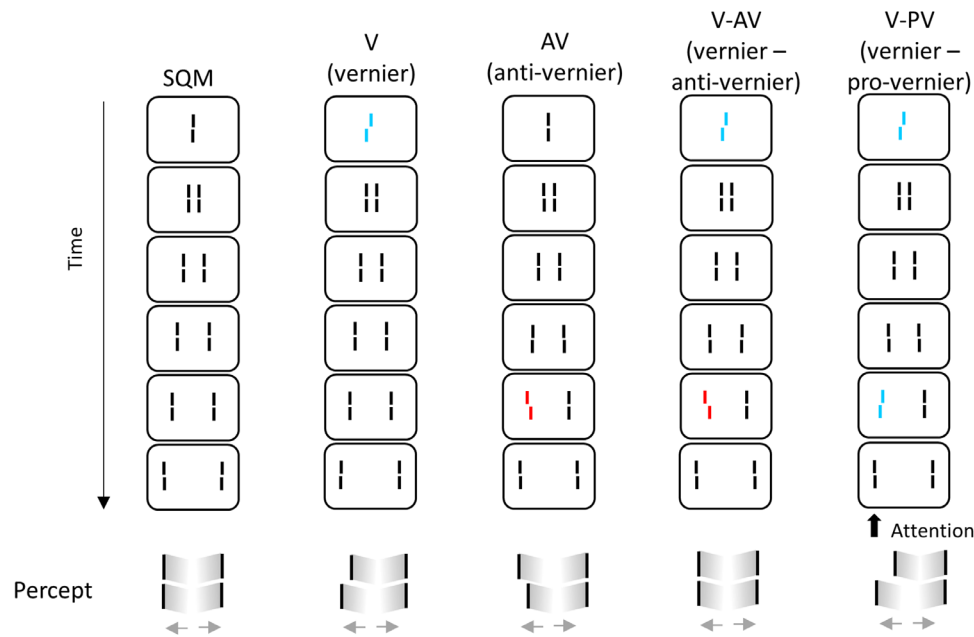


Figure 1. The SQM paradigm. A central line is followed by pairs of flanking lines. A percept of two moving streams diverging from the center is elicited. Observers attend to one of the streams (here, the left stream) and report the perceived offset direction (right/left) by pressing hand-held push buttons. In condition V (vernier), only the central line is offset; the offset is visible at the following lines and observers report the offset direction. In condition AV (anti-vernier), only a flanking line is offset; the offset is visible in the attended stream. In condition V-AV (vernier-anti-vernier), the central line and one of the flanking lines are offset in opposite directions and cancel each other; observers cannot report the individual vernier offsets. In condition V-PV (vernier-pro-vernier), the central line and one of the flanking lines are offset in the same direction and add up. Figure adapted from [Drissi-Daoudi et al. \(2019\)](#).

its reliability and integrated in a statistically optimal way ([Ganmor, Landy, & Simoncelli, 2015](#); [Hübner & Schütz, 2017](#); [Wijdenes, Marshall, & Bays, 2015](#); [Wolf & Schütz, 2015](#)). It has been proposed that information transfer across saccades occurs to establish object correspondence and maintain perceptual stability ([Deubel, Schneider, & Bridgeman, 1996](#); [Poth, Herwig, & Schneider, 2015](#); [Tas, Moore, & Hollingworth, 2012](#); [Tas, Moore, & Hollingworth, 2014](#)). Obviously, information survives saccades, and there must be trans-saccadic integration, updating, or similar mechanisms because otherwise we would never be able to perceive smooth motion trajectories of objects when we make a saccade.

In most trans-saccadic studies, the objects are static, and object continuity is the assumption. What happens if the object is also moving? Look at a moving car and move your eyes. A smooth motion trajectory is perceived even though on the retina the smoothness of the trajectory is strongly disrupted by the saccade. The pre-saccadic and post-saccadic retinal images are at very different locations. Experimentally, studies have found spatiotopic integration of motion and apparent motion across a saccade ([Fabius, Fracasso, & Van der Stigchel, 2016](#); [Fracasso, Caramazza, & Melcher, 2010](#); [Szinte & Cavanagh, 2011](#); [Zimmermann, Morrone,](#)

& Binda, 2018). In particular, [Melcher \(2008\)](#) found trans-saccadic transfer of a tilt adaptation after-effect when pre- and post-saccadic locations were linked by a continuously moving object.

Using the sequential metacontrast paradigm (SQM) ([Otto, Öğmen, & Herzog, 2006](#); [Otto, Öğmen, & Herzog, 2009](#)), we recently showed that features integrate mandatorily and unconsciously within a window of integration that can last for up to 450 ms, depending on observer ([Drissi-Daoudi, Doerig, & Herzog, 2019](#)). Here, we show that this unconscious, mandatory integration cannot be disrupted by a saccade in line with trans-saccadic integration to maintain object identity across eye movements.

In the SQM, a central line is followed by pairs of flanking lines ([Figure 1](#)). A percept of two moving streams diverging from the center is elicited. The central line is invisible because it is masked by the following lines. Surprisingly, if the central line is given a spatial offset (i.e., vernier, where the lower segment of the line is offset either to the right or to the left), the offset is perceived at the aligned flanking lines ([Figure 1](#), vernier condition [V]). If, in addition, a line in the stream is offset, the two offsets integrate. When the offsets are in opposite directions, they cancel each other ([Figure 1](#), vernier-anti-vernier condition [V-AV]). When

they are in the same direction, they add up (Figure 1, vernier–pro-vernier condition [V-PV]).

Here, we manipulated two independent variables, eye movements (fixation or saccade) and smoothness/continuity of object motion (retinotopically or spatiotopically smooth) in four different conditions. The fixation condition (Figure 2a, left) was similar to the classic SQM. Observers kept their eyes fixated on a fixation dot, and the SQM moved smoothly. Keeping the eyes steady created a smooth, continuous motion in both spatiotopic and retinotopic reference frames. In the second fixation condition, the jump condition (Figure 2a, right), the eyes were steady on the fixation point while the SQM halfway jumped upward, producing a displacement on the retina similar to the one caused by a vertical saccade. In this case, a discontinuous motion occurs both in spatiotopic and retinotopic coordinates. In two conditions, observers performed a saccade between the presentation of the central vernier and the flank vernier offsets (spatiotopic and retinotopic conditions) (Figure 2a, middle columns). In the spatiotopic condition, the SQM was displayed as in the fixation condition (i.e., in the center of the screen during the entire trial). Hence, the stimulus had a smooth motion in a spatiotopic reference frame but discontinuous motion in a retinotopic reference frame. In the retinotopic condition, the SQM was displaced when the saccade was detected in order to cancel the retinal displacement induced by the eye movement. In this case, the stimulus had a smooth and discontinuous motion in the retinotopic and spatiotopic reference frames, respectively.

We considered three hypotheses. According to the first two hypotheses, integration occurs based on object identity (i.e., smooth continuous motion) either in spatiotopic or in retinotopic coordinates, respectively. According to the third hypothesis, vision starts anew after each saccade.

- *First hypothesis*—If object identity in external (spatiotopic) coordinates determines feature integration in the SQM, we expect integration when the stimulus is smooth in spatiotopic coordinates (i.e., in the spatiotopic condition and in the fixation condition). However, we do not expect integration when the stimulus is continuous in retinotopic coordinates (i.e., the retinotopic condition).
- *Second hypothesis*—If integration is determined by the integrity of the SQM on the retina (i.e., according to continuity/smoothness in retinotopic coordinates), we expect the opposite result because the SQM is not displaced on the retina in the retinotopic condition.
- *Third hypothesis*—If processing starts anew with each saccade, we expect no integration in either saccadic condition. The two offsets are perceived separately.

The jump condition serves as a control condition. Independent of the hypotheses, we expect no integration because the SQM is displaced on both the retina and the screen.

To anticipate our results, we first found that features integrate in the SQM across saccades in both spatiotopic and retinotopic coordinates (Experiment 1). Hence, in a second experiment, we tested whether the integration is mandatory in both coordinate frames—that is, whether observers can report the individual offsets. We found that integration is mandatory in the spatiotopic condition but not in the retinotopic condition (Experiment 2). Overall, we suggest that trans-saccadic integration is determined by object identity in spatiotopic coordinates.

Methods

Observers

Data were obtained from paid students recruited from the Ecole Polytechnique Fédérale de Lausanne (EPFL) and one of the authors (LD-D, Experiment 1). Observers signed informed consents and had normal or corrected-to-normal vision as tested with the Freiburg visual acuity test (Bach, 1996). The experiments were undertaken with the permission of the local ethics committee and in accordance with the tenets of the Declaration of Helsinki. Eight observers took part in Experiment 1 (age, 21–29 years; four females). Six observers participated in Experiment 2 (age, 19–27 years; no females). Participants, except the author, were naïve to the specific purpose of the experiment.

Power analysis

The sequential metacontrast paradigm has been introduced in Otto et al. (2006), where Cohen's $d \approx 2.0$. To achieve a power of 90%, a sample size of five observers is needed. Here, eight and six observers participated in Experiment 1 and Experiment 2, respectively. The smallest effect size of a significant result in our experiments was 1.51 (Experiment 2: V-AV R[AV] compared to AV in the retinotopic condition). With a sample size of six observers, we achieved a power of 83.7%.

Apparatus

Stimuli were presented on a BenQ XL2540 LCD monitor (1920 × 1080 pixels, 240 Hz; BenQ, Taipei, Taiwan) using MATLAB (R2013a; MathWorks, Natick, MA) with Psychtoolbox (Brainard, 1997).

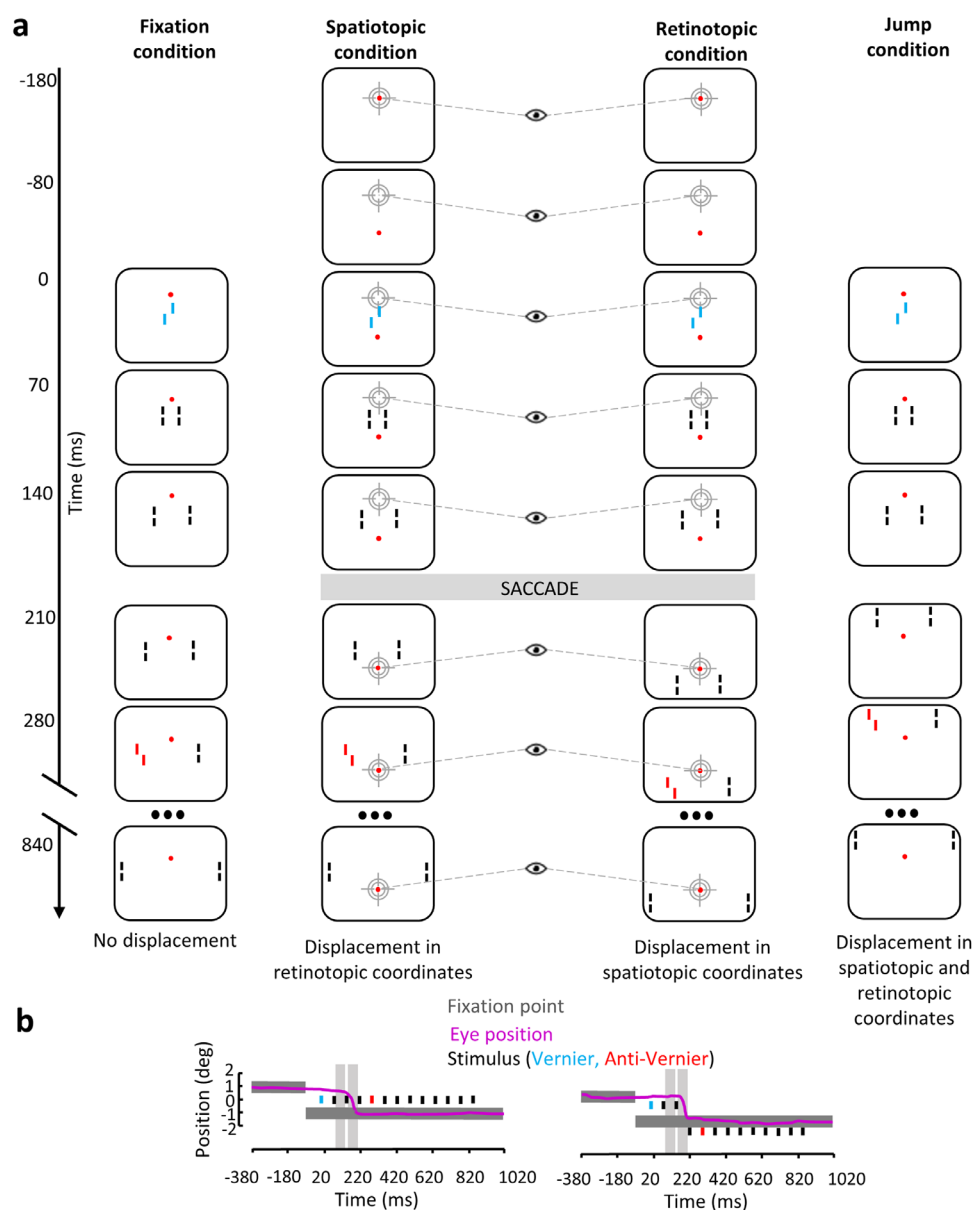


Figure 2. (a) In the fixation condition, participants fixated the fixation point displayed 1° above the center of the screen during the entire trial. The SQM was presented in the center of screen. Here is shown the V-AV condition in which the central line and the fourth flanking line have equal contribution and are offset in opposite directions. In the spatiotopic condition, observers were instructed to make a saccade to the fixation point as soon as it was displaced. The SQM was displayed in the center of the screen; thus, there was no displacement in spatiotopic coordinates compared to the fixation condition, but there was retinal displacement. In the retinotopic condition, observers were instructed to make a saccade to the fixation point as soon as it was displaced. The SQM was first presented in the center of the screen and then displayed 1° below the new location of the fixation point after the saccade was detected; thus, the retinal stimulation was the same as in the fixation condition, but there was displacement in spatiotopic coordinates. In the jump condition, observers fixated the fixation point displayed 1° above the center of the screen during the entire trial. The SQM was first presented in the center of the screen. At 210 ms SOA, the sequence continued 1° above the fixation point; thus, the SQM was displaced in both spatiotopic and retinotopic coordinates compared to the fixation condition. Only the vernier–anti-vernier configuration is shown. (b) Time course of events with a sample eye position trace (pink line) in the spatiotopic and retinotopic conditions. Black rectangles indicate presentation of the individual lines of the SQM. The vernier and anti-vernier offsets are represented in blue and red, respectively. Dark gray rectangles indicate presentation of the fixation point. As described in the Methods section, only saccades made during the blank periods (light gray areas) were kept for analysis; that is, no saccade-induced motion smear occurred. Colors are for illustration purposes only; stimuli were white on a black background.

Participants were seated in a well-lit room 2.50 meters from the screen. Eye movements were recorded with an SMI iViewX HiSpeed 1250 eyetracker (SensoMotoric Instruments, Tetlow, Germany), which was set up for binocular mode at 500-Hz sampling frequency. The eyetracker was calibrated before each block. Stimulus luminance was 98 cd/m^2 as measured with a Minolta LS-100 luminance meter (Osaka, Japan), and background luminance on the screen was below 1 cd/m^2 .

Stimuli

Stimuli were variations of the sequential metacontrast stimulus as introduced by Otto et al. (2006). The sequence started with a central line consisting of two vertical segments of $1200''$ (arcsec) length and separated by a vertical gap of $120''$. The central line was followed by 12 pairs of flanking lines presented one after the other farther away from the center (total stimulus duration, 840 ms). The length of the flanking lines was also $1200''$, separated by a vertical gap of $120''$. The horizontal distance between the central line and the first flanking lines, as well as between consecutive flanking lines, was $200''$. Each line was presented for 20 ms. The interstimulus interval (ISI) between the lines was 50 ms. A motion percept of two streams of lines diverging from the center was elicited. Observers attended to one of the streams. The central line and/or the fourth flanking line in the attended stream (280 ms after presentation of the central line) were offset; that is, the lower segment was offset either to the right or to the left with respect to the upper segment (vernier offset). Four configurations were presented: (1) *vernier* (V), where only the central line was offset; (2) *anti-vernier* (AV), where only the fourth flanking line was offset; (3) *vernier-anti-vernier* (V-AV), where the central line and the fourth flanking line were offset in opposite directions; and (4) *vernier-pro-vernier* (V-PV), where the central line and the fourth flanking line were offset in the same direction. A red fixation point (0.1° diameter) was presented throughout the trial. The intertrial interval was 1.5 seconds.

Offset sizes

Before the experiment proper, we calibrated the offsets for each participant to achieve comparable performances across observers and so that the central and the flanking lines offsets individually had the same contribution. We presented sequences with only the central or the flanking line offset. A parameter estimation by sequential testing (PEST) adaptive procedure (Taylor & Creelman, 1967) was used to determine the offset sizes necessary to yield around

75% to 80% dominance for the central offset and 20% to 25% dominance for the flanking line offset. The threshold of 75% to 80% allows detection of the effect of an additional offset. The offset sizes were determined with the fixation condition only and were kept the same in the other conditions.

Conditions

In all conditions, the fixation point was (initially) presented 1° above the center of the screen. There were four conditions (Figure 2a).

Fixation condition

The SQM started from the center of the screen and continued smoothly. Observers fixated the fixation point during the entire trial. The motion was smooth in both retinotopic and spatiotopic coordinates.

Spatiotopic condition

The fixation point disappeared 80 ms before stimulus onset and immediately reappeared 1° below the center of the screen. The SQM was presented in the center of the screen. Observers were instructed to make a saccade to the fixation point as soon as it was displaced. Thus, there was no displacement in spatiotopic coordinates compared to the fixation condition but there was a displacement in retinotopic coordinates.

Retinotopic condition

The fixation point disappeared 80 ms before stimulus onset and immediately reappeared 1° below the center of the screen. Observers were instructed to make a saccade to the fixation point as soon as it was displaced. The SQM was initially displayed in the center of the screen. When the saccade was detected, the rest of the sequence of the SQM was presented 1° below the fixation point. Thus, the retinal stimulation was the same as in the fixation condition, but there was a displacement in spatiotopic coordinates.

Jump condition

The SQM was presented in the center of the screen, and 210 ms after stimulus onset the rest of the sequence of the SQM was presented 1° above the fixation point. Observers fixated the fixation point during the entire trial. Thus, the SQM was displaced in both spatiotopic and retinotopic coordinates compared to the fixation condition.

Procedure

Observers were instructed to attend to one of the streams. Half of the participants attended to the left stream and the other half attended to the right stream. The task was to report the offset direction perceived in the attended stream by pressing one of two hand-held push buttons. Before each saccade condition, observers were trained to perform the saccade without offsets.

Experiment 1

The fixation, spatiotopic, retinotopic, and jump conditions were tested on 4 consecutive days, each condition on a separate day. For each condition, four blocks were completed (16 variants in total): (1) only the central line was offset (condition V); (2) only the flanking line in frame 4 in the attended stream was offset (condition AV); and (3 and 4) both lines were offset in either the same or in opposite directions in a randomly interleaved fashion—that is, half of the trials in a block were V-AV and the other half V-PV. Each block consisted of 80 trials. All participants started with the fixation condition and ended with the jump condition. The order of the spatiotopic and retinotopic conditions was randomized. The first testing session lasted around 1.5 hours. The following sessions lasted between 1 and 1.25 hours.

Experiment 2

The fixation, spatiotopic, and retinotopic conditions were tested on six new observers, each condition on 3 consecutive days. In the fixation condition, observers completed three blocks: (1) only the central line was offset (condition V), (2) only the flanking line in frame 4 in the attended stream was offset (condition AV), and (3) both lines were offset in opposite directions (condition V-AV). The fixation condition was used to calibrate the individual offsets and to verify that the two offsets integrated in condition V-AV. After the fixation condition, participants were informed about the paradigm and told that two offsets were presented in V-AV blocks. Four variants were performed in the spatiotopic and retinotopic conditions: (1) only the central line was offset (condition V); (2) only the flanking line in frame 4 in the attended stream was offset (condition AV); (3) both lines were offset in opposite directions, and observers were instructed to report the central offset (condition V-AV R[V]); and (4) both lines were offset in opposite directions, and observers were instructed to report the flanking line offset (condition V-AV R[AV]). Each block consisted of 80 trials. The order in which observers were tested with the spatiotopic and retinotopic conditions was randomized. The first testing session lasted around

1.5 hours. The following sessions lasted between 1 and 1.25 hours.

Fixation control and online saccade detection

For the trial to start, the gaze had to stay within a circular area of radius 0.5° for a duration of 1 second. The center of this area had to not deviate by more than 0.5° from the specified fixation position. After successful fixation, the gaze signal was baseline corrected to the average gaze position measured during start fixation. Within the stimulus presentation loop, the saccade detection algorithm was run each time and at the last possible moment before the stimulus animation had to be updated, thereby minimizing the delay between detecting a saccade and seeing the according stimulus change on the screen. A saccade was detected if all of the following were true:

1. No blinks were detected. A blink was defined as a signal loss for 40 ms or longer.
2. The average gaze position for the 120-ms time window preceding the saccade did not deviate by more than 0.8 arcdeg from the average gaze position during start fixation.
3. The first saccade sample deviated more than 1 arcdeg from the average gaze position for the preceding 120-ms time window.
4. The direction of the saccade, as indicated by the first saccade sample, was within $\pm 30^\circ$ of the expected direction, given the displacement of the fixation dot.

Trial selection

For analysis, we kept only trials that fulfilled four requirements: (a) there were no blinks or signal loss during the trial; (b) fixation was kept on the fixation point ($\pm 0.5^\circ$); (c) after a saccade, the eye correctly landed on the new fixation location and fixation was kept on the fixation point ($\pm 0.5^\circ$); and (d) the saccade was performed during two ISIs: either between 90 and 140 ms stimulus onset asynchrony (SOA) or between 160 and 210 ms SOA (in the saccade conditions). The third requirement implies that we kept only trials in which: (1) the saccade was performed between the presentation of the central offset and the presentation of the flanking line offset; (2) no line was presented on the screen during the saccade; and (3) after the saccade, at least one straight line was presented before the flanking line offset. As an example, the saccade dataset of one observer is displayed in Supplementary Figure S1.

A total of 5099 trials (49.8%) were kept in Experiment 1: in the *fixation condition*, V (398 trials, 62.2%), AV (390 trials, 60.9%), V-AV (383 trials, 59.8%), and V-PV

(358 trials, 55.9%); in the *spatiotopic condition*, V (283 trials, 44.2%), AV (253 trials, 39.5%), V-AV (211 trials, 33%), and V-PV (231 trials, 36.1%); in the *retinotopic condition*, V (292 trials, 45.6%), AV (241 trials, 37.7%), V-AV (268 trials, 41.9%), and V-PV (239 trials, 37.3%); and in the *jump condition*, V (377 trials, 58.9%), AV (389 trials, 60.8%), V-AV (385 trials, 60.2%), and V-PV (401 trials, 62.7%). In the fixation and jump conditions, 18.4% and 81.6% of the excluded trials were due to criteria (a) and (b), respectively. In the spatiotopic and retinotopic conditions, 8.1%, 4.5%, 30.8%, and 56.6% of the excluded trials were due to criteria (a), (b), (c), and (d), respectively.

In Experiment 2, 1962 trials (37.2%) were kept for analysis: in the *fixation condition*, V (245 trials, 51%), AV (196 trials, 40.8%), and V-AV (271 trials, 56.5%); in the *spatiotopic condition*, V (126 trials, 26.3%), AV (166 trials, 34.6%), V-AV R[V] (132 trials, 27.5%), and V-AV R[AV] (135 trials, 28.1%); and in the *retinotopic condition*, V (171 trials, 36.6%), AV (154 trials, 32.1%), V-AV R[V] (185 trials, 38.5%), and V-AV R[AV] (181 trials, 37.7%). In the fixation condition, 7.4% and 92.6% of the excluded trials were due to criteria (a) and (b), respectively. In the spatiotopic and retinotopic conditions, 4.2%, 1.3%, 26%, and 68.5% of the excluded trials were due to criteria (a), (b), (c), and (d), respectively. Analyses were also performed on the raw data and did not change the conclusions (Supplementary Figures S2 and S3).

Data analysis

Experiment 1

To assess integration, we compared condition V-AV to condition V and to the 50% dominance level using Holm-corrected, two-sided, one-sample *t*-tests. If there is integration, we would expect that dominance in the V-AV condition is non-significantly different from 50% and significantly different from condition V. Indeed, the dominance level in the V-AV condition should be different compared to when only one offset is present (condition V) if the two offsets integrate. Inversely, if there is no integration, we would expect that performance in V-AV is significantly different from 50% and non-significantly different from condition V. We corrected for eight comparisons (two tests each for the fixation, spatiotopic, retinotopic, and jump conditions).

Experiment 2

In Experiment 2, we tested whether integration is mandatory in the spatiotopic and retinotopic conditions. If the integration of the two offsets is mandatory, observers should not be able to report the individual offsets. In contrast, if the integration is not

mandatory, then observers should be able to report the individual offsets. Thus, dominance in V-AV R[V] and V-AV R[AV] should be at the same level as in V and AV, respectively. We therefore compared V-AV R[V] with V and V-AV R[AV] with AV using two-sided paired *t*-tests, Holm corrected for four comparisons.

Results

We tested how feature integration in the sequential metacontrast occurs across saccades. The central line was followed by 12 pairs of flanking lines. Observers attended to one of the streams. The central line and/or the fourth flanking line in the attended stream (280 ms after the presentation of the central line) were offset; that is, the lower segment was offset either to the right or to the left with respect to the upper segment (vernier offset). Four configurations were presented (Figure 1): (1) vernier (V), where only the central line was offset; (2) anti-vernier (AV), where only the fourth flanking line was offset; (3) vernier–anti-vernier (V-AV), where the central line and the fourth flanking line were offset in opposite directions; and (4) vernier–pro-vernier (V-PV), where the central line and the fourth flanking line were offset in the same direction. A red fixation point was presented throughout the trial.

We tested four conditions (Figure 2): fixation, spatiotopic, retinotopic, and jump. Performance was quantified in terms of dominance i.e., the percentage of observers' responses in accordance with the *central* vernier offset (Figure 3a). A dominance level above 50% means that the central vernier dominated performance, a dominance level below 50% indicates that the flanking line vernier dominated performance, and a dominance level around 50% indicates that none of the offsets was dominant.

Before each experiment, we calibrated each offset individually and for each participant to yield around 75% to 80% dominance (see Methods). Because performance was quantified in accordance with the *central* vernier offset, when 80% of the responses were in accordance with the *flanking* vernier offset direction dominance was 20%. The offsets were calibrated to yield 75% to 80% dominance to (1) provide the same contribution individually, (2) achieve comparable performance across observers, and (3) allow detection of the effect of an additional offset. Indeed, when both offsets were presented in opposite directions with the respective offset sizes, they canceled each other, and dominance was around 50% (Otto et al., 2006) (Figure 3a). When the offsets were in the same direction, dominance was above the 75% to 80% dominance threshold. The offset sizes were determined in the fixation condition and were kept the same throughout the other conditions.

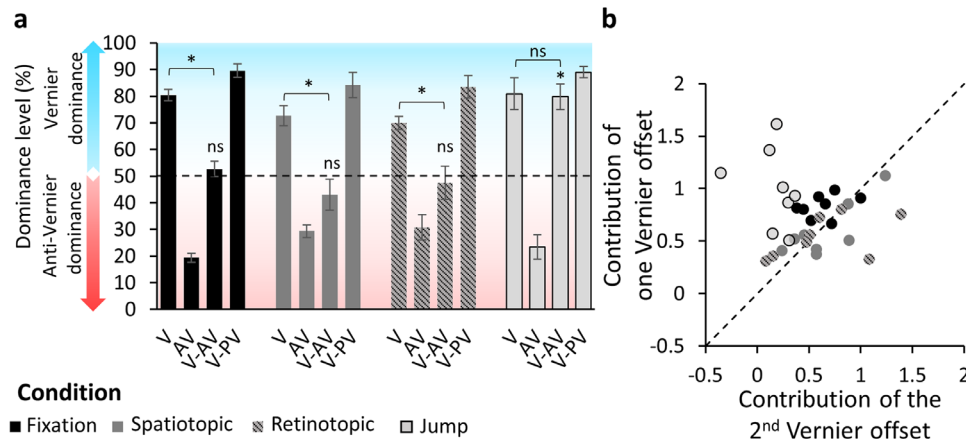


Figure 3. Experiment 1. (a) Dominance level as a function of the different configurations of the SQM for the fixation (black), spatiotopic (dark gray), retinotopic (dark gray/gray dashed), and jump (light gray) conditions. A dominance level above 50% (blue part of the plot) indicates that the central vernier dominated performance; a dominance level below 50% (red part of the plot) indicates that the anti-vernier dominated performance. Offset sizes were calibrated in the fixation condition to yield 75% to 80% dominance when only one line was offset (V or AV). The dominance level in the AV condition was 19.4%; that is, it was 80.6% correct in accordance with the flanking line offset. In the fixation condition V-AV, the dominance level was around 50%; that is, the offsets canceled each other. When the offsets were in the same direction (V-VP), the dominance level was around 90%; hence, the offsets integrated. The same pattern was observed in the spatiotopic and retinotopic conditions. In the jump condition, dominance in the V-AV configuration was around 80%, indicating that the offsets did not integrate. Error bars represent SEM. ns, $p_{Holm} > 0.05$; * $p_{Holm} \leq 0.05$. Individual data are presented in the Supplementary Material (Supplementary Figure S4). (b) Contribution of one vernier offset (average of percent correct in V and AV) as a function of the contribution of the other vernier offset $[(V-PV) - (V-AV)]/2$, when both offsets were present for each observer and each condition in z-score space. The diagonal represents perfect integration (both offsets contributed equally to the performance). The fixation (black), spatiotopic (dark gray), and retinotopic (gray dashed) conditions are close to the diagonal, showing integration of the offsets. The jump condition (light gray) forms a cluster above the diagonal, indicating that the central offset dominated performance.

Features integrate across saccades

In the fixation condition, observers looked at the fixation point located 1° above the center of the screen during the entire trial, while the SQM was displayed in the center of the screen. This condition is similar to the standard SQM. The dominance level in the V-AV condition was around 50%, indicating that the offsets integrated in this condition: for V-AV versus V, $t(7) = 9.6$, $p_{Holm} = 0.00023$, Cohen's $d = 3.4$, and 95% confidence interval (CI), -5.23 to -1.51 ; for V-AV versus 50%, $t(7) = 0.95$, $p_{Holm} = 1.0$, Cohen's $d = 0.34$, and 95% CI, -0.39 to 1.04 (Figure 3a).

In the spatiotopic and retinotopic conditions, observers performed a 2° downward saccade between the central offset and the flanking line offset. In the spatiotopic condition, the SQM was displayed in the center of the screen during the entire trial, whereas, in the retinotopic condition, the streams were displayed 1° below the fixation point after the saccade, making this condition retinotopically similar to the fixation condition. We found that the two offsets integrated in both the spatiotopic condition—V-AV versus V:

$t(7) = 4.12$, $p_{Holm} = 0.02$, Cohen's $d = 1.46$, 95% CI, -2.46 to -0.42 ; V-AV versus 50%: $t(7) = 1.2$, $p_{Holm} = 1$, Cohen's $d = 0.42$, and 95% CI, -1.14 to 0.32 —and the retinotopic condition—V-AV versus V: $t(7) = 8.04$, $p_{Holm} = 0.00062$, Cohen's $d = 2.84$, and 95% CI, -4.44 to -1.22 ; V-AV versus 50%: $t(7) = 0.63$, $p_{Holm} = 1$, Cohen's $d = 0.22$, and 95% CI, -0.92 to 0.49 .

In the jump condition, observers fixated on the fixation point during the entire trial. The streams were initially displayed in the center of the screen and then reappeared 1° above the center. The jump occurred between the presentation of the central line offset and the flanking line offset. We found no integration in the jump condition: V-AV versus V: $t(7) = 0.21$, $p_{Holm} = 1$, Cohen's $d = 0.074$, and 95% CI, -0.77 to 0.622 ; V-AV versus 50%: $t(7) = 6.3$, $p_{Holm} = 0.0024$, Cohen's $d = 2.23$, and 95% CI, 0.88 to 3.54 . Dominance was driven almost exclusively by the central vernier offset.

These results suggest that feature integration in the SQM occurs in both spatiotopic and retinotopic coordinates. To visualize all of the conditions in one graph, we plotted the mutual contributions of both verniers against each other in z-score space (Figure 3b).

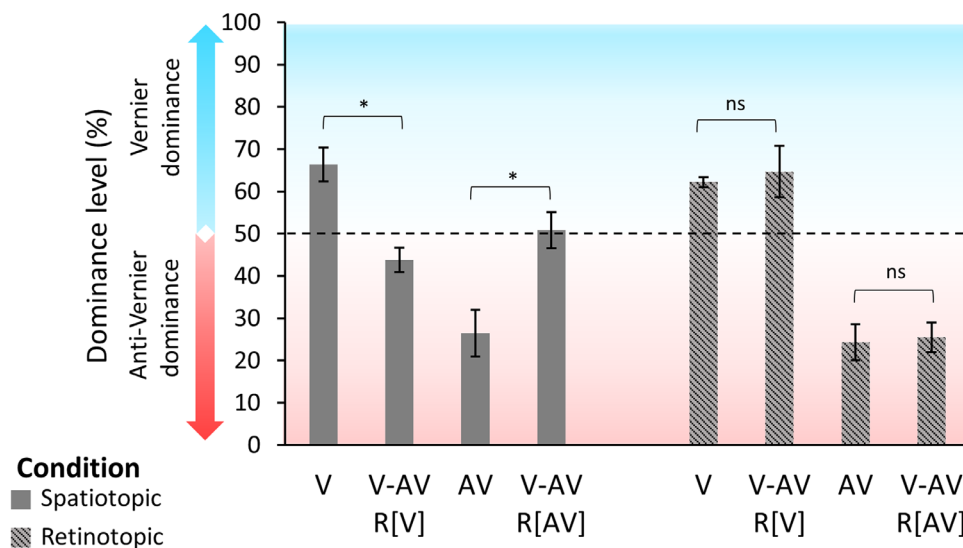


Figure 4. Experiment 2. Dominance level as a function of the different configurations of the SQM in the spatiotopic (dark gray) and retinotopic (gray dashed) conditions. In V-AV R[V], participants were instructed to report the central vernier offset direction; in V-AV R[AV], participants were instructed to report the anti-vernier offset direction. Performances were around 70% to 75% when only one line was offset (V and AV). Because the data are plotted in accordance with the central line offset, the dominance level in the AV condition is around 25%. In the spatiotopic condition, dominances in both the V-AV R[V] and V-AV R[AV] conditions were around 50%, indicating that observers were not able to report the individual offsets. In the retinotopic condition, observers reported the individual offsets in V-AV at a similar level as when only one offset was presented. ns, $p_{Holm} > 0.05$; * $p_{Holm} \leq 0.05$. Individual data are presented in the Supplementary Material (Supplementary Figure S5).

Integration is mandatory when object identity is preserved

In the jump condition of Experiment 1, we found, as expected, no integration; however, we found integration in both the spatiotopic and retinotopic conditions. We performed Experiment 2 to test whether integration is mandatory in these conditions (i.e., whether observers are able to report the individual offsets in the condition V-AV). For these reasons, we did not include the V-PV configuration or the jump condition.

First, six new observers were presented with the fixation condition to verify the integration of the two offsets (V: mean = 79.7, SEM = 2.12; AV: mean = 23.9, SEM = 1.92; V-AV: mean = 52.4, SEM = 3.29). Then, observers were informed about the paradigm and told that two offsets were present in condition V-AV. Participants were instructed to report, blockwise, either the central vernier offset (R[V]) or the flanking line offset (R[AV]) direction in the spatiotopic and retinotopic conditions. If the integration of the two offsets is mandatory, observers should not be able to report the individual elements, and the dominance level should be around 50%. In contrast, if the integration is not mandatory, then observers should be able to report the individual offsets. Thus, the dominance in V-AV R[V] and V-AV R[AV] should be at the same level as in V and AV, respectively.

Integration is mandatory in the spatiotopic condition; for two-sided, Holm-corrected, paired t -tests, V versus (V-AV R[V]): $t(5) = 5.73$, $p_{Holm} = 0.008$, Cohen's $d = 2.34$, and 95% CI, 0.71 to 3.94; AV versus (V-AV R[AV]): $t(5) = 3.7$, $p_{Holm} = 0.042$, Cohen's $d = 1.51$, and 95% CI, -2.7 to -0.27 (Figure 4). In the retinotopic condition, observers were able to report the individual offsets as dominances in the V-AV conditions were similar to the V and AV conditions; for two-sided, Holm-corrected, paired t -tests, V versus (V-AV R[V]): $t(5) = 0.37$, $p_{Holm} = 1.0$, Cohen's $d = 0.15$, and 95% CI, -0.95 to 0.66 ; AV versus (V-AV R[AV]): $t(5) = 0.36$, $p_{Holm} = 1.0$, Cohen's $d = 0.15$, and 95% CI, -0.95 to 0.67 (Figure 4).

Discussion

Here, we tested feature integration in the SQM across saccades. The trajectory of the SQM could be smooth on the screen or make an abrupt jump, and the conditions could or could not contain a saccade. Our results show that features integrate across saccades when object identity is preserved in external (i.e., spatiotopic) coordinates.

In the fixation condition, there was a smooth apparent-motion trajectory both on the screen and

on the retina. In line with previous results (Otto et al., 2006), offsets integrated. In the jump condition, there was no smooth trajectory on the screen nor on the retina. The pre-saccadic part of the streams was below the fixation dot, the post-saccadic part above fixation. There was no integration. Observers could easily report the central vernier offset. Interestingly, in the spatiotopic condition, motion was smooth on the screen (as in the fixation condition) but disrupted on the retina as in the jump condition. Despite this disruptive translation, not only did integration occur but it was also *mandatory*. In the retinotopic condition, the trajectory was not smooth on the screen (as in the jump condition) but was smooth on the retina (as in the fixation condition). Integration was optional. Observers were able to report each offset independently when instructed; however, the offsets integrated when individual reporting of offsets was not instructed.

Trans-saccadic integration is usually explained by remapping of receptive fields (Duhamel, Colby, & Goldberg, 1992) or of attentional pointers (Cavanagh, Hunt, Afraz, & Rolfs, 2010), where information from the pre-saccadic location is mapped onto the corresponding post-saccadic location. The basic idea is that, based on an efference copy, neurons respond to the post-saccadic location already before the saccade is executed and maintain the selectivity to that location until after the saccade, effectively bridging two consecutive fixations. In our experiments, not only the eyes but also the stimulus were in movement. Hence, integration did not depend solely on a correction of a static map operated by an efference copy, but stimulus motion needs to be taken into account. When there are multiple stimuli, these stimuli need to be grouped in both space and time. Accordingly, simple integration by corresponding locations across saccades cannot explain our results because whether or not the offsets integrate depends on the spatiotemporal configuration of the SQM.

Our findings are consistent with object-based remapping theories. For example, Deubel et al. (1996) found that introducing a post-saccadic blank cancels the insensitivity to trans-saccadic displacement (Bridgeman & Mayer, 1983). This finding shows that the pre-saccadic location is available after the saccade. It has been suggested that the blank breaks the object correspondence. Two objects are then represented and compared; thus, the displacement becomes noticeable (Deubel et al., 1996; Tas et al., 2012). Tas et al. (2012) found similar results when changing the contrast or the identity of the post-saccadic stimulus. Although these results show that the pre-saccadic stimulus information (location and features) is retained, Tas et al. (2012) interpreted these results within an object updating theory (Enns, Lleras, & Moore, 2010). In this theory, when object correspondence is established, the pre-saccadic object is overwritten and replaced by

the post-saccadic object. We, in contrast, propose that features are not overwritten but are integrated across saccades when object correspondence is established. Importantly, the integration is mandatory.

The mechanisms responsible for maintaining object identity across saccades might be a more general mechanism that maintains object identity across any visual disruption. Indeed, it has been shown that masking in the absence of eye movements produces qualitatively similar effects as saccades in terms of spatial and temporal compression as well as suppression of displacement (Zimmermann, Born, Fink, & Cavanagh, 2014).

Object identity also determines feature integration in paradigms without eye movements such as the classical SQM and the Ternus–Pikler display (Boi, Ögmen, Krummenacher, Otto, & Herzog, 2009; Ögmen, Otto, & Herzog, 2006). We proposed a two-stage model in which the first stage consists of Gestalt figure–ground segregation and grouping (Ögmen & Herzog, 2010). These processes establish a reference frame for each group. These reference frames determine object identities and are used to attribute features to stimuli and integrate them according to group (object) identities. Hence, mandatory integration across the motion trajectory in the SQM without eye movements may reflect a default operation of the human brain to obtain good estimates of features of moving objects. However, it is important to integrate only features of the same object and not mingle features of different objects. We propose that, for this reason, integration occurs only within a motion stream (Otto et al., 2006). Thus, features integrate non-retinotopically according to object identity with or without eye movements.

We found optional integration in the retinotopic condition. Analysis of reference frames for the perception of motion indicates that the visual system uses an amalgamation of different reference frames, both retinotopic and non-retinotopic, instead of a single reference frame that emerges as a “winner-take-all” (Agaoglu, Herzog, & Ögmen, 2015). We previously found that, although the majority of integration takes place according to a non-retinotopic motion-based reference frame, there is still some residual retinotopic integration (Ağaoğlu, Herzog, & Ögmen, 2012; Lauffs, Choung, Ögmen, & Herzog, 2018). Fabius et al. (2016), for example, also found that motion information can be stored both in a retinotopic and in a spatiotopic reference frame. Moreover, reference frames depend on the identity and strength of perceptual groups that generate them. For example, in the retinotopic condition, observers may group the fixation point with the SQM streams. Hence, the pre- and post-saccadic stimuli are grouped together via the fixation point, explaining why integration takes place in the retinotopic

condition. However, because the pre- and post-saccadic parts of the stimuli are not smoothly connected in space–time, they can be viewed as two event segments within the spatiotemporal group. In this case, integration is possible (because both offsets belong to the common overarching group) but optional (due to the fact that offsets belong to two different segments of the same overall group).

Traditionally, the effect of remapping has been documented with large saccades (typically 10°) (e.g., Fracasso et al., 2010; Melcher & Morrone, 2003) and simple, well-isolated stimuli which do not tax much of the spatial resolution of the remapping system. Previous estimates of the trans-saccadic receptive field have suggested that the remapping system may be quite coarse (Cicchini, Binda, Burr, & Morrone, 2012). In our paradigm, to keep stimuli parafoveal, we employed rather small saccades for which still integration occurred; thus, our findings indicate that the system for trans-saccadic integration has a finer spatial selectivity than previously thought.

To conclude, our results show that features integrate mandatorily even if they are not in proximity on the retina because of a saccade. This shows that the brain actively combines the pre- and post-saccadic information to obtain object identity.

Keywords: Trans-saccadic perception, reference frames, visual stability, feature integration

Acknowledgments

The authors thank Marc Repnow for technical support and discussions about the data. This work was supported by a grant from the Swiss SystemsX.ch initiative (2015/336); by a Swiss National Science Foundation grant (“Basics of visual processing: from elements to figures,” 176153), by grants from the European Union (EU) and Horizon 2020–ERC Advanced (“Spatio-temporal mechanisms of generative perception,” 832813–GenPercept); by a grant from the Italian Ministry of Education, University, and Research under the PRIN2017 programme (“Temporal context in perception: serial dependence and rhythmic oscillations,” 2017SBCPZY); and by a grant from the European Programmes Flag-ERA JTC 2019 (DOMINO).

Commercial relationships: none.

Corresponding author: Leila Drissi-Daoudi.

Email: leila.drissidaoudi@epfl.ch.

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

References

- Ağaoğlu, M. N., Herzog, M. H., & Ögmen, H. (2012). Non-retinotopic feature processing in the absence of retinotopic spatial layout and the construction of perceptual space from motion. *Vision Research*, 71, 10–17.
- Agaoglu, M. N., Herzog, M. H., & Ögmen, H. (2015). The effective reference frame in perceptual judgments of motion direction. *Vision Research*, 107, 101–112.
- Bach, M. (1996). The Freiburg Visual Acuity test—Automatic measurement of visual acuity. *Optometry and Vision Science*, 73(1), 49–53.
- Boi, M., Ögmen, 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):, 1–11, <https://doi.org/10.1167/9.13.5>.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Bridgeman, B. (2011). Visual stability. In S. P. Livensedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 511–521). Oxford, UK: Oxford University Press.
- Bridgeman, B., & Mayer, M. (1983). Failure to integrate visual information from successive fixations. *Bulletin of the Psychonomic Society*, 21(4), 285–286.
- 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.
- Cicchini, G. M., Binda, P., Burr, D. C., & Morrone, M. C. (2012). Transient spatiotopic integration across saccadic eye movements mediates visual stability. *Journal of Neurophysiology*, 109(4), 1117–1125.
- Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2010). Parametric integration of visual form across saccades. *Vision Research*, 50(13), 1225–1234.
- Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research*, 36(7), 985–996.
- Drissi-Daoudi, L., Doerig, A., & Herzog, M. H. (2019). Feature integration within discrete time windows. *Nature Communications*, 10(1), 1–8.
- Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040), 90–92.

- 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., Lleras, A., & Moore, C. M. (2010). Object updating: A force for perceptual continuity and scene stability in human vision. In R. Nihawan, & B. Khurana (Eds.), *Space and Time in Perception and Action* (pp. 50–520). Cambridge, UK: Cambridge University Press.
- Fabius, J. H., Fracasso, A., & Van der Stigchel, S. (2016). Spatiotopic updating facilitates perception immediately after saccades. *Scientific Reports*, 6(1), 1–11.
- Fornaciai, M., Binda, P., & Cicchini, G. M. (2018). Trans-saccadic integration of orientation information. *Journal of Vision*, 18(4):, 1–11, <https://doi.org/10.1167/18.4.9>.
- Fracasso, A., Caramazza, A., & Melcher, D. (2010). Continuous perception of motion and shape across saccadic eye movements. *Journal of Vision*, 10(13):, 1–17, <https://doi.org/10.1167/10.13.14>.
- Ganmor, E., Landy, M. S., & Simoncelli, E. P. (2015). Near-optimal integration of orientation information across saccades. *Journal of Vision*, 15(16):, 1–12, <https://doi.org/10.1167/15.12.1306>.
- Harrison, W. J., & Bex, P. J. (2014). Integrating retinotopic features in spatiotopic coordinates. *Journal of Neuroscience*, 34(21), 7351–7360.
- Hübner, C., & Schütz, A. C. (2017). Numerosity estimation benefits from transsaccadic information integration. *Journal of Vision*, 17(13):, 1–16, <https://doi.org/10.1167/17.13.12>.
- Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, 23(3), 420–456.
- Jonides, J., Irwin, D. E., & Yantis, S. (1983). Failure to integrate information from successive fixations. *Science*, 222(4620):188.
- 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.
- Melcher, D. (2005). Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Current Biology*, 15(19), 1745–1748.
- Melcher, D. (2007). Predictive remapping of visual features precedes saccadic eye movements. *Nature Neuroscience*, 10(7), 903–907.
- Melcher, D. (2008). Dynamic, object-based remapping of visual features in trans-saccadic perception. *Journal of Vision*, 8(14):, 1–17, <https://doi.org/10.1167/8.14.2>.
- Melcher, D. (2011). Visual stability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1564), 468–475.
- Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature Neuroscience*, 6(8), 877–881.
- Melcher, D., & Morrone, M. C. (2015). Nonretinotopic visual processing in the brain. *Visual Neuroscience*, 32, E017.
- 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.
- Ögmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, 46(19), 3234–3242.
- 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>.
- 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), 1670–1686.
- Paeye, C., Collins, T., & Cavanagh, P. (2017). Transsaccadic perceptual fusion. *Journal of Vision*, 17(1):, 1–11, <https://doi.org/10.1167/17.1.14>.
- Poth, C. H., Herwig, A., & Schneider, W. X. (2015). Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Frontiers in Systems Neuroscience*, 9, 176.
- Rayner, K., & Pollatsek, A. (1983). Is visual information integrated across saccades? *Perception & Psychophysics*, 34(1), 39–48.
- Stewart, E. E., & Schütz, A. C. (2018). Optimal trans-saccadic integration relies on visual working memory. *Vision Research*, 153, 70–81.
- Szinte, M., & Cavanagh, P. (2011). Spatiotopic apparent motion reveals local variations in space constancy. *Journal of Vision*, 11(2):, 1–20, <https://doi.org/10.1167/11.2.4>.
- Tas, A. C., Moore, C. M., & Hollingworth, A. (2012). An object-mediated updating account of insensitivity to transsaccadic change. *Journal of Vision*, 12(11):, 1–13, <https://doi.org/10.1167/12.11.18>.
- Tas, A. C., Moore, C. M., & Hollingworth, A. (2014). The representation of the saccade target object

- depends on visual stability. *Visual Cognition*, 22(8), 1042–1046.
- Taylor, M. M., & Creelman, C. D. (1967). PEST: efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, 41(4A), 782–787.
- von Helmholtz, H. (1896). *Theorie der Luftschwingungen in Röhren mit offenen Enden*. Leipzig, Germany: W. Engelmann.
- Wijdenes, L. O., Marshall, L., & Bays, P. M. (2015). Evidence for optimal integration of visual feature representations across saccades. *Journal of Neuroscience*, 35(28), 10146–10153.
- Wittenberg, M., Bremmer, F., & Wachtler, T. (2008). Perceptual evidence for saccadic updating of color stimuli. *Journal of Vision*, 8(14):, 1–9, <https://doi.org/10.1167/8.14.9>.
- Wolf, C., & Schütz, A. C. (2015). Trans-saccadic integration of peripheral and foveal feature information is close to optimal. *Journal of Vision*, 15(16):, 1–18, <https://doi.org/10.1167/15.16.1>.
- Wolfe, B. A., & Whitney, D. (2015). Saccadic remapping of object-selective information. *Attention, Perception, & Psychophysics*, 77(7), 2260–2269.
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, 48(20), 2070–2089.
- Zimmermann, E., Born, S., Fink, G. R., & Cavanagh, P. (2014). Masking produces compression of space and time in the absence of eye movements. *Journal of Neurophysiology*, 112(12), 3066–3076.
- Zimmermann, E., Burr, D., & Morrone, M. C. (2011). Spatiotopic visual maps revealed by saccadic adaptation in humans. *Current Biology*, 21(16), 1380–1384.
- Zimmermann, E., Morrone, M. C., & Binda, P. (2018). Perception during double-step saccades. *Scientific Reports*, 8(1), 1–9.