

Features integrate along a motion trajectory when object integrity is preserved

Leila Drissi-Daoudi

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



Haluk Öğmen

Department of Electrical & 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



Information about a moving object is usually poor at each retinotopic location because photoreceptor activation is short, noisy, and affected by shadows, reflections of other objects, and so on. Integration across the motion trajectory may yield a much better estimate about the objects' features. Using the sequential metacontrast paradigm, we have shown previously that features, indeed, integrate along a motion trajectory in a long-lasting window of unconscious processing. In the sequential metacontrast paradigm, a percept of two diverging streams is elicited by the presentation of a central line followed by a sequence of flanking pairs of lines. When several lines are spatially offset, the offsets integrate mandatorily for several hundreds of milliseconds along the motion trajectory of the streams. We propose that, within these long-lasting windows, stimuli are first grouped based on Gestalt principles of grouping. These processes establish reference frames that are used to attribute features. Features are then integrated following their respective reference frame. Here using occlusion and bouncing effects, we show that indeed such grouping operations are in place. We found that features integrate only when the spatiotemporal integrity of the object is preserved. Moreover, when several moving objects are present, only features belonging to the same object integrate. Overall, our results show that feature integration is a deliberate strategy of the brain and long-lasting windows of processing can be seen as periods of sense making.

Introduction

Sensory information needs to be integrated across space and time, to perceive motion, for example. A moving object remains at a retinotopic location briefly and, as a result, the computation of its features requires the integration of information along the motion path of the moving object. During its motion, the object is also often occluded by other stationary or moving objects. These observations raise two fundamental questions: 1) How does the visual system integrate feature information along motion pathways? and 2) How does the visual system avoid mixing up features of different objects when occlusions occur?

The “two-stage” model shown in [Figure 1](#) provides a possible answer to these questions. According to this model, stimuli are grouped at an “early stage” based on Gestalt principles of grouping. The grouping operations generate reference frames. For example, for a moving object, stimuli at different retinotopic locations will be grouped, and the set of these retinotopic locations together will constitute a spatiotemporal reference frame. In other words, for a moving object the reference frame corresponds to the path of motion. At the next stage, features are then attributed to objects according to these early grouping operations and they will be integrated following their respective reference frames.

To investigate spatiotemporal feature integration and test this model, we used the sequential metacontrast paradigm (SQM; [Otto, Öğmen, & Herzog, 2006](#); [Otto, Öğmen, & Herzog, 2009](#); [Piéron, 1935](#)). In the SQM, a central line is followed by pairs of flanking lines. Because of metacontrast masking ([Bachmann & Francis, 2013](#); [Breitmeyer & Öğmen, 2006](#)), the central

Citation: Drissi-Daoudi, L., Öğmen, H., & Herzog, M. H. (2021). Features integrate along a motion trajectory when object integrity is preserved. *Journal of Vision*, 21(12):4, 1–15, <https://doi.org/10.1167/jov.21.12.4>.



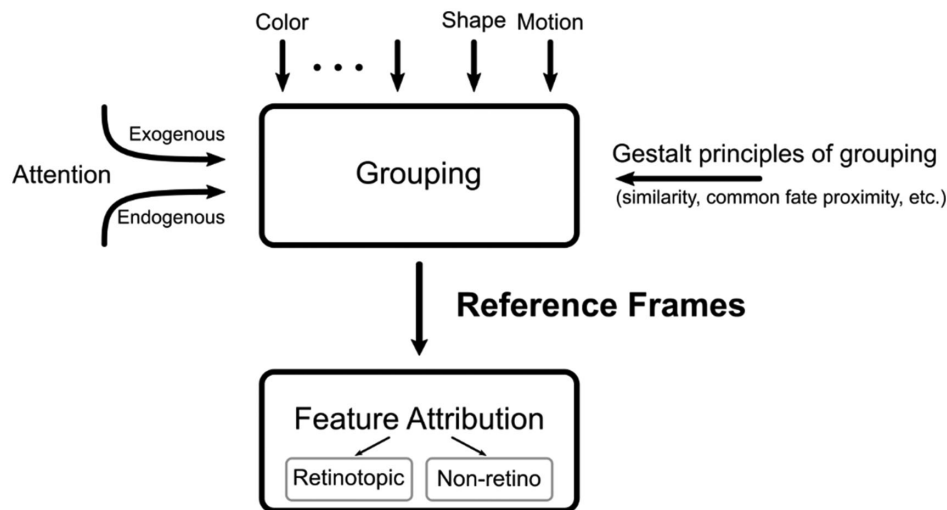


Figure 1. The “two-stage” model. The first stage consists of Gestalt grouping and segregation processes, which establish a reference frame for each group. These reference frames are then used to attribute features to stimuli. Features are then integrated following their respective reference frames.

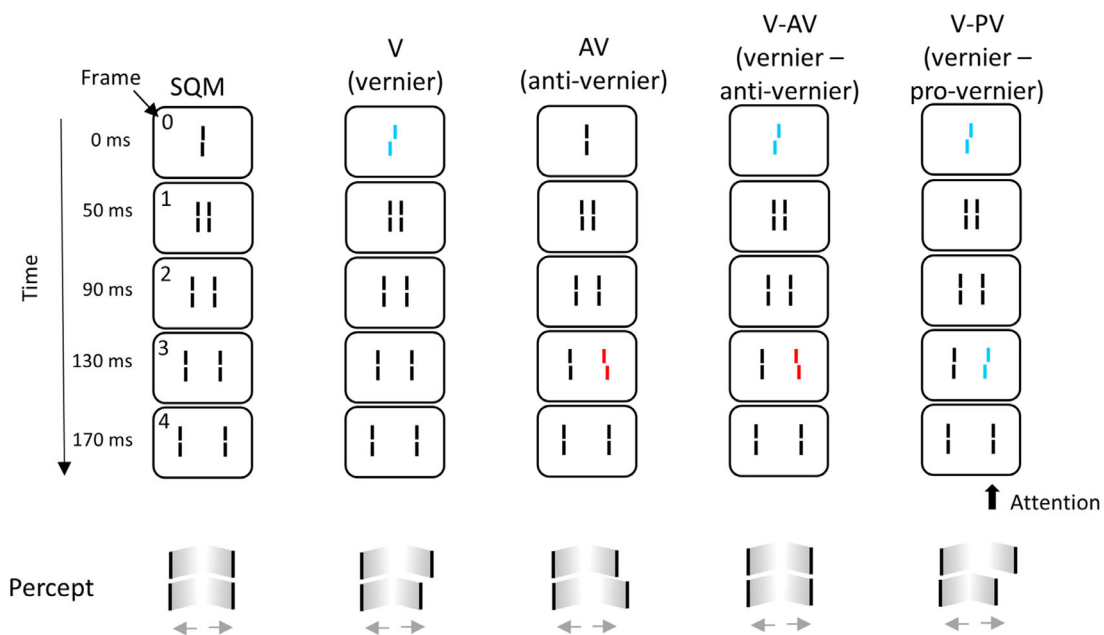


Figure 2. The sequential metacontrast (SQM). Each line was presented for 20 ms with an interstimulus interval (ISI) of 20 ms (30 ms for the first ISI to obtain strong masking of the central vernier). The percept is two streams of lines expanding from the center. The presentation of each pair of lines is a frame. Frame 0 corresponds to the presentation of the central line. *V (vernier)*: only the central line is offset, that is, the lower segment of the line is spatially offset to the right or to the left compared with the upper segment. *AV (anti-vernier)*: a flanking line is offset. *V-AV (vernier – anti-vernier)*: the central line and a flanking line are offset in opposite directions. *V-PV (vernier – pro-vernier)*: the central line and a flanking line are offset in the same direction. Observers are instructed to attend to one of the streams (here the right stream) and to report the direction (right or left) of the perceived offset. Colors are for illustration purpose. All stimulus elements were white or red on a black background. Figure adapted from (Drissi-Daoudi et al., 2019).

line is invisible. However, if the central line is spatially offset, that is, the lower segment is offset either to the right or to the left compared with the upper segment (Figure 2, vernier configuration), all the flanking lines

in the stream are perceived as offset, even though they are straight. Observers can report the direction of the perceived offset (right or left) in the stream they were instructed to attend to. When, in addition, a flanking

line is offset in the opposite direction as the central line, the two offsets integrate and cancel each other (Figure 2, vernier—anti-vernier configuration). When the offsets are in the same direction, discrimination improves (Figure 2, vernier—pro-vernier). The SQM show that features are integrated along the motion path.

Previously, we have also shown that offsets integrate in the SQM in discrete temporal windows lasting for several hundreds of milliseconds (Drissi-Daoudi, Doerig, & Herzog, 2019; Otto et al., 2009). Importantly, the integration is mandatory, that is, observers do not have access to the individual offsets. A window opens with stimulus onset, and once the first window closes, a second similar window opens (Drissi-Daoudi et al., 2019). These results support a two-stage model, in which, first, features are processed unconsciously and continuously with high spatiotemporal resolution in a long-lasting discrete window (Herzog, Drissi-Daoudi, & Doerig, 2020; Herzog, Kammer, & Scharnowski, 2016; see also Elliott & Giersch, 2016). After the window closes, we consciously perceive the output of the processing. We argue that these long-lasting windows can be seen as periods of sense making and, thus, feature integration is a deliberate strategy of the brain (Herzog et al., 2020).

According to this two-stage model, perceptual grouping should determine whether offsets integrate or not. When elements are perceived to belong to the same object, then it makes sense to integrate the information along the motion trajectory. However, as highlighted by the second question mentioned above, features of different objects should not be mingled. For example, vernier offsets in different streams of the SQM do not integrate but are perceived separately (Otto et al., 2006). Here, we used occlusions and the bouncing effect to manipulate perceptual grouping and thereby test the role of perceptual grouping in feature integration in the SQM.

Methods

Observers

Observers were students from the Ecole Polytechnique Fédérale de Lausanne. Participants signed informed consent, had normal or corrected-to-normal vision, and were paid for their participation. Visual acuity was 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 Declaration of Helsinki except for pre-registration.

Eight observers took part in Experiment 1a (age 20–28 years; 2 females) and 8 new observers participated

in Experiment 1b (age 19–28 years; 4 females). Eight new observers participated in Experiment 2 (age 20–33 years; 5 females) and 10 new observers participated in Experiment 3 (age 20–26 years; 5 females).

Apparatus

Stimuli were presented on a BenQ XL2540 24.5” LCD (1920 × 1080 pixels, used with 240 Hz; BenQ, Taipei, Taiwan; Experiment 1a) or on an ASUS VG248QE 24” LCD (1920 × 1080 pixels, used with 144 Hz; Asus Tek computer, Taipei, Taiwan; Experiments 1b, 2 and 3) monitor using Matlab (The MathWorks Inc.) with Psychtoolbox (Brainard, 1997; Pelli, 1997). Stimuli were white (luminance: 100 cd per m²) or red (20 cd/m²), on a black background with a luminance of 0.1 cd per m². Participants were seated 2.50 m from the screen in a dimly lit room. Viewing distance was kept constant by means of a chinrest.

Stimuli

The stimuli were variations of the sequential metacontrast stimulus (SQM; Figure 2; Otto, Ögmen, & Herzog, 2006). The sequence started with a central line consisting of two vertical segments of a length of 20° (arcmin), separated by a vertical gap of 2°. The line was followed by pairs of flanking lines presented one after the other further away from the center. The distance between the central line and the first flanking lines as well as between consecutive flanking lines was 3.3°. Each line was presented for 20 ms. The interstimulus interval (ISI) between the central line and the first pair of flanking lines was 30 ms (to obtain strong masking of the central vernier) and the ISI between consecutive pairs of flanking lines was 20 ms. A motion percept of two streams of lines diverging from the center is elicited. These presentation times values are nominal values that could only be presented approximately due to the refresh rates. For example, a presentation time of 20 ms was actually 20.8 ms (5 refresh cycles at 244 Hz and 3 refresh cycles at 144 Hz). ISIs and line duration were each rounded to a multiple of refresh cycles before computing the times of the line presentations. The width of the lines was 70” (3 pixels) and anti-aliasing was used to draw the lines.

One or more lines were spatially offset (vernier offset); that is, the lower segment of the line was offset either to the right or to the left with respect to the upper segment. Each trial was preceded by a fixation dot in the center of the screen for 1 second (Experiment 1a) or 500 ms (Experiments 1b, 2, and 3) followed by a blank screen for 500 ms. Then, the stimulus sequence was presented and participants responded by pressing

one of two hand-held push buttons. There were four configurations (Figure 2): 1) V (vernier): only the central line was offset; 2) AV (anti-vernier): only a flanking line was offset; 3) V–AV (vernier—anti-vernier): the central line and a flanking line were offset in opposite directions; and 4) V–PV (vernier—pro-vernier): the central line and a flanking line were offset in the same direction.

Experiment 1a (Figure 3a)

We tested whether features integrate across an occluder versus when there is a gap in the sequence of the SQM. The SQM was presented with eight pairs of lines (total stimulus duration: 350 ms). The anti-vernier was presented 290 ms after the central line (frame 7). We tested three conditions: Classic, Occluded and Gap (Figure 3a). In the Occluded condition, a grey rectangle ($42^\circ \times 9.9^\circ$, 40 cd/m^2) occluded the third, fourth, and fifth flanking lines of the attended stream. The occluder was presented during the entire trial, that is, from the fixation point presentation to the end of the SQM presentation. The center of the rectangle was 13.3° from the center of the screen. In the Gap condition, the third, fourth, and fifth flanking lines of the attended stream were not displayed.

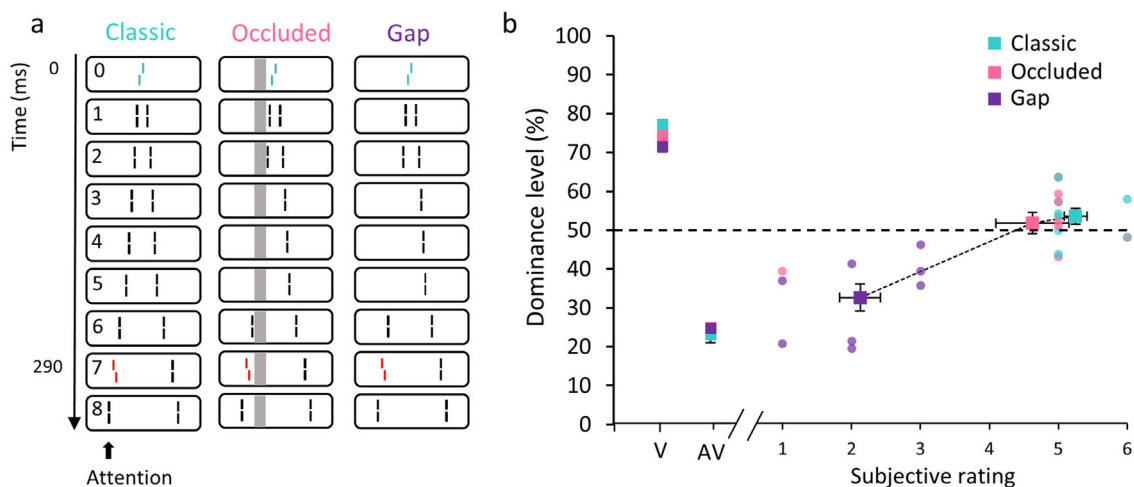


Figure 3. Experiment 1a. (a) The anti-vernier was presented in frame 7. In the Occluded condition, the lines of the attended stream in frames 3, 4, and 5 were occluded by a grey rectangle. The same three lines were missing in the Gap condition. Colors are for illustration purpose only. (b) V and AV show the offset calibrations with either the vernier (V) or the anti-vernier only (AV; the symbol of the occluded AV configuration is invisible because of the overlap with the other symbols; likewise, error bars are often too small to be visible). In the next conditions, both the vernier and the anti-vernier were presented together. We plot performance with respect to the subjective ratings. Observers rated the stream in the Classic condition (blue) as unified, and offsets integrated indicated by a dominance level of about 50%. Similarly, the stream was perceived as unified in the Occluded condition (pink), and offsets integrated. In the Gap condition (purple), the stream was perceived more disjointed than in the other conditions, and offsets integrated less. Observers reported mainly the offset of the anti-vernier, as they were instructed to report the perceived offset direction at the end of the motion trajectory. Thus, offsets integrate across the occluder. However, if the spatiotemporal integrity of the stream is not preserved, the offsets integrate less. Circles indicate individual data. Error bars represent standard errors of the mean.

Experiment 1b (Figure 4a)

Experiment 1b was identical to Experiment 1a except for the trajectory of the streams. First, the streams were diverging from the center, then, from frame 4 on, they switched direction to converge back to the center (Figure 4a).

Experiment 2 (Figure 5a)

We tested whether offsets integrate when the stream changes color in the middle of the trajectory, with and without an occluder. Additionally, we tested whether integration is mandatory across occlusion. The SQM was presented with nine pairs of flanking lines (total stimulus duration: 390 ms). The anti-vernier was presented 290 ms after the central line (frame 7). Four conditions were tested: Classic, Occluded, Classic_red, and Occluded_red (Figure 5a). In the Occluded and Occluded_red conditions a grey rectangle ($42^\circ \times 9.9^\circ$, 40 cd/m^2) occluded the third, fourth, and fifth flanking lines of the attended stream. The occluder was presented during the entire trial, that is, from the fixation point presentation to the end of the SQM presentation. The center of the rectangle was 13.3° from the center of the screen. In the Classic_red and Occluded_red conditions, the lines in the attended

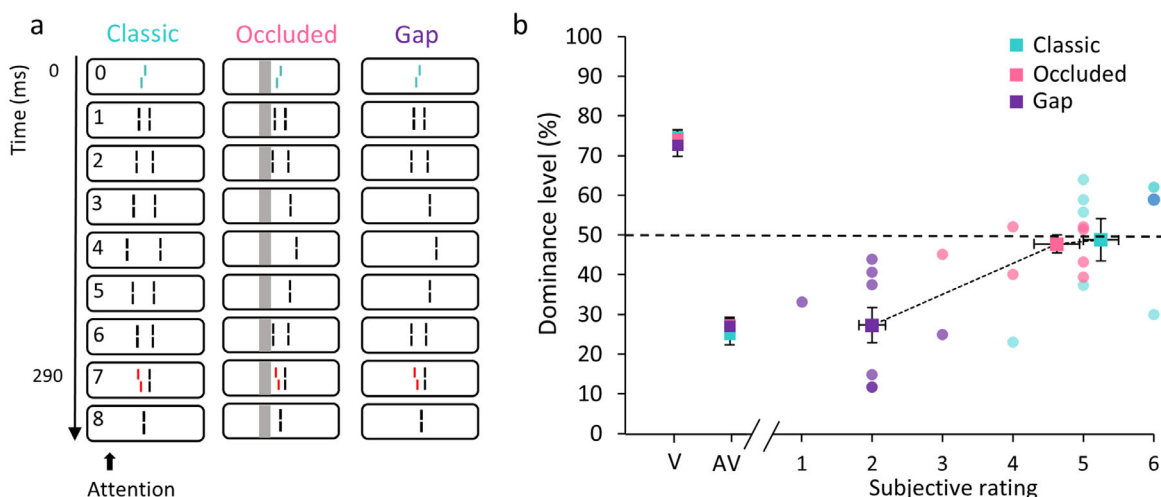


Figure 4. Experiment 1b. (a) Experiment 1b was identical to Experiment 1a except that the streams diverged until frame 4 and then converged back to the center. (b) Similar to Experiment 1a, observers perceived the stream as unified in the Classic (blue) and Occluded (pink) conditions. The offsets integrated in these conditions. In the Gap condition (purple), the stream appeared more disjointed than in the other conditions. The offsets integrated less. Circles indicate individual data. Error bars represent standard errors of the mean.

stream were first white, then were red from the fifth frame on.

Experiment 3 (Figure 6a)

We investigated the integration when two SQMs were presented. One SQM started 16.7° to the right of the center of the screen and the other 16.7° to the left of the center. The sequence contained 10 flanking lines after the central line (total stimulus duration: 430 ms). Only one stream of each sequence was displayed: the left stream of the right sequence and the right stream of the left sequence (Figure 6a). The lines of the right sequence were red and the lines of the left sequence were white. The two streams merged in frame 5 and continued their trajectory in two ways. In the Crossing condition, the red stream, starting on the right, continued its trajectory toward the left border of the screen after frame 5. Similarly, the white stream continued its trajectory toward the right border of the screen after frame 5. In the Bouncing condition, after the streams merged in frame 5, the streams were displayed from the center to their respective starting position. The lines of the red stream were longer than the lines of the white stream (segment length of 33.3°) to reinforce a percept of two streams that bounce against each other than a percept of two crossing streams.

Procedure

The different conditions were tested blockwise. The order of the conditions was randomized across observers to decrease the influence of learning and

fatigue effects in the averaged data. For each observer, each experimental condition was measured twice. Hence, each observer performed two blocks with each condition. After each condition had been measured once, the order of blocks was reversed for the second set of measurements. For example in Experiment 1a an observer performed the blocks in the following order: Occluded, Classic, Gap, Gap, Classic, and Occluded. The results of two identical conditions were collapsed. Each block contained 80 trials, yielding 160 trials per condition in total. The task was to report the perceived vernier offset direction (right or left) at the end of the motion trajectory by using hand-held push buttons.

Experiments 1a and 1b

Three conditions were tested: Classic, Occluded, and Gap (see the section on Stimuli). One-half of the observers were instructed to attend to the right stream and the other half to the left stream.

Experiment 2

In the first part of the experiment, four conditions were tested: Classic, Occluded, Classic_{red}, and Occluded_{red}, with the observers being naïve about the stimulus. One-half of the observers were instructed to attend to the right stream and the other half to the left stream. In the second part of the experiment, observers were informed of the paradigm, that is, that two offsets were presented. Observers were then instructed to report the direction of the anti-vernier offset (labelled R[AV]). The rest of the procedure was identical to the first part of the experiment.

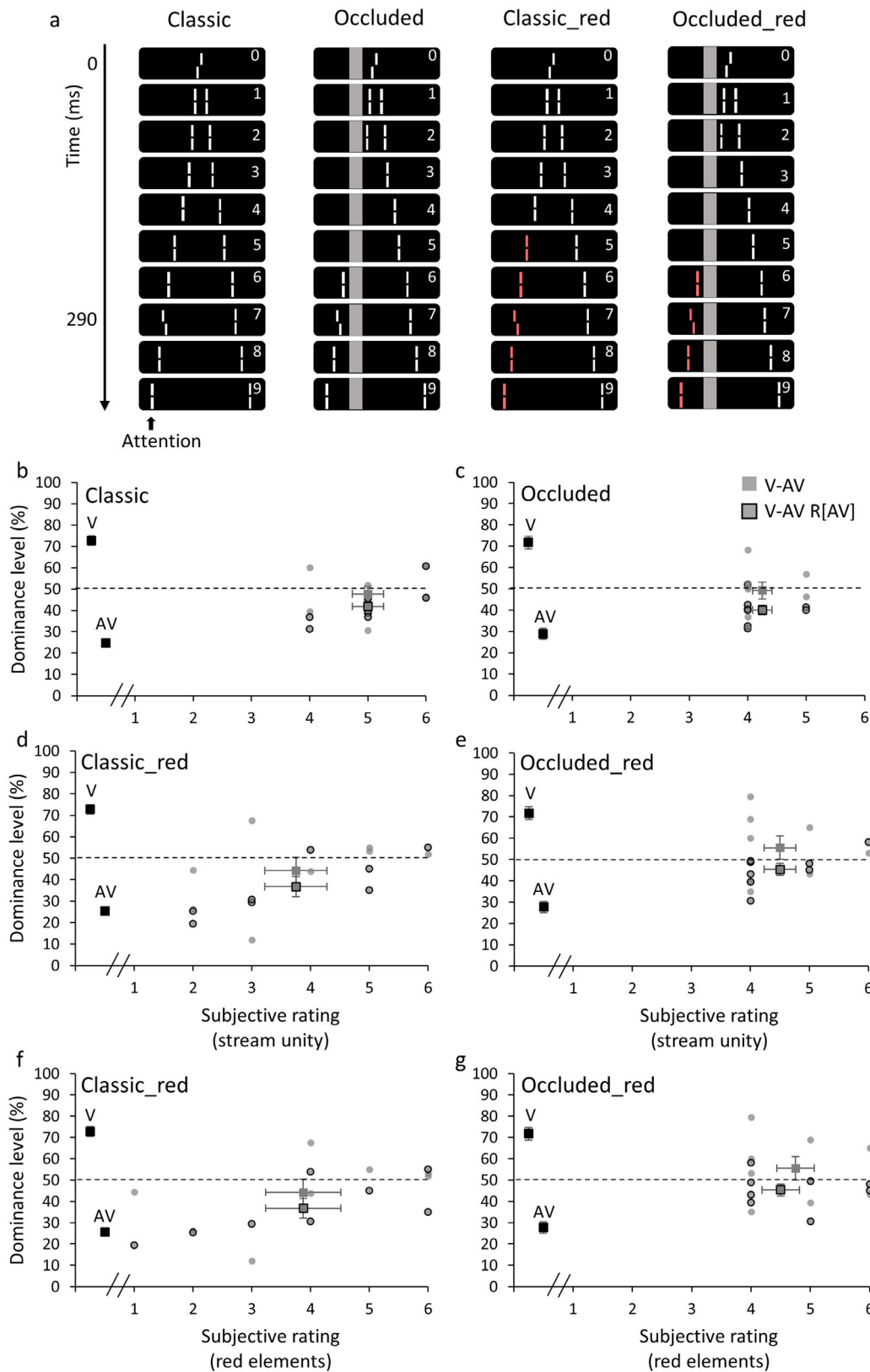


Figure 5. Experiment 2. (a) The anti-vernier was presented in frame 7 (290 ms). In the Occluded and Occluded_red conditions, the lines of the attended stream in frames 3, 4, and 5 were occluded by a grey rectangle. In the Classic_red and Occluded_red conditions, the lines of the attended stream from frame 5 to the end of the stimulus were red. (b, c, d and e) Dominance level as a function of the subjective ratings regarding stream unity (1 [“The motion stream appears completely disjointed”] to 6 [“The motion stream appears completely unified”]) in the Classic (b), Occluded (c), Classic_red (d) and Occluded_red (e) conditions. (f and g) Dominance level as a function of the subjective ratings regarding the integration of the red elements (1 [“The red elements appear to be completely



← separated from the motion stream”] to 6 [“The red elements appear to be completely part of them motion stream”]) in the Classic_red (f) and Occluded_red (g) conditions. V and AV show offset calibration. V–AV: observers were naïve. V–AV R[AV]: observers were informed of the paradigm and instructed to report the anti-vernier. Offsets largely integrated mandatorily across occlusion, even when there was a color change after the occluder. When there was no occluder and the stream chanded color mid-trajectory, integrations seemed to follow the subjective ratings. Offset integrated less when the stream was perceived as less unified and the red elements as less part of the stream. Circles indicate individual data. Error bars represent standard errors of the mean.

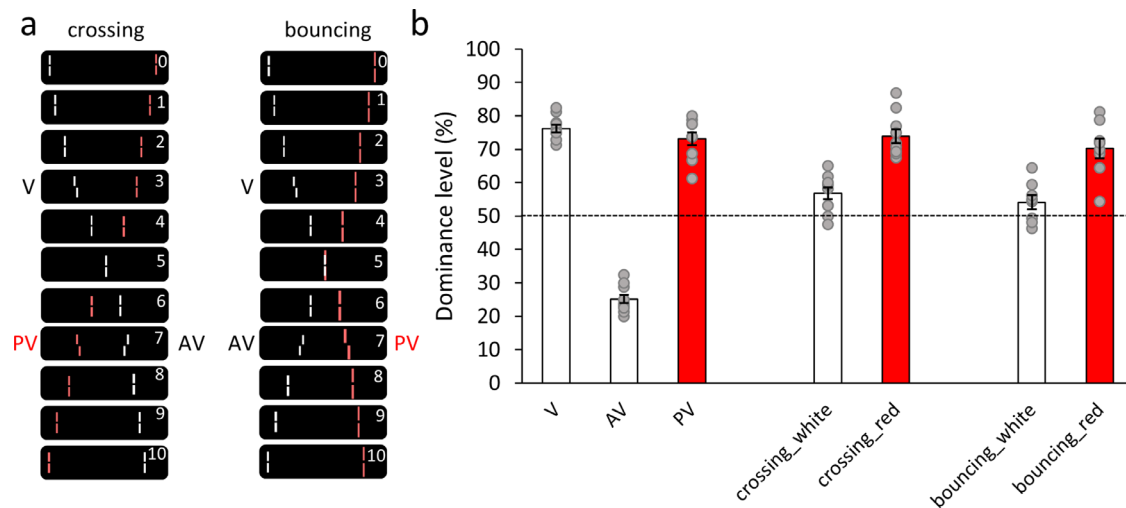


Figure 6. Experiment 3. (a) Crossing condition. The percept was a white and a red stream crossing. Bouncing condition. The red stream had longer lines than the white stream to reinforce the percept of two streams bouncing on each other. (b) Dominance level in the different conditions. White and red bars represent observers attending to the white and red stream, respectively. In the Crossing condition, the vernier (V) and anti-vernier (AV) offsets presented in the white stream integrated when observers attended to the white stream (crossing_white). When attending to the red stream, observers reported the direction of the offset presented in the red stream (crossing_red), which was in the same direction as V (PV). In the Bouncing condition, two observers perceived a mixture of crossing and bouncing streams. Data from these two observers is not included. For the other eight observers, V and AV integrate in the white stream (bouncing_white) and observers reported the direction of PV in the red stream (bouncing_red). Thus, only features that belong to the same object integrate. Circles indicate individual data. Error bars represent standard errors of the mean.

Experiment 3

Four conditions were tested: Crossing_white, Crossing_red, Bouncing_white, and Bouncing_red. In the Crossing_white and Crossing_red conditions, observers were presented the Crossing stimulus (see the Stimuli section) and instructed to attend to the white and red stream, respectively. In the Bouncing_white and Bouncing_red conditions, observers were presented with the Bouncing stimulus (see the Stimuli section) and instructed to attend to the white and red stream, respectively. Observers performed first the Crossing blocks and then the Bouncing blocks.

Subjective rating and report

Experiment 1

At the end of the experiment, observers were shown, in random order, the stimuli used in the three

conditions (Classic, Occluded, and Gap) without any offset. Observers were asked to rate their percept on a scale from 1 (“The motion stream appears completely disjointed”) to 6 (“The motion stream appears completely unified”).

Experiment 2

At the end of the experiment, observers were shown, in random order, the four stimuli presented during the experiment without any offset. Observers were asked to rate their percept on a scale from 1 (“The motion stream appears completely disjointed”) to 6 (“The motion stream appears completely unified”). When the conditions containing red elements were presented (Classic_red and Occluded_red), observers were also asked to rate their percept on a scale from 1 (“The red elements appear to be completely separated from the motion stream”) to 6 (“The red elements appear to be completely part of the motion stream”).

Experiment 3

Observers were shown the stimulus of the Crossing condition five times (without any offset) and asked to verbally describe the stimulus. Then, in a block of 80 trials, observers reported after each trial whether they perceived two crossing streams or two bouncing streams by pressing hand-held push buttons, that is, the right button when perceiving crossing streams and the left button when perceiving bouncing streams. All the participants reported perceiving crossing streams in every trial. Observers then completed the blocks of the Crossing condition without reporting their percept. Before the Bouncing condition blocks, observers were shown the stimulus of the Bouncing condition five times and asked to verbally describe the stimulus. Then, in a block of 80 trials, observers reported after each trial whether they perceived two crossing streams or two bouncing streams by pressing push buttons. Two observers perceived a mixture of crossing and bouncing streams. These two observers completed the Bouncing condition blocks with a report of their percept after each trial. The other eight participants perceived bouncing streams in all the trials and performed the Bouncing condition blocks without reporting their percept.

Analysis

Performance is quantified in terms of dominance, that is, the percentage of responses in accordance with the central vernier offset direction. Thus, dominance above 50% means that the central vernier offset dominates performance, dominance below 50% means that the anti-vernier dominates the performance, and a dominance around 50% means that none of them is dominant. For example, a dominance level of 25% means that the observer responded in 75% of the trials according to the direction of the anti-vernier offset.

Integration in the SQM is largely linear and well-predicted by the sum of the dominance levels in configurations V and AV when plotted between -50% and 50% (Otto et al., 2009). Here, we calculated the predicted integration dominance level as $[(V - 50) + AV]$ as dominance levels are plotted between 0% and 100% (Table 1). If the two offsets integrate, we expect the dominance in configuration V–AV to be not significantly different from the predicted integration dominance and significantly different from the dominance in configuration AV, in which only one offset is presented. Inversely, if the two offsets do not integrate, we expect the dominance in configuration V–AV to be significantly different from the predicted integration dominance and not significantly different from the dominance in configuration AV.

Experiment_ Condition	No. of observers	Mean predicted integration dominance level	SEM
1a_Classic	8	50.40	3.38
1a_Occluded	8	49.40	0.89
1a_Gap	8	47.50	2.25
1b_Classic	8	49.61	1.71
1b_Occluded	8	51.70	2.00
1b_Gap	8	49.63	4.00
2_Classic	8	46.8	2.1
2_Occluded	8	50.62	2.59
2_Classic_red	8	48.24	2.89
2_Occluded_red	8	49.40	2.39
3_Crossing_white	10	52.02	1.88
3_Crossing_red	10	48.97	2.5
3_Bouncing_white	8	51.3	2.03
3_Bouncing_red	8	47.15	2.92

Table 1. Predicted integration dominance levels from configurations V and AV. Integration in the SQM is largely linear and well predicted by the sum of the dominance levels in configurations V and AV when plotted between -50% and 50% (Otto et al., 2009). Here, we calculated the predicted integration dominance level as $[(V - 50) + AV]$ as dominance levels are plotted between 0% and 100% . The values are means and standard errors of the mean (SEM) of the number of observers that took part in each experiment.

Experiments 1a and 1b

Dominances in configuration V–AV in each condition (Classic, Occluded, and Gap) were compared with their respective predicted dominances and to the dominances in configuration AV.

Experiment 2

Dominances in configuration V–AV when observers were naïve in the Classic, Occluded, and Occluded_red conditions were compared with their respective predicted dominances and to dominances in configuration AV. When observers were aware of the paradigm and instructed to report the anti-vernier, the dominances in configuration V–AV $R[AV]$ was compared with the dominances when observers were naïve and to the dominances in configuration AV.

Experiment 3

Dominances in configurations V–AV in conditions Crossing_white and Bouncing_white were compared with their respective predicted dominances and to the dominance in configuration AV_white. Dominances in configurations V–AV in conditions Crossing_red and Bouncing_red were compared with their respective

predicted dominances and to the dominance in configuration PV_{red}.

For all comparisons, we used false discovery rate (FDR) corrected two-sided paired samples *t*-tests. Additionally, we combined the data of Experiments 1a and 1b and performed the same comparisons with Bayesian paired samples *t*-tests (Supplementary Table S2).

Dominance levels in each condition and for each observer in all Experiments are provided in Supplementary Table S4.

Offset calibration

Before each experiment, we calibrated the offsets for each participant to achieve comparable performances across observers and so that the central and the flanking lines' offsets have individually the same contribution. We presented the SQM with only one offset, that is, the central line or a flanking line was offset (configurations V and AV). A PEST adaptive procedure (Taylor & Creelman, 1967) was used to determine the offset sizes, aiming for around a 75% correct response rate and stopping after 80 trials, thereafter taking the respective value from the psychometric function that fitted the collected data best. The threshold and slope of the psychometric function (cumulative Gaussian) were estimated by means of a maximum likelihood analysis, taking all trials into account. The guessing rate was set to 50%, the rate of motor errors was set to 3%. A parametric bootstrap method was used to assess confidence intervals. Analysis was done on a logarithmic test level scale. Data across both offset directions (left and right) were pooled for the analysis. If the estimated threshold was computed by pure extrapolation of the experimental data, that is if the threshold was found to lie outside of the range of tested values, the block was discarded and repeated. The dominance levels of configurations V and AV plotted in each experiment's graph are the dominance levels obtained when using the offset size previously acquired from the PEST procedure. In Experiments 1a, 1b, and 2, the central (V) and flanking line (AV) offsets were calibrated for each condition tested in the experiment. In Experiment 3, the offsets were calibrated using the Crossing condition. The same offset sizes were used for both the Crossing and Bouncing conditions. The sizes of all the offsets are provided in Supplementary Table S1.

Results

Predicted dominance levels

Otto et al. (2009) showed that integration in the SQM is largely linear and well predicted by the sum of

the dominance levels in configurations V and AV when plotted between -50% and 50% . Here, the predicted integration dominance level is calculated as $[(V - 50) + AV]$ as dominance levels are plotted between 0% and 100% . The predicted dominance levels are close to 50% , which indicates integration when the vernier and the anti-vernier have the same weight (Table 1). We show only the 50% line in the graphs.

Experiment 1

We tested whether features integrate across an occluder versus when there is a gap in the sequence of the SQM (Experiments 1a and 1b). The rationale is that an occluder does not disrupt significantly the spatiotemporal contiguity of a moving object whereas a gap does. Accordingly, pre- and post-occlusion segments of the moving object should be grouped as a single motion stream, whereas the pre- and post-gap segments of the moving object should be grouped as different motion streams. Hence, the model predicts integration of features across the occluder but not across the gap.

Three conditions were presented (Figure 3a): the classic SQM, an Occluded condition, and a Gap condition. In the Occluded condition, a grey rectangle occluded the lines in frames 3, 4, and 5 of the attended stream. In the Gap condition, the same three lines were not presented. The anti-vernier was displayed in frame 7, thus after the occluder or the gap. At the end of the experiment, observers rated the spatiotemporal integrity (i.e., whether they are grouped into one continuous motion stream or not) of the three stimuli on a scale from 1 ("The motion stream appears completely disjointed") to 6 ("The motion stream appears completely unified").

In the Classic condition, observers perceived the stream as being unified and features integrated (Figure 3b, blue; V-AV [classic] vs. V-AV [classic_prediction]: $t(7) = 0.97$, $p = 0.37$, $p_{FDR} = 0.44$; V-AV [classic] vs. AV [classic]: $t(7) = 10.13$, $p = 1.96e-5$, $p_{FDR} = 8.8e-5$). Similarly, in the Occluded condition, observers perceived the stream as unified (Figure 3b, pink). In this condition, the offsets before and after the occluder integrated (V-AV [occluded] vs. V-AV [occluded_prediction]: $t(7) = 0.75$, $p = 0.48$, $p_{FDR} = 0.48$; V-AV [occluded] vs. AV [occluded]: $t(7) = 9.53$, $p = 2.94e-5$, $p_{FDR} = 8.8e-5$). However, there was less integration in the Gap condition (Figure 3b, purple; V-AV [gap] vs. V-AV [gap_prediction]: $t(7) = 3.01$, $p = 0.02$, $p_{FDR} = 0.04$; V-AV [gap] vs. AV [gap]: $t(7) = 1.21$, $p = 0.27$, $p_{FDR} = 0.4$). Observers were able to report the offset direction of the anti-vernier more often than in the Classic and Occluded conditions. In the Gap condition, the stream appeared more disjointed to the observers than in the other conditions.

Experiment 1b was identical to Experiment 1a, except for the trajectory of the streams. The streams diverged from the center until frame 4 and then converged back to the center (Figure 4a). Thus, there were two main differences between Experiments 1a and 1b: 1) reversal of motion direction and 2) the fact that the second vernier was on the same side of the occluder/gap as the central vernier. According to the two-stage model, because the first stage provides a reference frame along the motion path of grouped elements, integration should persist even when the direction of motion changes, as long as elements are grouped into the same motion stream. Similarly, the exact locations of features do not matter for integration as long as they are part of the same perceptual group. Results are similar to those of Experiment 1a (Figure 4b). The offsets integrated in the Classic (V–AV [classic] vs. V–AV [classic_prediction]: $t(7) = 0.17$, $p = 0.87$, $p_{FDR} = 0.94$; V–AV [classic] vs. AV [classic]: $t(7) = 3.54$, $p = 0.01$, $p_{FDR} = 0.02$) and Occluded conditions (V–AV [occluded] vs. V–AV [occluded_prediction]: $t(7) = 1.15$, $p_{FDR} = 0.29$; V–AV [occluded] vs. AV [occluded]: $t(7) = 5.57$, $p = 8.45e-4$, $p_{FDR} = 0.005$), whereas less integration happened in the Gap condition (V–AV [gaps] vs. V–AV [gap_prediction]: $t(7) = 3.97$, $p = 0.005$, $p_{FDR} = 0.016$; V–AV [gap] vs. AV [gap]: $t(7) = 0.074$, $p = 0.94$, $p_{FDR} = 0.94$). Perceptually, the stream appeared more unified in the Classic and Occluded conditions than in the Gap condition.

We also combined the data of Experiments 1a and 1b, and performed Bayesian paired samples t -tests. The results are in the same direction as the main analysis (Supplementary Table S2).

Experiment 2

Here, we tested whether integration is mandatory across occlusion, that is, whether observers are able to report the direction of the anti-vernier ignoring the central vernier. We also tested whether offsets integrate when the stream changes color in the middle of the trajectory, with and without an occluder.

In the first part of the experiment, observers were naïve. In the second part of the experiment, observers were informed of the paradigm, that is that two offsets were presented, and observers were instructed to report the direction of the anti-vernier (labelled R[AV]). These instructions were used to assess whether integration was mandatory in the different conditions. At the end of the experiment, observers rated their percept on a scale from 1 (“The motion stream appears completely disjointed”) to 6 (“The motion stream appears completely unified”). When the conditions containing red elements were presented (Classic_red and Occluded_red), observers were also

asked to rate their percept on a scale from 1 (“The red elements appear to be completely separated from the motion stream”) to 6 (“The red elements appear to be completely part of the motion stream”).

In all conditions, except the Classic_red condition, observers rated the stream as unified (i.e., subjective ratings between 4 and 6; Figures 5b, 5c, and 5e). The offsets integrated in these conditions (V–AV[classic] vs. V–AV[classic_prediction]: $t(7) = 0.18$, $p = 0.86$, $p_{FDR} = 0.86$; V–AV[classic] vs. AV[classic]: $t(7) = 5.01$, $p = 0.002$, $p_{FDR} = 0.005$; V–AV[occluded] vs. V–AV[occluded_prediction]: $t(7) = 0.39$, $p = 0.71$, $p_{FDR} = 0.77$; V–AV[occluded] vs. AV[occluded]: $t(7) = 4.05$, $p = 0.005$, $p_{FDR} = 0.012$; V–AV[occluded_red] vs. V–AV[occluded_red_prediction]: $t(7) = 1.3$, $p = 0.24$, $p_{FDR} = 0.28$; V–AV[occluded_red] vs. AV[occluded_red]: $t(7) = 5.63$, $p = 7.94e-4$, $p_{FDR} = 0.005$). Moreover, integration was mandatory, that is, observers were largely unable to report the direction of the anti-vernier only (V–AV[classic R[AV]] vs. V–AV[classic]: $t(7) = 1.3$, $p = 0.23$, $p_{Holm} = 0.28$; V–AV[classic R[AV]] vs. AV[classic]: $t(7) = 5$, $p = 0.002$, $p_{FDR} = 0.005$; V–A [occluded R[AV]] vs. V–AV[occluded]: $t(7) = 1.98$, $p = 0.088$, $p_{FDR} = 0.15$; V–AV[occluded R[AV]] vs. AV [occluded]: $t(7) = 3.86$, $p = 0.006$, $p_{FDR} = 0.012$; V–AV[occluded_red R[AV]] vs. V–AV[occluded_red]: $t(7) = 1.82$, $p = 0.11$, $p_{FDR} = 0.17$; V–AV[occluded_red R[AV]] vs. AV[occluded_red]: $t(7) = 5.43$, $p = 9.78e-4$, $p_{FDR} = 0.005$). Hence, we found that offsets largely integrated mandatorily across the occlusion, even when the color of the stream changed after the occluder. The stream was still perceived as being unified and the red elements as part of the stream (subjective ratings between 4 and 6; Figure 5g). It is noteworthy that, although integration is largely mandatory, the dominance when observers are instructed to report the anti-vernier is lower than the dominance when observers are naïve in the three conditions. This finding suggest that observers reported the anti-vernier in more trials when they were instructed to.

In the Classic_red condition, some observers perceived the stream as being unified and the red elements as part of the stream, whereas others did not (subjective ratings between 1 and 6; Figures 5d and 5f). Integration seems to follow the subjective ratings, particularly when observers were instructed to report the anti-vernier. We found less integration when the stream was perceived as less unified and the red elements as less part of the stream.

It is noteworthy that, for one observer, the offsets did not integrate in any condition. When the observer was naïve, the central vernier dominated performance (dominance level between 60% and 79.4%) whereas the anti-vernier was reported when the observer was instructed to (dominance level between 31.25% and 39.4%).

Experiment 3

Here, we investigated integration when two objects were presented. We used two streams that can be perceived as either crossing (Figure 6a, left) or bouncing against each other (Figure 6a, right). This manipulation allows two perceptual grouping outcomes: In the bouncing case, each moving stimulus is perceived to reverse direction whereas in the Crossing condition they maintain the same motion direction. Accordingly, the reference frame reverses direction in the Bouncing condition and not in the Crossing condition. Hence, the two-stage model predicts integration according to these reference frames. One stream was white and the other one was red. Observers were instructed to attend to either the white or the red stream and to report the direction of the offset perceived at the end of the trajectory of the stream. The white stream contained a vernier offset in frame 3 (which we label V) and an anti-vernier in frame 7 (AV). The red stream contained only one offset in frame 7. This offset was in the same direction as V (PV).

In the Bouncing condition, the red stream had longer lines to reinforce a percept of streams that bounce on each other as there is usually a bias toward crossing percepts. However, when featural differences increase after the collision point, the proportion of bouncing percepts increases (e.g., Feldman & Tremoulet, 2006). In the Crossing condition, all participants perceived crossing streams in every trial. Data from all observers was pooled together. In the Bouncing condition, two

observers perceived a mixture of crossing and bouncing streams. Data of these two observers were not pooled with the data of the other eight observers who only perceived bouncing streams.

When observers attended to the white stream in the Crossing condition, offsets integrated (Figure 6b; crossing_white vs. crossing_white_prediction: $t(9) = 1.1, p = 0.31, p_{FDR} = 0.49$; crossing_white vs. AVwhite: $t(9) = 16.1, p = 6.13e-8, p_{FDR} = 4.91e-7$). When attending to the red stream, observers reported the direction of the only offset presented in the red stream (crossing_red vs. crossing_red_prediction: $t(9) = 7.3, p = 4.58e-5, p_{FDR} = 1.2e-4$; crossing_red vs. PVred : $t(9) = 0.22, p = 0.83, p_{FDR} = 0.83$). Similarly, offsets in the white stream integrated in the Bouncing condition when observers attended to the white stream (bouncing_white vs. bouncing_white_prediction: $t(7) = 0.8, p = 0.45, p_{FDR} = 0.6$; bouncing_white vs. AVwhite: $t(7) = 10.8, p = 1.3e-5, p_{FDR} = 5.2e-5$). Observers reported the direction of the only offset in the red stream when instructed to attend to this stream (bouncing_red vs. bouncing_red_prediction: $t(7) = 4.3, p = 0.004, p_{FDR} = 0.007$; bouncing_red vs. PV7red: $t(7) = 0.59, p = 0.57, p_{FDR} = 0.65$). Thus, only features that belong to the same object integrate.

As mentioned, two observers perceived a mixture of crossing and bouncing streams in the Bouncing condition. The data did not match the expected results in which integration follows the percept of crossing or bouncing streams (Figure 7). Data from each observer and the number of trials that are perceived as

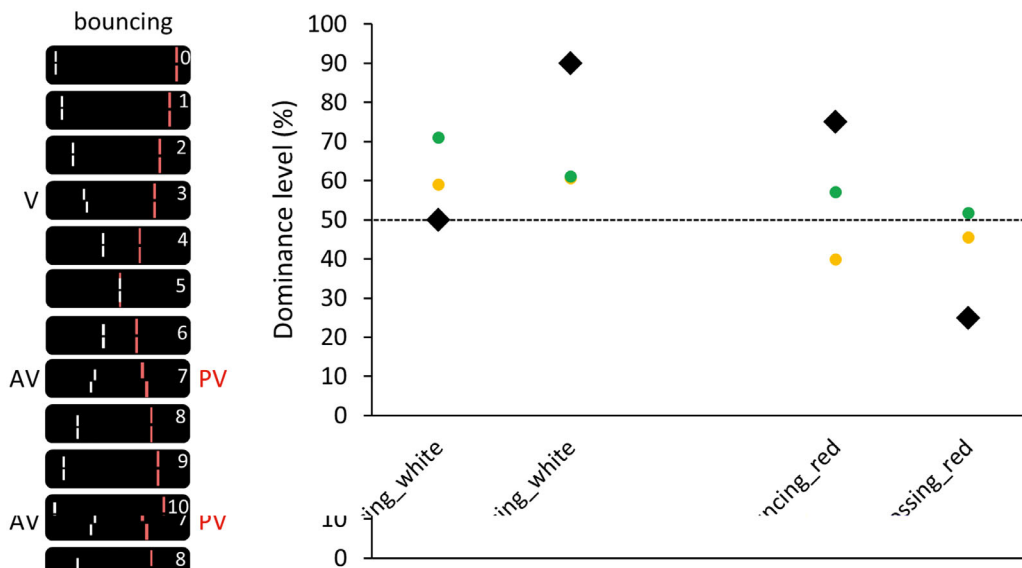


Figure 7. Data from the two observers that perceived a mixture of crossing and bouncing streams in the Bouncing condition. Data from one observer are in orange and the other observer’s data are in green. Black diamonds indicate the expected dominance level when integration follows the perceived trajectory of the streams. For example, when attending to the white stream, if a bouncing trajectory is perceived,

V and AV should integrate, and dominance level should be around 50%. If a crossing trajectory is perceived, V and PV should integrate, and dominance level should be above 75%. The data do not match the expected results.

crossing or bouncing in each condition are provided in Supplementary Table S3.

Discussion

We previously showed that features integrate in long-lasting windows of unconscious processing (Drissi-Daoudi et al., 2019). We argued that these windows determine periods during which the brain tries to make sense of incoming information (Herzog et al., 2016, 2020). We propose that, as shown in Figure 1, during such a temporal window, Gestalt figure-ground segregation and grouping operations are at play, establishing a reference frame for each group (Ögmen & Herzog, 2010). Object identities are then determined by these reference frames, which are used to attribute features to stimuli. Finally, features are integrated according to object identities.

In the SQM, grouping is indeed important. Otto et al. (2006) presented parallel streams with a central vernier offset and a flank vernier offset in the right stream (attention to the right stream). When an additional line was presented to the left of the central line, the offsets integrated, whereas when the additional line was presented to the right of the central line, the offsets did not integrate. The authors suggest that the additional line grouped differently the central vernier offset with the streams. Hence, the offsets integrated only when they were part of the same stream.

Here, we first investigated whether offsets integrate across occlusion and across a gap in the motion trajectory (Experiment 1). At the end of the experiment, observers were asked to rate the spatiotemporal integrity of the streams. When some lines were missing, creating a gap in the motion trajectory, observers perceived the stream as less unified. Accordingly, the offsets integrated less. However, offsets integrated across an occluder and observers rated the motion stream as unified when part of the trajectory was occluded. Even when the stream changed color after the occluder, the offsets integrated (Experiment 2). These results illustrate well the principle of spatiotemporal priority, that is, that spatiotemporal factors usually trump featural considerations for object persistence (Flombaum, Scholl, & Santos, 2009; Scholl, 2001). For example, if a moving disk passes behind an occluder and reappears with a different color, it will be perceived as one single object having changed color, provided that the disk comes out at the right place and time (Burke, 1952; Michotte, Thine, & Crabbé, 1964). Similarly, here, the pre- and post-occluded parts of the SQM are perceived as being the same object, even though the offsets are not in the same direction, and even when the stream changes color after the occluder.

Hence, the features integrated. Using the Ternus Pikler display, we have shown that vernier offset information can be held in memory and attributed across non-retinotopic positions (Scharnowski, Hermens, Kammer, Ögmen, & Herzog, 2007). A similar mechanism may be at work here (see also Ögmen & Herzog, 2016).

Importantly, integration is mandatory in the SQM. When observers were informed of the paradigm and instructed to report the direction of the offset presented after the occluder, they were largely unable to access this offset independently from the central vernier offset (Experiment 2). It is noteworthy that the dominance when observers are instructed to report the anti-vernier is lower than the dominance when observers are naïve, suggesting that the observers reported the direction of the anti-vernier more often. It is possible that the anti-vernier becomes “stronger” due to the knowledge of the stimulus and the instruction to report the anti-vernier.

Spatiotemporal factors are also usually prioritized in apparent motion. For example, if a disk and a star are flashed rapidly one after the other at different locations, the percept is one moving object that changes its shape from a disk to a star, rather than two separate objects with different shapes (Kolers, 2013; Kolers & Pomerantz, 1971). In Experiment 2, we also presented streams that changed color in the middle of the trajectory, however, without an occluder (Classic_red condition). As predicted by the principle of spatiotemporal priority, some observers perceived the stream in the Classic_red condition as being unified and the red elements as part of the stream, however others did not. It might be that, for the latter observers, the featural change was too important and too abrupt to be ignored. It has been suggested that the visual system also uses featural, and not only spatiotemporal, information to resolve object correspondence (Hein & Moore, 2012; Moore, Stephens, & Hein, 2020). Interestingly, dominance levels seem to follow observers’ subjective ratings. Offsets integrated less when the streams were rated as less unified and the red elements as less part of the stream.

Finally, we presented two distinct streams using color, or color and size, as grouping cues (Experiment 3). Although the trajectories of the streams were somewhat blended, as they were crossing each other or bouncing against each other, the offsets did not get mingled. Only offsets belonging to the same stream integrated. For two observers, however, this was not the case in the Bouncing condition. In contrast with the other observers, they perceived a mixture of crossing and bouncing streams. Their data did not show a link between integration and their percept. It is noteworthy that both these observers reported that the stimulus was highly ambiguous and the percept hard to define. This

might explain the observed pattern. Moreover, these data came from only two participants. More observers are needed to further investigate the case in which both percepts occur.

In a previous study, we investigated feature integration in the SQM across saccades (Drissi-Daoudi, Ögmen, Herzog, & Cicchini, 2020). We found that features mandatorily integrate when object identity is preserved in the external world, even when an eye movement was executed between the presentation of the central vernier and the anti-vernier. Hence, object identity determines feature integration with and without eye movements.

Here, we changed the grouping of elements using space. It is also possible to modulate grouping using time, such as in the Ternus–Pikler display (Boi, Ögmen, Krummenacher, Otto, & Herzog, 2009; Boi, Vergeer, Ögmen, & Herzog, 2011; Pikler, 1917; Ternus, 1926). In the Ternus–Pikler display, two central elements are aligned horizontally. When a third element is presented alternately to the right and to the left of the two central elements, two percepts are possible depending on the ISI. For a short ISI (e.g., 50 ms), the two central elements seem to flicker at the same position and the external element appears to jump for right to left (“element motion”). For a longer ISI (e.g., 200 ms), the three elements appear to move together to the left and to the right (“group motion”). Ögmen, Otto, & Herzog (2006), using the Ternus–Pikler display with verniers, showed that feature attribution and feature integration follows perceptual grouping, that is, whether observers perceived element or group motion. Moreover, feature integration, like in the SQM, was mandatory in the Ternus–Pikler display and unconscious information plays an important role (Lauffs, Choung, Ögmen, & Herzog, 2018).

Taken together, these results show that the visual system uses motion information to establish reference frames that allow the attribution and integration of features to objects whether the moving stimuli overlap retinotopically (as in Ternus–Pikler displays) or not (as in SQM), or whether motion is induced by the observer (as in the aforementioned saccade study). This ability seems to be a natural outcome of ecological perception, where motion (observer and/or object motion) and occlusions are abundant. The essence of spatiotemporal integration, based on object identity, is also expressed in the object-file theory (Kahneman, Treisman, & Gibbs, 1992), which took inspiration from file-based storage systems in computers. According to this theory, an object-file is opened for each object and the attributes of the moving object are inserted into this file based on spatiotemporal continuity. However, beyond the conceptual computer analogy, the theory gives no details or mechanisms to explain how object

files are opened and how information is inserted over the motion pathway.

Motion grouping based reference frames of the two-stage model can be viewed as a mechanistic expression of the concept of spatiotemporal priority (Agaoglu, Clarke, Herzog, & Ögmen, 2016; Clarke, Ögmen, & Herzog, 2016; Ögmen & Herzog, 2010). Instead of object files, it relies on the geometric concept of reference frames that are synthesized by perceptual-grouping operations. Vision starts with an egocentric reference frame (retinotopic reference frame which is based on the retina of the observer) whereas perception is dominated by exocentric reference frames (e.g., a reference frame based on the motion path of a stimulus) (Agaoglu, Herzog, & Ögmen, 2015; Huynh, Tripathy, Bedell, & Ögmen, 2017). Exocentric reference frames have been identified in the primate nervous system (Olson, 2003; Zaehle et al., 2007). The two-stage model suggests that Gestalt grouping and reference frame synthesis go hand-in-hand and constitute the early stages of computations that allow the processing of dynamic stimuli by taking into account both observer’s and external objects’ motion.

Keywords: feature integration, grouping, reference frames

Acknowledgments

The authors thank Marc Repnow for technical support. This work was supported by a grant from the Swiss SystemsX.ch initiative (2015/336) and by the Swiss National Science Foundation grant ‘Basics of visual processing: from elements to figures’ (176153).

Commercial relationships: none.

Corresponding author: Leila Drissi-Daoudi.

Email: leila.drissidaoudi@gmail.com.

Address: Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 19, Lausanne 1015, Switzerland.

References

- Agaoglu, M. N., Clarke, A. M., Herzog, M. H., & Ögmen, H. (2016). Motion-based nearest vector metric for reference frame selection in the perception of motion. *Journal of Vision*, 16(7), 14–14.
- 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: Official Publication of the American Academy of Optometry*, 73(1), 49–53.
- Bachmann, T., & Francis, G. (2013). *Visual masking: Studying perception, attention, and consciousness*. Cambridge, MA: Academic Press.
- 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), 5–5, <https://doi.org/10.1167/9.13.5>.
- Boi, M., Vergeer, M., Ogmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology*, 21(20), 1732–1737.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436, <https://doi.org/10.1163/156856897X00357>.
- Breitmeyer, B., & Ögmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision*. Oxford, UK: Oxford University Press.
- Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, 4(3), 121–138.
- 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.
- Drissi-Daoudi, L., Doerig, A., & Herzog, M. H. (2019). Feature integration within discrete time windows. *Nature Communications*, 10(1), 1–8, <https://doi.org/10.1038/s41467-019-12919-7>.
- Drissi-Daoudi, L., Ögmen, H., Herzog, M. H., & Cicchini, G. M. (2020). Object identity determines trans-saccadic integration. *Journal of Vision*, 20(7), 33–33, <https://doi.org/10.1167/jov.20.7.33>.
- Elliott, M. A., & Giersch, A. (2016). What happens in a moment. *Frontiers in Psychology*, 6, 1095, <https://doi.org/10.3389/fpsyg.2015.01905>.
- Feldman, J., & Tremoulet, P. D. (2006). Individuation of visual objects over time. *Cognition*, 99(2), 131–165, <https://doi.org/10.1016/j.cognition.2004.12.008>.
- Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2009). Spatiotemporal priority as a fundamental principle of object persistence. *The Origins of Object Knowledge*, Oxford, England: Oxford University Press, 135–164.
- 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., Drissi-Daoudi, L., & Doerig, A. (2020). All in good time: Long-lasting postdictive effects reveal discrete perception. *Trends in Cognitive Sciences*, 24(10), 826–837, <https://doi.org/10.1016/j.tics.2020.07.001>.
- Herzog, M. H., Kammer, T., & Scharnowski, F. (2016). Time slices: What is the duration of a percept? *PLoS Biology*, 14(4), e1002433, <https://doi.org/10.1371/journal.pbio.1002433>.
- Huynh, D., Tripathy, S. P., Bedell, H. E., & Ögmen, H. (2017). The reference frame for encoding and retention of motion depends on stimulus set size. *Attention, Perception, & Psychophysics*, 79(3), 888–910.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175–219, [https://doi.org/10.1016/0010-0285\(92\)90007-O](https://doi.org/10.1016/0010-0285(92)90007-O).
- Kolers, P. A. (2013). *Aspects of motion perception: International series of monographs in experimental psychology*. New York: Elsevier.
- Kolers, P. A., & Pomerantz, J. R. (1971). Figural change in apparent motion. *Journal of Experimental Psychology*, 87(1), 99.
- 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>.
- Michotte, A., Thine, G. O., & Crabbé, G. (1964). *Les compléments amodaux des structures perceptives*. Ottignies-Louvain-la-Neuve, Belgium: Institut de psychologie de l'Université de Louvain.
- Moore, C. M., Stephens, T., & Hein, E. (2020). Object correspondence: Using perceived causality to infer how the visual system knows what went where. *Attention, Perception, & Psychophysics*, 82(1), 181–192, <https://doi.org/10.3758/s13414-019-01763-y>.
- 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>.
- Ögmen, H., & Herzog, M. H. (2016). A new conceptualization of human visual sensory-memory. *Frontiers in Psychology*, 7, 830, <https://doi.org/10.3389/fpsyg.2016.00830>.
- 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.

- Olson, C. R. (2003). Brain representation of object-centered space in monkeys and humans. *Annual Review of Neuroscience*, 26(1), 331–354.
- 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.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10 (4), 437–442.
- Piéron, H. (1935). Le processus du métacontraste. *Journal de Psychologie Normale et Pathologique*, 32, 5–24.
- Pikler, J. (1917). *Sinnesphysiologische Untersuchungen*, Leipzig, Germany: J.A. Barth.
- 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>.
- Scholl, B. J. (2001). Spatiotemporal priority and object identity. *Current Psychology of Cognition*, 20(5), 359–372.
- Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 41(4A), 782–787, <https://doi.org/10.1121/1.1910407>.
- Ternus, J. (1926). Experimentelle untersuchungen über phänomenale Identität. *Psychologische Forschung*, 7(1), 81–136.
- Zaehle, T., Jordan, K., Wüstenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2007). The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, 1137, 92–103.