



How voluntary spatial attention influences feature biases in object correspondence

Madeleine Y. Stepper¹ · Bettina Rolke¹ · Elisabeth Hein¹

Published online: 28 June 2019

© The Psychonomic Society, Inc. 2019

Abstract

Our visual system is able to establish associations between corresponding images across space and time and to maintain the identity of objects, even though the information our retina receives is ambiguous. It has been shown that lower level factors—as, for example, spatiotemporal proximity—can affect this correspondence problem. In addition, higher level factors—as, for example, semantic knowledge—can influence correspondence, suggesting that correspondence might also be solved at a higher object-based level of processing, which could be mediated by attention. To test this hypothesis, we instructed participants to voluntarily direct their attention to individual elements in the Ternus display. In this ambiguous apparent motion display, three elements are aligned next to each other and shifted by one position from one frame to the next. This shift can be either perceived as all elements moving together (group motion) or as one element jumping across the others (element motion). We created a competitive Ternus display, in which the color of the elements was manipulated in such a way that the percept was biased toward element motion for one color and toward group motion for another color. If correspondence can be established at an object-based level, attending toward one of the biased elements should increase the likelihood that this element determines the correspondence solution and thereby that the biased motion is perceived. Our results were in line with this hypothesis providing support for an object-based correspondence process that is based on a one-to-one mapping of the most similar elements mediated via attention.

Keywords Attention · Motion: Apparent · Perceptual organization

Introduction

Imagine several kittens playing, jumping across each other, and all looking very much alike. Keeping track of one of these kittens subjectively seems to be easy for us. But the task is not as easy as it seems, as the information our visual system receives is ambiguous and not continuous—for example, because the kittens could occlude each other while moving around, or they could reappear spatially shifted from behind a cupboard. Building up correspondence that is, establishing associations between images across space and time and maintaining the identity of an object, like our individual kittens, is therefore difficult. How our visual system solves this “correspondence problem” (Ullmann, 1979) has been a topic of research for decades.

Much of this research has used apparent motion displays (Wertheimer, 1912), in which no physical motion is present, but two successively presented objects are perceived as one single moving object. It has been shown that spatiotemporal factors, like the specific spatial distance and the time interval between the occurrences of the objects, is important to establish correspondence between the objects and to perceive apparent motion (e.g., Korte, 1915). The correspondence problem is especially obvious for ambiguous apparent motion displays, like, for example, the motion quartet (von Schiller, 1933), for which different correspondence solutions are possible. The motion quartet consists of two elements presented at opposing edges of a fictive square alternating with two more elements at the other two edges. Depending on the distance and the temporal interval between successively presented elements, the elements can be perceived as moving horizontally or vertically. For example, reducing the horizontal distance between the elements results in the perception of more horizontal movements, and reducing the vertical distance results in the perception of more vertical movements (e.g., Hock, Kelso, & Schöner, 1993; von Schiller, 1933). Temporal factors have also been shown to strongly influence another

✉ Madeleine Y. Stepper
madeleine.stepper@uni-tuebingen.de

¹ Evolutionary Cognition—Cognitive Science, Department of Psychology, University of Tübingen, Schleichstrasse 4, D - 72076 Tübingen, Germany

ambiguous apparent motion display, the Ternus display (Pikler, 1917; Ternus, 1926). The Ternus display usually consists of three elements presented next to each other, shifted by one element position in the next frame (see Fig. 1a). Depending on how correspondence is solved, all elements can be perceived as moving together—that is, each element moving to the position of the adjacent element (group motion)—or one element can be perceived as jumping across the others that remain stationary (element motion). Which type of motion is perceived strongly depends on the time between the successively presented stimuli frames (ISI), as the probability to perceive group motion increases with increasing ISI (Pantle & Petersik, 1980; Petersik & Pantle, 1979). Taken together, it has been shown that the spatiotemporal relationship between stimulus occurrences strongly influences the way correspondence is established and apparent motion is perceived.

Another factor that influences correspondence is feature information (e.g., Alais & Lorenceau, 2002; Casco, 1990; Dawson, Nevin-Meadows, & Wright, 1994; Hein & Cavanagh, 2012; Hein & Moore, 2012; Kramer & Rudd, 1999; Kramer & Yantis, 1997; Moore & Enns, 2004; Petersik & Rice, 2008; Wallace & Scott-Samuel, 2007). For example, Alais and Lorenceau (2002) used a Ternus display with Gabor patches as Ternus elements. These patches could be oriented either collinearly (i.e., gratings oriented horizontally) or parallel (i.e., gratings oriented vertically). The authors showed that more group motion was perceived for the collinearly oriented

elements compared with the elements oriented in parallel, suggesting that the feature information of the elements within a frame could influence the correspondence solution. Hein and Moore (2012) manipulated the appearance of the individual elements in the Ternus display in a way that the elements were either compatible with the element motion percept (element bias; see Fig. 1b, top display) or compatible with the group motion percept (group bias; see Fig. 1b, bottom display). They showed that the motion percept is shifted in the direction of the bias: For the group bias, more group motion was perceived, and for the element bias, more element motion was perceived compared with a display without such biases (i.e., all elements were identical; see Fig. 1a). Different feature biases (i.e., color, polarity, orientation, hue and luminance) all strongly influenced the correspondence solution (Hein & Moore, 2012). These findings suggest that the identity of the elements across frames also strongly influences correspondence.

Finally, there is evidence that in addition to these rather lower level factors—spatiotemporal and feature information—even more complex, higher level information can influence how correspondence is determined. For example, lexical information (Chen & Zhou, 2011; Tse & Cavanagh, 2000), the global context (He & Ooi, 1999; Ramachandran & Anstis, 1983), the perceived size and lightness (He & Nakayama, 1994; Hein & Moore, 2014), semantic information (Hsu, Taylor, & Pratt, 2015; Yu, 2000), as well as attention (Aydın, Herzog, & Ögmen, 2011; Kohler, Haddad, Singer, & Muckli, 2008; Suzuki & Peterson, 2000; Wertheimer, 1912; Xu, Suzuki,

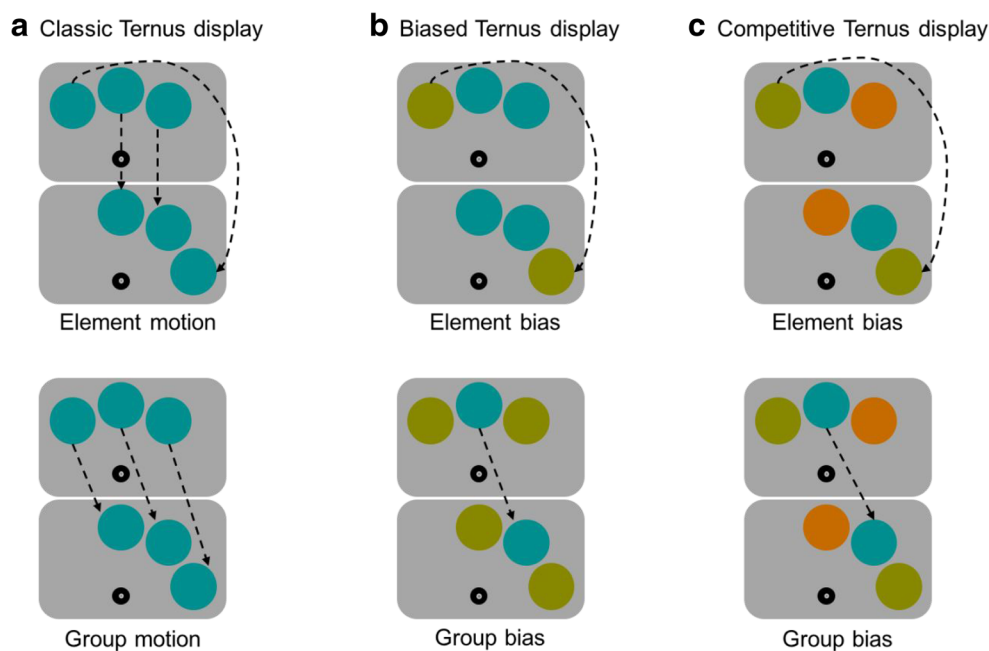


Fig. 1 Three different Ternus display types. **a** Classic Ternus display, in which all elements have the same color and either element motion or group motion can be perceived. **b** Biased Ternus display, in which one differently colored element is either compatible with the element motion percept (element bias, here olive green) or the group motion percept

(group bias, here cyan). **c** Competitive Ternus display, in which all elements have different colors, arranged in such a way that the display contains an element bias (here olive green) and a group bias (here cyan) at the same time. (Color figure online)

& Franconeri, 2013) modulate the perception of apparent motion. Regarding the influence of attention, Kohler et al. (2008), for example, instructed participants to voluntarily control the perceived motion direction in the motion quartet (von Schiller, 1933). The results showed that participants were able to hold an intended motion direction twice as long than in a passive viewing condition, in which participants were instructed to just report their motion percept. Moreover, they were also able to switch between the vertical and horizontal moving directions twice as fast as compared with the automatic switching in a passive viewing condition. Specifically for the Ternus display, Aydın et al. (2011) investigated if the availability of attentional resources influences the apparent motion percept using a dual-task paradigm. In the dual-task condition, participants had to detect and count the occurrence of a particular form in a stream of different forms at fixation, in addition to judging the motion of the Ternus display in the periphery. They showed that in the dual-task condition, less group motion was perceived compared with a control condition in which attention was fully available for the Ternus display. The authors concluded that more attention is needed for perceiving group motion compared with element motion. Thus, studies have shown that besides spatiotemporal factors and feature information, even higher level factors, as, for example, attention directed toward a particular motion percept can influence how correspondence is solved.

To explain the influence of these different factors, several theories have been developed. Some of these theories emphasize the importance of spatiotemporal factors, as, for example, motion energy models (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Werkhoven, Sperling, & Chubb, 1993). According to these models, low-level motion detectors—that is, Reichardt detectors (Reichardt, 1961)—compute motion energy based on spatiotemporal activation changes. The direction of these changes then constitutes the basis for determining apparent motion. These theories thus can account particularly well for effects of the ISI and the spatial distance. In line with motion energy models other theories, as for example the spatiotemporal priority theory (Flombaum, Scholl, & Santos, 2012) or the object-file theory (Kahneman, Treisman, & Gibbs, 1992) have highlighted spatiotemporal information as the most important factor to establish correspondence, whereas the identity of an object in terms of its feature information should play no or only a minor role.

To account more directly for the influence of feature information on correspondence, grouping theories have been proposed (e.g., Alais & Lorenceau, 2002; He & Ooi, 1999; Kramer & Yantis, 1997). These theories suggest that correspondence depends on how strongly the objects are associated or grouped with one another based on their features, following, for example, general grouping principles, as the similarity or proximity of the objects (e.g., Wertheimer, 1923). In particular, Kramer and Yantis (1997) suggested that the stronger the spatial grouping of the elements within a Ternus frame is, the

more group motion should be perceived. Moreover, the stronger the temporal grouping of the overlapping elements across Ternus frames is, the more element motion should be perceived. This grouping mechanism could explain, for example, the findings by Alais and Lorenceau (2002) that collinearly oriented elements increased group motion percepts, as these elements should lead to more spatial grouping due to facilitated contour interactions. In addition, element biases, as shown, for example, by Hein and Moore (2012), could be easily explained as the spatial grouping (i.e., within a Ternus frame) should be decreased, and the temporal grouping (i.e., across both Ternus frames) increased in these conditions.

Finally, to account for the influence of higher level factors and feature-based biases, Hein and colleagues suggested an object-based theory of correspondence (Hein & Cavanagh, 2012; Hein & Moore, 2014). According to this theory correspondence is established by a one-to-one mapping—that is, each individual element in one frame is connected with the perceptually most similar element in the next frame. Thus, in contrast to grouping theories, perceived motion is not based on the similarity of all elements within a frame (spatial grouping), but all individual elements across frames are matched based on their identity. Such an object-based theory could explain feature biases as well as high-level influences of lexical or semantic knowledge on correspondence that are difficult to explain with grouping or motion energy theories. Hein and Cavanagh (2012) suggested that attentional pointers (Cavanagh, 1992; Cavanagh et al. 2010)—that is, spatiotopically organized location pointers that are based on identity information—could connect the most similar elements across frames and track them over space and time, attention thus being a key mechanism for correspondence. Such a correspondence process could happen at a relatively high level of processing such that the similarity of the objects, even in terms of lexical or semantic knowledge and the global context could be taken into account by this type of correspondence process.

The aim of the current study was to directly test the object-based theory of correspondence (Hein & Cavanagh, 2012; Hein & Moore, 2014) by further investigating the potential influence of spatial attention on object correspondence. Following this theory orienting attention to an object should make this object more likely to determine the correspondence solution, as it should orient the attentional pointers toward that object. To test this idea, we run two experiments in which we used a biased Ternus display and instructed participants to direct their attention to one of the elements. In particular, we created a Ternus display containing a competitive bias (Hein & Schütz, 2019), for which differently colored elements were arranged in a way that, across both frames, the percept was biased toward group motion by one color and element motion by another color (see Fig. 1c). Additionally, we used a classic Ternus display, in which all elements had the same color (see

Fig. 1a). Attention was manipulated by using a precue to one of the Ternus elements (e.g., Posner, 1980). The precue consisted of a written word presented at the beginning of each trial that indicated which element of the first Ternus frame participants should attend (left, center, right, or all). Participants had to indicate whether they perceived group or element motion in the Ternus display (Ternus task). As orienting attention was not necessary to solve the Ternus task, we used an additional discrimination task to independently verify whether attention had been oriented successfully. The Ternus task was identical in the two experiments, but they differed concerning this additional discrimination task. In Experiment 1, participants were asked to discriminate the orientation of a Landolt C that was briefly presented on one of the Ternus elements. Ternus task and discrimination task were randomly intermixed and thus participants could not anticipate the specific task of a given trial when processing the cue. In Experiment 2, we separated the Ternus task from the discrimination task to avoid potential dual-task costs. In addition, instead of a difficult Landolt C discrimination task, we used a simple gap detection task.

According to the object-based correspondence theory with its attentional pointers (Hein & Cavanagh, 2012; Hein & Moore, 2014), we expected an influence of attention in the competitive display condition, as attending a specific element should make this element more likely to be connected with the element of that particular color feature in the next frame, and thus it should be more likely that this element determines the correspondence solution. In particular, we expected more perceived group motion when the element containing the group bias was attended (GB-match condition; center element; cyan element in Fig. 1c) compared with when the element containing the element bias was attended (EB-match condition; left element; green element in Fig. 1c). If attention was oriented to the third element, no particular effect was expected, as there was no direct feature-based match toward the element or group motion percept. In the classic display condition, with all elements being the same, orienting attention to one of the elements should not have any specific effect, as there would be no particular good match of the attended object in the first frame with one of the objects in the second frame in this case.

In contrast to the object-based theory, motion energy models (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Werkhoven et al., 1993) and grouping theories (Alais & Lorenceau, 2002; He & Ooi, 1999; Kramer & Yantis, 1997) do not rely on attention. We nevertheless expected that attention should have a general effect on motion energy or grouping and thus also affect correspondence. In particular, attention studies have shown that the orienting of attention might affect the appearance of the attended stimulus by increasing its contrast (e.g., Carrasco, Ling, & Read, 2004; Carrasco, Penpeci-Talgar, & Eckstein, 2000; Posner, 1980) and its perceived duration (Rolke, Dinkelsbach, Hein, & Ulrich, 2008;

Rolke, Ulrich, & Bausenhardt, 2006; Yeshurun, 2004; Yeshurun & Levy, 2003). These attentional effects on the stimulus should affect the motion energy and grouping of the elements in the Ternus display in the same way in both displays, as these attention effects should be independent of the features of the elements. In particular, for grouping theories, if one of the elements is attended and thus appears to have a higher contrast than the other two elements, the spatial grouping strength between the elements should decrease, thereby increasing the amount of perceived element motion in all cue position and display conditions. For motion energy models, the temporal effect of attention should be most important. When attended, the elements in the first frame should be perceived as longer lasting, thus orienting attention to the second or third element of the first Ternus frame, should “close the temporal gap” between the two successive frames. Because this should decrease the perceived length of the ISI, we expected that group motion percepts should decrease. It is less clear what one would predict for the situation when attention is oriented to the first element, as motion energy theories are usually based on the central elements. We think, however, that in that case attention should rather increase group motion percepts, as there is no element at that location in the second frame, and thus the system should signal motion of the first element to the adjacent element.

To summarize our hypothesis, for motion energy and grouping theories we predict general attention effects independent of the particular features of the elements—that is, the same effects for the two display conditions and no interaction between the cue position (first, second, third, or all elements) and the display type (classic or competitive display). In contrast, for the object-based correspondence theory we expect such an interaction, as for the competitive display we predict more element motion percepts in the EB-match condition (attention oriented to the first element) compared with the GB-match condition (attention oriented to the second element), while for the classic display, we predict no particular effect of the attentional manipulation.

Experiment 1

Method

Participants A group of 14 participants (nine female) took part in the experiment. The sample size was chosen based on previous studies investigating correspondence, and in particular attentional effects, on correspondence (Hein & Moore, 2014; Kohler et al., 2008). Their ages ranged between 19 and 25 years ($M = 20.57$ years, $SD = 1.88$ years) and they were mostly students of the University of Tübingen. For their participation, they were compensated with money (8 € per hour) or course

credit. All of them were naïve as to the purpose of the experiment and reported normal or corrected-to-normal vision.

Apparatus The experiments were controlled by a PC with Windows XP as the operating system, on which a self-written program running in MATLAB (Version R2012a, 7.14, MathWorks Inc., MA, USA) using the Psychtoolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) and the EyeLink Toolbox (Cornelissen, Peters, & Palmer, 2002). A desk-mounted video-based eye tracker (EyeLink 1000 Desktop Mount, SR Research Ltd., Ontario, Canada) was used to monitor central fixation. Eye movements were measured monocular on the right eye with a sampling frequency of 500 Hz. Stimuli were presented on a 17-inch color cathode ray tube monitor ($1,024 \times 768$ pixels) with a refresh rate of 100 Hz. Participants conducted the experiment in a dimly lit individual testing room with a fixed viewing distance of 60 cm, and their heads were stabilized by a chin rest with forehead support.

Stimuli We use a modified version of a Ternus display (Pikler, 1917; Ternus, 1926) that consisted of two frames with three elements each with a diameter of 1.6° (see Fig. 2). The elements were aligned on a fictive circle with a diameter of 5.6° centered on the fixation point in the middle of the screen to ensure equal distances from the fixation point to each of the elements. Elements were separated by a center to center distance of 2° . The first frame was always presented horizontally centered around the fixation point. The second frame was

shifted by one element position to the left or to the right. In the classic Ternus display all elements had the same color (see Fig. 1a) and were cyan (RGB: 0, 142, 142; 15.2 cd/m^2), green (RGB: 136.5, 136.5, 4.5; 15.2 cd/m^2) or orange (RGB: 197, 107, 0; 15.2 cd/m^2), the color being randomly assigned across trials. In the competitive Ternus display, the three elements were presented in different colors in the following way: The first element in the first frame and the last element in the second frame were identical, the second elements in both frames were identical, and the last element in the first frame and the first element in the second frame were identical (see Fig. 1c). The same three colors as described above for the classic Ternus display were used (cyan, green, and orange), and which element was given which color was randomly assigned across trials. The background was presented in gray (RGB: 130.5, 130.5, 130.5; 14.7 cd/m^2) with a luminance set to be as equal as possible to the colors of the Ternus elements. The fixation point was black (0.07 cd/m^2) with a diameter of 0.59° and a smaller gray point in the center (0.15°) to facilitate precise fixations. As an attention cue the word “left,” “center,” “right,” or “all” (Arial font with type size 14) was presented centered 1.5° above the fixation point.

To test whether attention was successfully oriented, we replaced the Ternus task with a discrimination task in one third of the trials. For this discrimination task, a Landolt C with a diameter of 0.5° and a line width of 0.03° (one pixel) was presented centered on one of the elements in the first Ternus frame. The gap of the Landolt C pointed to the left or to the right with a fixed gap size. Gap size depended on the individual

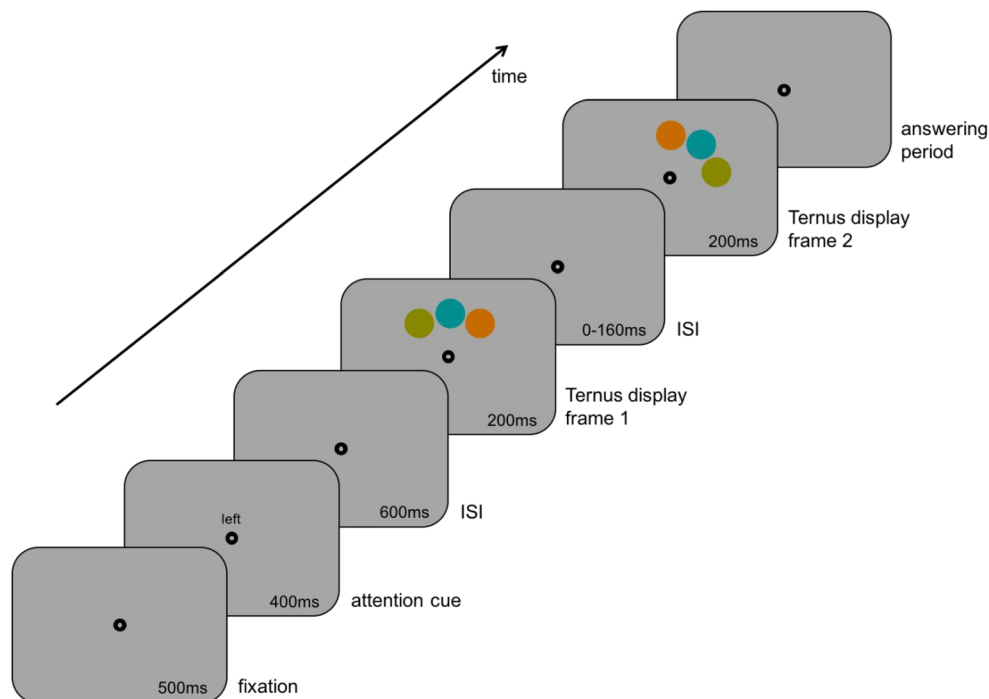


Fig. 2 Time course of a single Ternus task trial (shown here is a competitive Ternus display, with a cue to the left element, and a motion direction of the Ternus display to the right). (Color figure online)

performance of the participants in a pretest and ranged between 0.06° (two pixels) and 0.19° (six pixels). During the answering period, two Landolt Cs with a diameter of 1.6° and a gap size of 0.22° (seven pixels) were presented 3° to the left and right of fixation, one with a gap to the right and one with a gap to the left. Which Landolt C was presented on which side was randomly chosen. All Landolt Cs were black.

Task For the Ternus task, participants had to judge if they perceived all elements as moving together (group motion) or one element as moving separately, jumping across the other two elements (element motion) by pressing the “j” or “f” key, respectively. For the discrimination task, participants had to indicate as correctly as possible the side on which the Landolt C with the same orientation as the one they saw previously was presented, by pressing the “j” or “f” key for the right or left side, respectively.

Procedure Participants were informed about the experimental procedure and gave informed consent according to the ethical principles of the World Medical Association (2013; Declaration of Helsinki) prior to their participation. The experiment comprised two sessions of about 2 to 2.5 hrs, run on two different days. In the first session, a pretest was conducted prior to the main experiment in order to determine individual performance for the Landolt C discrimination. A “1 up 3 down” adaptive staircase (Kaernbach, 1991) was used to find the gap size, for which participants’ performance was about 75% correct in discriminating the Landolt C. This gap size was used for the first session and further adjusted for the second session, if the error rate in the first session was less than 10% or above 40%, by decreasing or increasing the gap size by one pixel. After this, each session began with written instructions and clearly distinguishable demonstrations of group and element motion (using extreme ISI of zero ms and 160 ms). Each session started with two practice trial blocks of 20 trials. Central fixation was not monitored during the first practice block to familiarize participants stepwise with the experimental procedure and the eye tracker. Participants completed 24 experimental blocks of 20 trials in each of the two sessions.

The time course of a Ternus task trial is displayed in Fig. 2. Each trial began with a fixation point. Participants were asked to fixate it and then confirm their fixation by pressing the “j” key. Following this confirmation, the fixation point was presented for another 500 ms, after which the cue (left, right, center, or all) was added to the display for 400 ms. Following the cue, the fixation point was presented alone again for another 600 ms. The first Ternus frame then added for 200 ms. After a variable ISI of 0–160 ms, during which again only the fixation point was presented, the second Ternus frame was added for another 200 ms. During the answering period only the fixation point was then presented until the participants gave their response. The next trial started after 500 ms.

The trial sequence for the discrimination task was similar to the Ternus task with the following exceptions: After the first Ternus frame was presented for 100 ms, a Landolt C with a gap either to the left or to the right side was added to the display at one of the Ternus elements for 40 ms. After the Landolt C disappeared, the Ternus frame was presented for another 60 ms. During the answering period, a screen with the two possible Landolt Cs and the fixation point was presented until a response was recorded.

Participants’ fixations on the fixation point were controlled from the beginning of each trial (key press by the participant) until the answering period started. To this end a fictive square (2.5°) that was centered around the fixation point was defined, within which participants had to fixate for the fixation to be accepted as valid. Between blocks and if necessary within a block, the eye tracker was calibrated with a five-point calibration. If fixation was lost, a written message reminded participants to fixate (presented for 1,500 ms). Trials, in which fixation was lost, were aborted and immediately repeated.

Design For the Ternus task a 2 (display type: competitive, classic) $\times 5$ (ISI: 0 ms, 20 ms, 40 ms, 80 ms, 160 ms) $\times 4$ (cue position: first, second, third, all) $\times 2$ (motion direction: right, left) within-subjects design was used. All factors were counterbalanced and randomly mixed within all trials. Each participant completed 640 Ternus task trials, resulting in eight observations for each condition. For the discrimination task, a 2 (display type: competitive, classic) $\times 4$ (cue position: first, second, third, all) $\times 2$ (target position: cued, noncued) within subject design was used. All factors were counterbalanced and randomly mixed within all trials. For the factor target position and all cue position conditions, but the “all” condition, in the cued condition the Landolt C was presented at one of the elements indicated by the cue (valid condition). In the noncued condition, the target was randomly presented at one of the two elements not indicated by the cue (invalid condition). If the cue oriented attention to all elements, then the target was randomly presented at one of them (neutral condition). Each participant completed 320 discrimination trials, resulting in 80 observations for the neutral condition and 120 observations each for the valid and invalid condition for both display types together.

Results

All statistical analysis were done with R (R Development Core Team, 2008). For analysis of variance (ANOVA), whenever necessary, Greenhouse–Geisser corrections were computed to account for violations of the sphericity assumption. For further post hoc analysis, Holm-corrected *t* tests were conducted. Prior to the inferential analysis trials in which other keys were used than one of the two possible response keys were excluded from the data, 20 trials in the Ternus task (<1%) and two trials in the discrimination task (<1%).

Ternus task For the analysis of the Ternus task, we first submitted the individual mean percentage of group motion responses to a three-factorial analysis of variance (ANOVA), with the factors ISI (0 ms, 10 ms, 20 ms, 40 ms, 80 ms, 160 ms), cue position (first, second, third, all) and display type (competitive, classic). In a next step, we looked at the competitive and the classic Ternus display separately, using a two-factorial analysis of variance (ANOVA) with the factors ISI (0 ms, 10 ms, 20 ms, 40 ms, 80 ms, 160 ms) and cue position (first, second, third, all). For the competitive Ternus display, if attention was directed towards the first element (i.e., the element containing the element bias), we will refer to it as the EB-match condition. If attention was directed to the second element (i.e., the element containing the group bias), we will refer to it as the GB-match condition. This distinction makes no sense for the classic Ternus display; as there are no biases in this display, we will therefore just refer to these conditions as attention being oriented to the first and second element. For both display types, if attention is directed to all elements, we will refer to it as the neutral condition, and if attention is directed to the third element, we will refer to it as the third element condition.

The omnibus analysis across all factors revealed a significant main effect of ISI, $F(4, 52) = 28.52, p < .001, \eta_p^2 = .69$ and of cue position, $F(3, 39) = 14.93, p < .001, \eta_p^2 = .53$, but no interaction between both factors, $F(12, 156) = 1.47, p = .142, \eta_p^2 = .10$. In addition, there was a trend for the display type, $F(1, 13) = 3.94, p = .069, \eta_p^2 = .23$, and there was an interaction between the factor display type and ISI, $F(4, 52) = 40.29, p < .001, \eta_p^2 = .76$. Most importantly, there was an interaction between display type and cue position, $F(3, 39) = 9.93, p < .001, \eta_p^2 = .43$. Finally, a trend for the three-way interaction, $F(12, 156) = 2.10, p = .063, \eta_p^2 = .14$, occurred.

To further investigate these interactions separate ANOVAs were conducted for the classic and the competitive Ternus display. Figure 3 shows the mean percentages of group motion responses as a function of the attention manipulation and the ISI for both display types separately. The analysis for the classic Ternus display revealed the typical main effect of ISI, $F(4, 52) = 42.91, p < .001, \eta_p^2 = .77$, with an increase of group motion percepts with increasing ISI. The main effect of the factor cue position was also significant, $F(3, 39) = 6.29, p = .001, \eta_p^2 = .33$, but the interaction between ISI and cue position was not significant, $F(12, 156) = 0.85, p = .536, \eta_p^2 = .06$. Post hoc tests for the cue position showed that in the neutral condition more group motion was perceived than when attention was oriented to the first element, $t(13) = 4.53, p_{\text{Holm}} = .003, d = 1.21$. There was also a trend that more group motion was perceived in the neutral compared with the third element condition, $t(13) = 2.99, p_{\text{Holm}} = .052, d = 0.80$. All other comparisons did not reach significance, $t_s \leq 2.26, p_{s;\text{Holm}} \geq .168, d_s \leq 0.60$.

In contrast to the classic Ternus display, the analysis of the competitive display revealed no significant effect for the factor ISI, $F(4, 52) = 2.09, p = .157, \eta_p^2 = .14$, as the overall percentage of group motion percepts were very similar for all ISI conditions (see Fig. 3, right graph). There was, however, a trend for an interaction between ISI and cue position, $F(12, 156) = 2.21, p = .052, \eta_p^2 = .15$. To further investigate this interaction, we conducted ANOVAs with the factor ISI for each cue position separately. This revealed a significant effect of the ISI for the GB-match condition, $F(4, 52) = 2.57, p = .049, \eta_p^2 = .17$, and the neutral condition, $F(4, 52) = 2.88, p = .031, \eta_p^2 = .18$. For the other cue positions no significant effects of the ISI were found, $F_s(4, 52) \leq 2.18, p_s \geq .134$. Most importantly, the analysis revealed a main effect of the factor cue position, $F(3, 39) = 15.52, p < .001, \eta_p^2 = .54$. Post hoc tests for this factor revealed significant differences between all conditions: As can be seen in Fig. 3 (right graph), the percentage of group motion percepts for the GB-match condition was higher than for all other conditions: neutral condition, $t(13) = 2.60, p_{\text{Holm}} = .045, d = 0.70$; third element, $t(13) = 3.41, p_{\text{Holm}} = .019, d = 0.91$; and most importantly EB-match condition, $t(13) = 4.82, p_{\text{Holm}} = .002, d = 1.29$. In addition, more group motion was perceived in the neutral condition compared with the third element condition, $t(13) = 2.81, p_{\text{Holm}} = .045, d = 0.75$, and the EB-match condition, $t(13) = 5.18, p_{\text{Holm}} = .001, d = 1.39$, as well as for the third element condition compared with the EB-match condition, $t(13) = 2.75, p_{\text{Holm}} = .045, d = 0.73$.

Discrimination task For the analysis of the discrimination task a two-factorial analysis of variance (ANOVA) with the factors display type (classic, competitive) and cueing condition (valid, invalid, neutral) was conducted for each of the two dependent variables, our main dependent variable error rates, but also on reaction times. We found no significant effects, neither for the mean error rates nor for the reaction times. In particular, for the error rates, the results showed no differences between the different cueing conditions (valid: 24.23%, invalid: 23.72%, neutral: 22.68%), $F(2, 26) = 0.72, p = .497, \eta_p^2 = .05$. There were also no other significant effects, $F_s \leq 1.38, p_s \geq .261$. For the reactions times the results also showed no differences between the different cueing conditions (valid: 1,003 ms; invalid: 1,019 ms; neutral: 1,014 ms), $F(2, 26) = 0.44, p = .652, \eta_p^2 = .03$, and no other significant effects, $F_s \leq 2.40, p_s \geq .145$.

Discussion

The pattern of results for the Ternus task differed depending on the display type. In particular, participants reported more group motion percepts in the GB-match condition than in the EB-match condition in the competitive display, but there was no difference between the comparable cueing conditions in

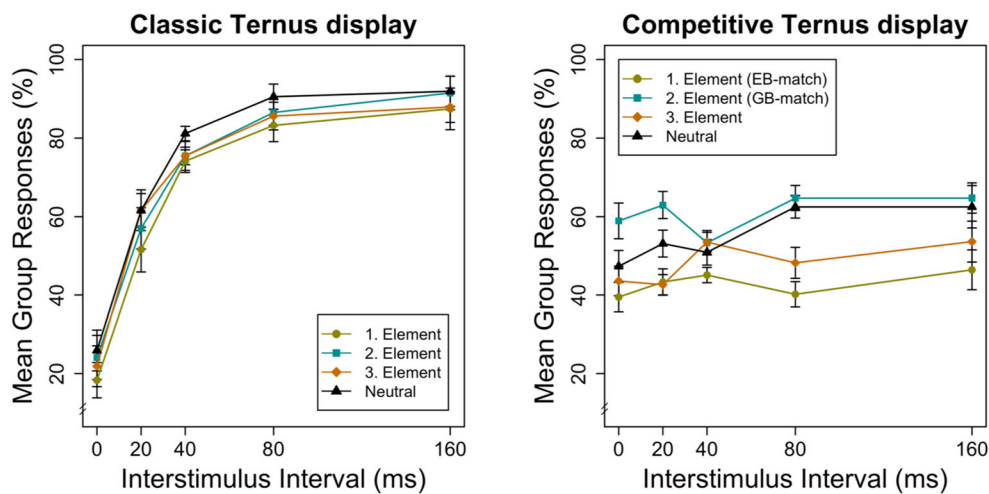


Fig. 3 Results of Experiment 1. Mean percentage of group motion responses as a function of ISI and cue position. The left graph shows the classic Ternus display and the right graph shows the competitive

Ternus display. The error bars represent the within-subject standard errors of the means in each condition (Cousineau, 2005; Morey, 2008). (Color figure online)

the classic display. The results therefore support the assumptions under the object-based correspondence theory. In addition to this difference between the GB-match and the EB-match condition, the specific pattern of attentional influences in the competitive Ternus display is also interesting. The results showed more group motion in the neutral compared with the EB-match condition and less group motion in the neutral compared with the GB-match condition. The difference between the neutral and the GB-match condition was, however, much smaller than the difference between the neutral and the EB-match condition. This could be due to more attentional resources being available for the whole Ternus frame in the neutral condition. Based on Aydın et al. (2011), more attention is needed to perceive group motion than to perceive element motion. Regarding the GB-match condition, this could have led to less group motion percepts due to less attentional resources available for the whole display, working against the bias toward more group motion. In contrast, regarding the EB-match condition, both effects should go in the same direction—that is, one would expect less group motion percepts due to less attentional resources, as well as due to the element bias. The finding that the neutral and the GB-match conditions were much more similar to each other than were the neutral and the EB-match conditions could also be explained within the framework of grouping theories (e.g., Kramer & Yantis, 1997), if one assumes that orienting attention toward an element changes its appearance—that is, making it more salient by increasing its contrast (e.g., Carrasco et al., 2004; Carrasco et al., 2000; Posner, 1980). This would reduce spatial grouping and therefore would decrease group motion percepts in the EB-match and the GB-match conditions compared with the neutral condition. The general difference between the GB-match and EB-match conditions, however, should remain constant, as the availability of less attentional resources or the reduced spatial grouping should have

led to the same decrease of group motion percepts in both conditions, which means that our main conclusion that attentional pointers seem to strongly influence the correspondence solution in the competitive display is supported by the specific pattern of results.

In addition, the results for the competitive display also showed more group motion in the neutral compared with the third element condition, the third element condition being more similar to the EB-match condition. As described above, directing attention to an individual element could have influenced the availability of general attentional resources or the spatial grouping, which could explain this effect at least to some degree. The third element condition, however, is also a special case concerning its feature, as the third element in the first frame is compatible with the first element in the second frame. Thus, if these elements are connected across frames via attentional pointers, this could be perceived as a special case of element motion, in which the center elements swap places, at least for some participants. Further studies have to investigate this possibility more closely.

Unexpected under the object-based correspondence theory, the analysis also revealed that attention had an effect in the classic Ternus display, as more group motion was perceived in the neutral condition, in which attention was directed to all elements, compared with all other conditions, although this difference only reached significance compared with when attention was directed to the first element. This attentional effect is in line with the study by Aydın et al. (2011), as if more attentional resources were available for the whole Ternus frames in the neutral condition when attention was directed to all elements, this could have led to more perceived group motion, than when attention was directed towards an individual element. This attentional effect in the classic display could

also be explained within the framework of grouping theories (e.g., Kramer & Yantis, 1997), as more group motion should be perceived in the neutral condition due to stronger spatial grouping compared with when attention is oriented to individual elements.

As a control for the allocation of attention on the Ternus element, participants had to perform a discrimination task at cued and noncued elements. We found, however, no effect of the attentional manipulation on discrimination performance. This was unexpected, as it has been shown that orienting attention to a target evokes faster responses and better performance in similar discrimination tasks (e.g., Posner, 1980; Posner, Snyder, & Davidson, 1980; Yeshurun & Carrasco, 1999). We assume that we did not obtain the expected cueing effect because the discrimination task was intermixed with the Ternus task. This dual task situation might have produced switch costs (e.g., Monsell, 2003), as participants might have focused on the main task—that is, the Ternus task that occurred in two thirds of all trials. This assumption is supported by the rather high RTs in the discrimination task. To test this idea that the intermixing of the discrimination task with the Ternus task prevented the attentional effect to be measurable in the discrimination task, we ran Experiment 2 and blocked the two tasks.

Experiment 2

In this experiment we used a blocked instead of a mixed design for the discrimination and the Ternus task. Moreover, to increase the importance of the cue, we made the cue predictive by presenting the target in 75% of the trials at the cued position (valid condition), instead of 50%, as in Experiment 1. Finally, we made the discrimination task easier and focused on reaction times as a measure of attentional allocation. In particular, we asked participants to detect a large cut-out on the top or the bottom of one of the Ternus elements, instead of a difficult Landolt C discrimination as in Experiment 1.

Method

Participants A new sample of 20 participants (13 female) contributed. We increased the sample size compared with Experiment 2 for the following reasons: First, we balanced the order of the tasks (Ternus and discrimination task) as well as the finger-to-key assignment for the RT-based discrimination task, which resulted in a multiple-of-four sample size number. Second, we expected that the block-wise separation of the Ternus task and the discrimination task might weaken the attentional effect in the Ternus task block, as directing attention was not necessary to solve this task. In increasing the sample size, we aimed to discover the potentially smaller effect. Participants'

ages ranged between 19 and 33 years ($M = 24.15$ years, $SD = 3.80$ years). Originally, 24 participants took part in this experiment. We excluded three participants from our analysis because they could not maintain fixation in more than 30% of the trials. One additional participant was excluded because this participant showed an atypical decrease of group motion responses with increasing ISI in the neutral condition of the classic Ternus display. This pattern is in the opposite direction of the typical increase of group motion with increasing ISI, suggesting that this participant might have mixed up the response keys.

Apparatus The apparatus was the same as in Experiment 1.

Stimuli The stimuli were the same as in Experiment 1 for the Ternus task. For the discrimination task we used a circular cut-out at the top or bottom of one of the Ternus elements. This cut-out was created by presenting a background-colored circle (diameter of 1.2°) on top, centered either at the top or at the bottom edge of the Ternus element.

Task The Ternus task was identical to that in Experiment 1. For the discrimination task, participants had to indicate as quickly and as correctly as possible with their index fingers, whether the cut-out in the Ternus element appeared at the top or the bottom, by pressing the “z” key for top (“z” on the German keyboard corresponds to “y” on the American keyboard) and the “b” key for bottom. The assignment of the fingers to the keys was counterbalanced across participants.

Procedure The general procedure was identical to Experiment 1, but with the following exceptions. First, no pretest for the discrimination task was necessary, as we used an easy cut-out discrimination. Second, the Ternus task and the discrimination task were run in two different sessions on two different days (order balanced across participants). In each session, participants completed 32 experimental blocks of 20 trials.

The time course of the Ternus task was identical to that of Experiment 1 (see Fig. 2). For the discrimination task, after the first Ternus frame was presented for 100 ms, the cut-out at one of the Ternus elements was presented for 100 ms, before the Ternus display disappeared. The next trial started 500 ms after a response was recorded.

Design The design for the Ternus task was exactly the same as for Experiment 1. For the discrimination task, the design was the same with the exception that the cue was now predictive, as the target was presented at the cued position in 75% of the trials (valid condition). Participants completed 640 discrimination task trials. This resulted in 160 observations for the neutral condition, 120 observations for the invalid condition and 360 trials for the valid condition for both display types together.

Results

Prior to the inferential analysis, we excluded trials in which other keys than one of the possible response keys were used. These were 71 trials in the Ternus task (<1%) and 44 trials in the discrimination task (<1%).

Ternus task As in Experiment 1, the omnibus analysis with the factors ISI, cue position, and display type revealed a significant main effect for ISI, $F(4, 76) = 23.53, p < .001, \eta_p^2 = .55$, and for cue position, $F(3, 57) = 9.16, p = .001, \eta_p^2 = .33$, but no interaction between both factors, $F(12, 228) = 0.83, p = .616, \eta_p^2 = .04$. In addition, and in contrast to Experiment 1, the factor display type was significant, $F(1, 19) = 9.18, p = .007, \eta_p^2 = .33$, with overall more group motion percepts in the classic ($M = 75.54\%$) compared with the competitive display ($M = 58.45\%$). As in Experiment 1, there was an interaction between the factor display type and ISI, $F(4, 76) = 33.06, p < .001, \eta_p^2 = .64$. Most importantly, and replicating Experiment 1, we found a significant interaction between display type and cue position, $F(3, 57) = 4.40, p = .019, \eta_p^2 = .19$. The three-way interaction between all three factors was also significant, $F(12, 228) = 1.83, p = .044, \eta_p^2 = .09$.

As in Experiment 1 we conducted separate ANOVAs for the classic and the competitive Ternus display to gain insights into the specific pattern of results for each display type. Figure 4 shows the mean percentages of group motion responses for the Ternus task as a function of the attention manipulation and the ISI for each display type condition. For the classic Ternus display there was an effect of ISI, $F(4, 76) = 35.22, p < .001, \eta_p^2 = .65$, with an increasing percentage of group motion percepts with increasing ISI. There was also a main effect of cue position, $F(3, 57) = 4.30, p = .019, \eta_p^2 = .18$. Descriptively the pattern of results was similar to those of

Experiment 1, as the largest difference in group motion percepts was between the neutral condition and the first element, followed by the third element condition and the second element condition. Holm-corrected post hoc tests, however, revealed no significant difference between any of the individual comparisons, $t_s \leq 2.54, p_{s;Holm} \geq .119, d_s \leq 0.56$. As in Experiment 1, the interaction of the factor ISI and cue position was not significant, $F(12, 228) = 1.69, p = .132, \eta_p^2 = .08$.

For the competitive Ternus display, consistent with Experiment 1, there was no main effect for the factor ISI, $F(4, 76) = 0.61, p = .559, \eta_p^2 = .03$. In contrast to Experiment 1, there was no trend for an interaction between the factor ISI and cue position, $F(12, 228) = 1.06, p = .393, \eta_p^2 = .05$. Most importantly, as in Experiment 1, there was a main effect of the factor cue position, $F(3, 57) = 8.52, p = .002, \eta_p^2 = .31$ (see Fig. 4, right graph). Holm-corrected post hoc tests for this factor revealed the following differences: Most importantly, group motion percepts in the GB-match condition were higher than in the EB-match condition, $t(19) = 2.73, p_{Holm} = .040, d = 0.61$. In addition, the group motion percepts were higher in the GB-match compared with the third element condition, $t(19) = 3.29, p_{Holm} = .019, d = 0.74$, and higher for the neutral compared with the EB-match condition, $t(19) = 3.16, p_{Holm} = .020, d = 0.71$, and the third element condition, $t(19) = 3.77, p_{Holm} = .008, d = 0.84$. In contrast to Experiment 1, the GB-match condition did not differ from the neutral condition, $t(19) = 0.76, p_{Holm} = .919, d = 0.17$, and there was also no difference between the EB-match and the third element condition, $t(19) = 0.48, p_{Holm} = .919, d = 0.11$.

Discrimination task We excluded trials in which participants made an error (6.06%). We additionally excluded trials with RTs ± 3 * standard deviations of the mean RT for each participant and condition (1.29%). In contrast to Experiment

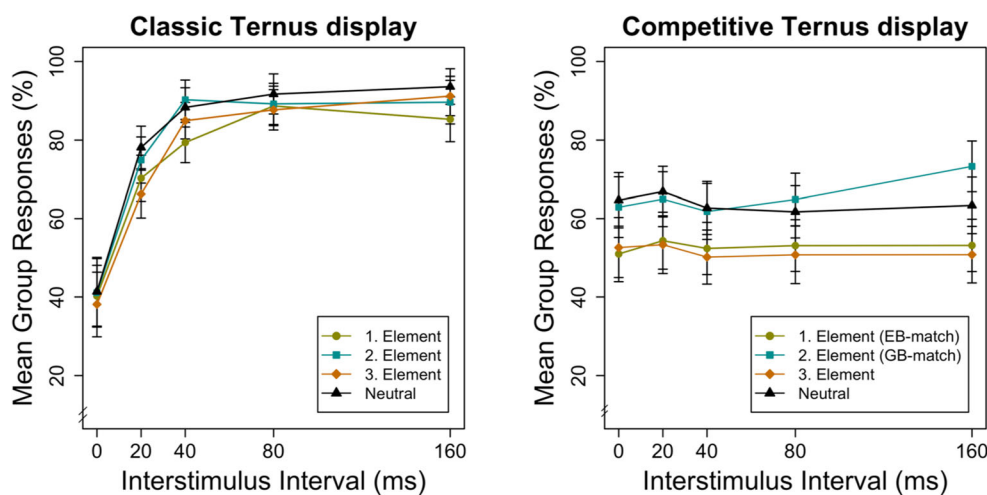


Fig. 4 Results of Experiment 2. Mean percentage of group motion responses as a function of ISI and cue position. The left graph shows the classic Ternus display and the right graph shows the competitive

Ternus display. The error bars represent the within-subject standard errors of the means in each condition (Cousineau, 2005; Morey, 2008). (Color figure online)

1, there was a significant main effect for the factor display type, $F(1, 19) = 20.89$, $p < .001$, $\eta_p^2 = .52$ and, most importantly, a significant effect of the factor cueing condition (valid: 457 ms, invalid: 496 ms, neutral: 465 ms), $F(2, 38) = 13.29$, $p = .001$, $\eta_p^2 = .41$. Post hoc tests revealed that participants reacted significantly faster in the valid compared with the invalid cueing condition, $t(19) = 3.71$, $p_{\text{Holm}} = .003$, $d = 0.83$, and in the neutral compared with the invalid cueing condition, $t(13) = 3.93$, $p_{\text{Holm}} = .003$, $d = 0.88$. There was a trend for a difference in RT between the valid compared with the neutral cueing condition, $t(19) = 1.81$, $p_{\text{Holm}} = .085$, $d = 0.41$. The interaction between the factors cueing condition and display type was not significant, $F(2, 38) = 1.07$, $p = .354$, $\eta_p^2 = .05$. The analysis for the mean error rates revealed no effect for the factor cueing condition (valid: 6.05%, invalid: 5.81%, neutral: 6.35%), $F(2, 38) = 0.58$, $p = .566$, $\eta_p^2 = .03$. There was also no difference in error rates between the two display types, $F(1, 19) = 0.13$, $p = .725$, $\eta_p^2 = .01$, but a trend for an interaction between display type and cueing condition, $F(2, 38) = 2.87$, $p = .069$, $\eta_p^2 = .13$.

Discussion

In this Experiment there was a cueing effect in the RT-based discrimination task. This result shows that the cue was in principle able to orient attention toward specific elements of the Ternus display. Why did we find the expected attentional effect in this experiment, but not in Experiment 1? We think that several factors might have contributed to the occurrence of the cueing effect. First, we strengthened the impact of the cue by enhancing its predictive value; second, we measured RT in a simple discrimination task instead of error rates; and third, we separated the discrimination task from the Ternus task. This latter change might have reduced switching costs (e.g., Monsell, 2003) between the Ternus task and the discrimination task, which might have masked attentional effects in Experiment 1. No matter which change might be the most important factor for establishing the cueing effect, we successfully showed that the cue employed in the two experiments had the potential to orient attention in the Ternus display.

Importantly, in this experiment, we replicated our most interesting result for the Ternus task—that is, the differential effects of attention in the competitive compared with the classic Ternus display. In particular, there were only small influences of attention in the classic Ternus display, but much larger attention effects in the competitive display. The reason why the attentional influence in the classic display and partly in the competitive display condition were slightly reduced in Experiment 2 might be due to the blocked design employed in Experiment 2. Here, an attentional orientation was task irrelevant for the Ternus task session and thus participants might have sometimes neglected the cueing instruction in this session. Overall, however, the most important effect of the

attentional orientation—that is, that in the competitive display more group motion percepts were reported for the GB-match condition compared with the EB-match condition, was replicated and is in line with the object-based correspondence theory (Hein & Cavanagh, 2012; Hein & Moore, 2014).

General discussion

The object-based correspondence theory (Hein & Cavanagh, 2012; Hein & Moore, 2014) suggests that correspondence is established by a one-to-one mapping of the elements that are perceived as most similar across both frames and that attention could mediate this process. In order to test this theory, we investigated if directing attention voluntarily to a specific object influences the correspondence solution. We used the Ternus display (Pikler, 1917; Ternus, 1926), an ambiguous apparent motion display, for which either element or group motion can be perceived depending on how correspondence is solved, and which has been shown to be strongly influenced by feature information (Casco, 1990; Hein & Moore, 2012; Kramer & Yantis, 1997; Petersik & Rice, 2008). In particular, we created a competitive Ternus display containing a bias toward element motion and a bias toward group motion within the same display by changing the color of the elements (Hein & Schütz, 2019). We also used a classic Ternus display, in which all elements had the same color. Attention was either directed to one individual or to all Ternus elements in both display conditions. Based on the object-based correspondence theory, we expected that attending an individual element would increase the impact of that element for solving correspondence in the competitive Ternus display, but not in the classic Ternus display. For the competitive Ternus display this should lead to more group motion percepts, if the element containing the group bias was attended (GB-match condition) compared with less group motion percepts, if the element containing the element bias was attended (EB-match condition). In the classic display condition, however, no effect of attention was expected under the object-based correspondence theory, as all elements had the same color and thus a similarity based one-to-one mapping across frames would not find a specific match. Therefore, directing attention to a particular element should not affect the correspondence solution. The results were in line with the predictions under the object-based correspondence theory (Hein & Cavanagh, 2012; Hein & Moore, 2014), as we found different effects of the attentional manipulation in the two display conditions in both experiments. In particular, across all ISI, more group motion was perceived in the GB-match compared with the EB-match condition in the competitive display, while no difference in group motion responses was found, when attention was oriented to the first and the second Ternus element in the classic display. This suggests that the attended element was weighted stronger

for solving correspondence that is, the corresponding one-to-one mapping across frames was more likely to be selected.

Interestingly, we also found different effects of the ISI in the two display types for the Ternus task. For the classic Ternus display we found the typical ISI effect (e.g., Pantle & Petersik, 1980; Petersik & Pantle, 1979), suggesting that spatiotemporal factors had a strong effect on correspondence in this case (Adelson & Bergen, 1985; Flombaum et al., 2012; Kahneman et al., 1992; van Santen & Sperling, 1985; Werkhoven et al., 1993). In the competitive Ternus display, however, the motion percept was nearly independent of the ISI. The competitive feature biases in this display seem to mostly override the ISI effect, which provides further evidence that under the right circumstances features can have a strong effect on solving correspondence (e.g., Alais & Lorenceau, 2002; Casco, 1990; Dawson et al., 1994; Hein & Cavanagh, 2012; Hein & Moore, 2012; Kramer & Rudd, 1999; Kramer & Yantis, 1997; Moore & Enns, 2004; Petersik & Rice, 2008; Wallace & Scott-Samuel, 2007). In line with the object-based theory of correspondence (Hein & Cavanagh, 2012; Hein & Moore, 2014), the independence from the ISI in the competitive display could be explained in the way that the feature information (i.e., the color of the elements) was more dominant than spatiotemporal factors and therefore correspondence might have been established mainly on the one-to-one mapping of the elements in this display condition.

Moreover, these different ISI effects in the two display conditions were modulated by the attentional manipulation in different ways. While for the classic display there was a very strong increase of group motion percepts with increasing ISI for all cueing conditions, for the competitive display only in the neutral condition a reduced (Experiment 1) or even no (Experiment 2) influence of ISI was found. This finding suggests that the influence of attention on correspondence was rather minimal when no features differentiated the Ternus elements and correspondence was more likely to be mediated by motion energy (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Werkhoven et al., 1993) and/or temporal grouping (He & Ooi, 1999; Kramer & Yantis, 1997). When the features made the elements distinguishable, however, the effect of attention on correspondence became very strong, suggesting that correspondence was more likely to be mainly mediated by object-based mechanisms. Thus, overall, the effect of attention on correspondence seems to be dependent on the correspondence mechanism. And which correspondence mechanism(s) is/are at work seems to be dependent on the complexity of the display (i.e., in our case, whether the elements are distinguishable by different features or not).

With the discrimination task we used in Experiment 2, we showed that attention was oriented successfully, which is in contrast to the lack of a cueing effect in Experiment 1. However, as the cue was the same in both Experiments and

as the main results in the Ternus task, especially, the difference between the EB-match and the GB-match condition in the competitive Ternus display were the same in both Experiments, one can assume that the attentional manipulation was also successful in Experiment 1. We hypothesize that we could not measure an attentional effect in the discrimination task because of switch costs between the Ternus task and the discrimination task, concealing the attention effects in this experiment. In contrast to the failure to measure an attentional effect in the discrimination task, the attentional manipulation seems to have even been stronger in Experiment 1 than in Experiment 2 for two reasons: First, concerning the effects in the competitive display in Experiment 1, the difference between the EB-match and the GB-match was larger and there were more modulations of the group motion percepts in the other cueing conditions compared with Experiment 2. Second, in the classic Ternus display there was a general attentional effect in Experiment 1, which can be explained by grouping theories (e.g., Kramer & Yantis, 1997) or effects of the availability of attentional resources (Aydın et al., 2011) that was smaller in Experiment 2. It is possible that due to the blocked design in Experiment 2 attention was oriented less strongly in the Ternus task, in which the attentional orienting did not help to do the task, reducing the strength of the attentional effects in the competitive display, and even eliminating some of the smaller, more general attentional effects in both display conditions.

To summarize, we found that spatial attention influences how feature biases are weighted to determine correspondence. Up to now studies have mainly shown that attention can influence correspondence if participants voluntarily envision a particular motion percept or voluntarily track a certain motion path (Kohler et al., 2008; Suzuki & Peterson, 2000; Wertheimer, 1912; Xu et al., 2013). Our results extend these findings by showing that voluntarily attending to a certain object also influences how correspondence is determined. Moreover, we found small general attention effects, on the motion percept in the Ternus display, especially in Experiment 1, which were present in both display types and could be explained by grouping theories (e.g., Kramer & Yantis, 1997) or a general effect of the availability of attentional resources (Aydın et al., 2011). Finally, our findings of an increase of the motion percepts in the direction of the bias for the competitive Ternus display in both Experiments support the object-based theory of correspondence (Hein & Cavanagh, 2012; Hein & Moore, 2014), which suggests that correspondence in this display condition is based on the perceived similarity of the individual objects in a one-to-one mapping, likely connecting these objects across space and time via attentional pointers (Cavanagh, 1992; Cavanagh, Hunt, Afraz, & Rolfs, 2010). Taken together, correspondence seems to be a complex process that can happen at different levels of processing depending on the specific requirements and the complexity of

the particular display the visual system has to interpret. Moreover, the effect of orienting attention toward individual elements seems to be dependent on the type of correspondence mechanism that is engaged.

Acknowledgements We thank Michael Hoyer, Maren Klimm, Marius Mackamul, and Carlos Medina for help with data collection. We also would like to thank the editor Michael D. Dodd and three anonymous reviewers for their very constructive feedback helping to clarify our ideas and suggesting the changes for the discrimination task in the second experiment. This research was supported by DFG project HE 7543/1-1.

Open practices statement None of the data or materials for the experiments reported here is available online, but can be obtained upon request. None of the experiments was preregistered.

References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284–299. <https://doi.org/10.1364/JOSAA.2.000284>
- Alais, D., & Lorenceau, J. (2002). Perceptual grouping in the Ternus display: Evidence for an “association field” in apparent motion. *Vision Research*, 42(8), 1005–1016. [https://doi.org/10.1016/S0042-6989\(02\)00021-4](https://doi.org/10.1016/S0042-6989(02)00021-4)
- Aydin, M., Herzog, M. H., & Ögmen, H. (2011). Attention modulates spatio-temporal grouping. *Vision Research*, 51(4), 435–446. <https://doi.org/10.1016/j.visres.2010.12.013>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. <https://doi.org/10.1163/156856897X00357>
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7, 308–313. <https://doi.org/10.1038/nn1194>
- Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research*, 40, 1203–1215. [https://doi.org/10.1016/S0042-6989\(00\)00024-9](https://doi.org/10.1016/S0042-6989(00)00024-9)
- Casco, C. (1990). The relationship between visual persistence and event perception in bistable motion display. *Perception*, 19(4), 437–445. <https://doi.org/10.1068/p190437>
- Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257(5076), 1563–1565. <https://doi.org/10.1126/science.1523411>
- 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. <https://doi.org/10.1016/j.tics.2010.01.007>
- Chen, L., & Zhou, X. (2011). Visual apparent motion can be modulated by task-irrelevant lexical information. *Attention, Perception, & Psychophysics*, 73(4), 1010–1015. <https://doi.org/10.3758/s13414-010-0083-5>
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, 34(4), 613–617. <https://doi.org/10.3758/BF03195489>
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45. <https://doi.org/10.20982/tqmp.01.1.p042>
- Dawson, M. R., Nevin-Meadows, N., & Wright, R. D. (1994). Polarity matching in the Ternus configuration. *Vision Research*, 34(24), 3347–3359. [https://doi.org/10.1016/0042-6989\(94\)90069-8](https://doi.org/10.1016/0042-6989(94)90069-8)
- Flombaum, J. I., Scholl, B. J., & Santos, L. R. (2012). Spatiotemporal priority as a fundamental principle of object persistence. In B. Hood & L. Santos (Eds.), *The origins of object knowledge* (pp. 135–164). Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199216895.003.0006>
- He, Z. J., & Nakayama, K. (1994). Perceived surface shape not features determines correspondence strength in apparent motion. *Vision Research*, 34(16), 2125–2135. [https://doi.org/10.1016/0042-6989\(94\)90322-0](https://doi.org/10.1016/0042-6989(94)90322-0)
- He, Z. J., & Ooi, T. L. (1999). Perceptual organization of apparent motion in the Ternus display. *Perception*, 28(7), 877–892. <https://doi.org/10.1068/p2941>
- Hein, E., & Cavanagh, P. (2012). Motion correspondence in the Ternus display shows feature bias in spatiotopic coordinates. *Journal of Vision*, 12(7), 1–14. <https://doi.org/10.1167/12.7.16>
- Hein, E., & Moore, C. M. (2012). Spatio-temporal priority revisited: The role of feature identity and similarity for object correspondence in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 975–988. <https://doi.org/10.1037/a0028197>
- Hein, E., & Moore, C. M. (2014). Evidence for scene-based motion correspondence. *Attention, Perception, & Psychophysics*, 76, 793–804. <https://doi.org/10.3758/s13414-013-0616-9>
- Hein, E., & Schütz, A. C. (2019). Competition between color and luminance in motion correspondence. *Attention, Perception, & Psychophysics*, 81(1), 310–322. <https://doi.org/10.3758/s13414-018-1589-5>
- Hock, H. S., Kelso, S. J. A., & Schöner, G. (1993). Bistability and hysteresis in the organization of apparent motion patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 19(1), 63–80. <https://doi.org/10.1037/0096-1523.19.1.63>
- Hsu, P., Taylor, J. E. T., & Pratt, J. (2015). Frogs jump forward: Semantic knowledge influences the perception of element motion in the Ternus display. *Perception*, 44(7), 779–789. <https://doi.org/10.1177/0301006615596903>
- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics*, 49(3), 227–229. <https://doi.org/10.3758/BF03214307>
- Kaheman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219. [https://doi.org/10.1016/0010-0285\(92\)90007-O](https://doi.org/10.1016/0010-0285(92)90007-O)
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What’s new in Psychtoolbox-3? *Perception*, 36 ECVF Abstract Supplement.
- Kohler, A., Haddad, L., Singer, W., & Muckli, L. (2008). Deciding what to see: The role of intention and attention in the perception of apparent motion. *Vision Research*, 48(8), 1096–1106. <https://doi.org/10.1016/j.visres.2007.11.020>
- Korte, A. (1915). Kinematoskopische Untersuchungen [Kinematoscopic examinations]. *Zeitschrift für Psychologie*, 72, 194–296.
- Kramer, P., & Rudd, M. (1999). Visible persistence and form correspondence in Ternus apparent motion. *Perception & Psychophysics*, 61(5), 952–962. <https://doi.org/10.3758/BF03206909>
- Kramer, P., & Yantis, S. (1997). Perceptual grouping in space and time: Evidence from the Ternus display. *Perception & Psychophysics*, 59(1), 87–99. <https://doi.org/10.3758/BF03206851>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Science*, 7(3), 134–140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychological Science*, 15(12), 866–871. <https://doi.org/10.1111/j.0956-7976.2004.00768.x>
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau. *Tutorial in Quantitative Methods for Psychology*, 4(2), 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>
- Pantle, A. J., & Petersik, T. J. (1980). Effects of spatial parameters on the perceptual organization of a bistable motion display. *Perception & Psychophysics*, 27(4), 307–312. <https://doi.org/10.3758/BF03206119>

- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. <https://doi.org/10.1163/156856897X00366>
- Petersik, T. J., & Pantle, A. (1979). Factors controlling the competing sensations produced by a bistable stroboscopic motion display. *Vision Research*, 19(2), 143–154. [https://doi.org/10.1016/0042-6989\(79\)90044-0](https://doi.org/10.1016/0042-6989(79)90044-0)
- Petersik, T. J., & Rice, C. M. (2008). Spatial correspondence and relation correspondence: Grouping factors that influence perception of the Ternus display. *Perception*, 37(5), 725–739. <https://doi.org/10.1068/p5900>
- Pikler, J. (1917). *Sinnesphysiologische Untersuchungen* [Sensoryphysiological examinations]. Leipzig, Germany: Barth.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25. <https://doi.org/10.1080/00335558008248231>
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109(2), 160–174. <https://doi.org/10.1037/0096-3445.109.2.160>
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org>
- Ramachandran, V. S., & Anstis, S. M. (1983). Perceptual organization in moving patterns. *Nature*. <https://doi.org/10.1038/304529a0>
- Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). Concord, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262518420.001.0001>
- Rolke, B., Dinkelbach, A., Hein, E., & Ulrich, R. (2008). Does attention impair temporal discrimination? Examining non-attentional accounts. *Psychological Research*, 72(1), 49–60. <https://doi.org/10.1007/s00426-006-0092-0>
- Rolke, B., Ulrich, R., & Bausenhart, K. M. (2006). Attention delays perceived stimulus offset. *Vision Research*, 46, 2926–2933. <https://doi.org/10.1016/j.visres.2006.02.022>
- Suzuki, S., & Peterson, M. A. (2000). Multiplicative effects of intention on the perception of bistable apparent motion. *Psychological Science*, 11(3), 202–209. <https://doi.org/10.1111/1467-9280.00242>
- Ternus, J. (1926). Experimentelle Untersuchungen über phänomenale Identität. *Psychologische Forschung*, 7, 81–136. Translated to English in W. D. Ellis (Ed.), 1950. *A sourcebook of Gestalt Psychology*. New York, NY: Humanities Press. <https://doi.org/10.1007/BF02424350>
- Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition*, 74, B27–B32. [https://doi.org/10.1016/S0010-0277\(99\)00065-7](https://doi.org/10.1016/S0010-0277(99)00065-7)
- Ullmann, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2(2), 300–321. <https://doi.org/10.1364/JOSAA.2.000300>
- von Schiller, P. (1933). Stroboskopische Alternativversuche [Stroboscopic alternative experiments]. *Psychologische Forschung*, 17, 179–214. <https://doi.org/10.1007/BF02411959>
- Wallace, J. M., & Scott-Samuel, N. E. (2007). Spatial versus temporal grouping in a modified Ternus display. *Vision Research*, 47(17), 2353–2366. <https://doi.org/10.1016/j.visres.2007.05.016>
- Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, 33(4), 463–485. [https://doi.org/10.1016/0042-6989\(93\)90253-S](https://doi.org/10.1016/0042-6989(93)90253-S)
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung [Experimental studies on the perception of motion]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 61(1), 161–265.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung*, 4(1), 301–350. <https://doi.org/10.1007/BF00410640>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. *The Journal of the American Medical Association*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>
- Xu, Y., Suzuki, S., & Franconeri, S. L. (2013). Shifting selection may control apparent motion. *Psychological Science*, 24(7), 1368–1370. <https://doi.org/10.1177/0956797612471685>
- Yeshurun, Y. (2004). Isoluminant stimuli and red background attenuate the effects of transient spatial attention on temporal resolution. *Vision Research*, 44, 1375–1387. <https://doi.org/10.1016/j.visres.2003.12.016>
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39(2), 293–306. [https://doi.org/10.1016/S0042-6989\(98\)00114-X](https://doi.org/10.1016/S0042-6989(98)00114-X)
- Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science*, 14(3), 225–231. <https://doi.org/10.1111/1467-9280.02436>
- Yu, K. (2000). Can semantic knowledge influence motion correspondence? *Perception*, 29(6), 693–707. <https://doi.org/10.1068/p3063>

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