

## Chapter 6

### Shape-Independent Processing of Biological Motion

Nikolaus F. Troje and

Dorita H. F. Chang

The chapter was published in:

Johnson, K. and Shiffrar, M. (2013) People Watching: Social, Perceptual, and Neurophysiological Studies of Body Perception. Oxford University Press.

Since Johansson (1973) first demonstrated that several light patches placed to represent the motion of the joints of an agent are sufficient to elicit the percept of a walking figure, many efforts have been made to demonstrate the rich amount of information that can be extracted from biological motion patterns. Biological motion, as represented by a set of moving dots, can not only readily invoke the percept of an animate entity, but can also enable the retrieval of characteristics such as the agent's gender (Barclay, Cutting, & Kozlowski, 1978; Kozlowski & Cutting, 1977), emotion (Dittrich, Troscianko, Lea, & Morgan, 1996; Pollick, Paterson, Bruderlin, & Sanford, 2001), and identity (Cutting & Kozlowski, 1977). The ability of the visual system to access such information is not restricted to human patterns but extends to the perception of point-light animations depicting other legged animals (Mather & West, 1993). What is the nature of the mechanisms that allow the retrieval of such a large wealth of information from biological motion?

Motion in point-light stimuli can play a number of different roles. A role that has been discussed extensively in the past is how motion can help to recover the global, articulated structure of a moving body. Some debate is currently under way in the literature about whether recovery of motion-mediated structure requires local motion signals at all, or whether this is done by tracking a deformable template (e.g., Beintema & Lappe, 2002; Neri, Morrone, & Burr, 1998, see also Fleischer & Giese, Chapter 22, this volume). Depending on which view is assumed, motion is directly involved in recovering partly rigid body segments and their articulation (Aggarwal & Cai, 1999), or it is only required to interpret the particular deformations of a previously learned template that already implements the articulated structure of the body

(Lange & Lappe, 2006). However, in both cases, the goal is to explain how the visual system recovers the global, articulated shape of the body and to which extent the local motion of body parts contributes to that process. The very interesting discussion around this topic is covered in more detail in Chapter 18, by Thompson, this volume.

Here, we are concerned with a different aspect of local biological motion. Biological motion seems to contain information beyond recovering the articulated structure of a body in action. As we will show in this chapter, it turns out that signatures imprinted on the movements of an animate agent by the physical laws related to inertia and gravity are used to label motion as biological and to derive information about the facing direction of an agent. These signatures are inherent in the local motion of individual body parts, and they are therefore entirely independent of the body's shape.

The importance of signatures in the local motion of individual dots first became obvious from studies investigating the effect of different masks on detection and recognition of point-light walkers (Cutting, Moore, & Morrison, 1988). Masks that consist of dots that mimic the local motion of the dots constituting the target point-light walker are much more efficient than any other kinds of masks. The difference between such a “scrambled walker mask” and other dynamic dot masks is that the former deprives the observer from the use of local motion to segregate the target walker from the background mask. The relative efficiency of scrambled masks therefore argues for an important role of cues contained in the local motion of individual dots (Bertenthal & Pinto, 1994; Hiris, Humphrey, & Stout, 2005).

A role of local motion independent of the recovery of motion-mediated structure is also evident from two different experiments reported by Mather, Radford, and West (1992). In the first experiment, these authors manipulated their display's spatial and temporal properties by

changing interframe intervals and frame increments and asked observers to complete coherence and direction discrimination tasks. They found that performances on their tasks were reliable only when the interframe interval and frame increments were short—consistent with the restricted temporal and spatial properties of local motion detection processes (e.g., Baker & Braddick, 1985). In a second experiment, Mather et al. (1992) measured effects of omitting individual dots of the point-light walker. The motivation of this experiment was to distinguish between two different models that explain how articulated structure can be retrieved from point-light displays. Models assuming a hierarchical encoding, in which the motion of any segment on the body is represented with respect to its parent segment (e.g., Cutting, 1981), would predict that omission of the most proximal dots (shoulders, hips) would be most detrimental. On the other hand, models that assume that individual segments would be identified by means of two or more dots that are rigidly connected (e.g., Webb & Aggarwal, 1982) and then composed into an articulated figure would predict that omission of dots representing more than one segment (e.g., dots representing elbows or knees) would result in the largest effects. The results contradicted both predictions. Performance in Mather et al.'s (1992) experiments remained good in all conditions except when wrist and ankle points were omitted. The authors point out that this finding speaks to a role of the local motion of the distal parts of the limbs. As wrist and ankle were omitted together in these experiments, they were not able to further specify whether the crucial motion is the one of the wrists (or hands), the motion of the ankles (or feet), or both. However, these experiments support the idea that at least some aspects of biological motion perception are controlled by the local motion of distal body segments, and that they are thus independent of the specific shape of the actor. Therefore, they would apply to the perception of humans and animals alike.

In our own work (Troje, 2008), we have suggested a framework that proposes a processing level associated with the detection of another animal in the visual environment that is distinct from a second level responsible for recovering its specific nature and actions. These two levels, in our view, are closely linked to the distinction between the roles of local motion and global structure, respectively. Our suggestion was initially motivated by findings of two distinct inversion effects for biological motion perception (Troje & Westhoff, 2006). In this chapter, we will summarize these first experiments. We will then discuss a number of predictions they make about the role of a proposed visual filter that responds to invariants of the local motion of the feet. We will eventually review more recent studies—both from our own lab and from other research groups—that test these predictions and that further characterize the particular properties of the proposed visual filter.

## Two Inversion Effects and a “Local” Visual Filter

Inversion effects, that is, a decrease in recognition performance if a stimulus is turned upside-down, have been described for several different object classes. They are particularly well studied in the context of face recognition. A large body of evidence demonstrates that inversion effects in faces can be attributed to the processing of configural information, that is, to the processing of the spatial layout of facial features rather than to the characterization of the features themselves (Diamond & Carey, 1986; Leder, Candrian, Huber, & Bruce, 2001; Maurer, Grand, & Mondloch, 2002). Inversion effects have also been reported in the context of biological motion perception. Performances on behavioral tasks are generally impaired if point-light displays are presented upside-down. For example, observers report that the perceived quality or naturalness of movement is degraded with inversion (Sumi, 1984). Accuracies for the detection of a coherent point-light target in a mask of scrambled walkers drop to chance level when the target is inverted

(Bertenthal & Pinto, 1994). Observers' abilities to judge human action type (Dittrich, 1993) and the agent's gender (Barclay et al., 1978) are also disrupted upon inverting veridically organized point-light displays. Drawing on analogies between biological motion perception and face perception, inversion-related impairments in perceiving biological motion have often been attributed to impaired processing of the familiar configural shape of the walker as conveyed by the display's overall global spatio-temporal organization (Bertenthal & Pinto, 1994).

Now, this view seems to contradict observations we made a few years ago in our lab (Troje & Westhoff, 2006). There, we demonstrated a pronounced inversion effect for biological motion stimuli that were entirely devoid of any spatial or temporal structure between individual dots. In the study, observers were presented with point-light displays that were scrambled such that the individual dot trajectories were displaced randomly in space, thereby removing global spatial structure. Noise was also added to both the phase and the frequency of the individual dots, thus disrupting their temporal structure, too. The displays were presented at both upright and inverted orientations, and observers were asked to discriminate the walker's facing direction. The first interesting result of this study was that observers were very well able to discriminate walking direction from scrambled displays (see Saunders, Suchan, & Troje, 2009, for some thoughts on why this wasn't described earlier). Moreover, we showed that inversion impaired performance for these scrambled displays almost to the same degree as for unscrambled displays. A second experiment revealed that the cues to direction and the associated inversion effect for scrambled displays were carried entirely by the dots representing the feet (Figure 6-1)—a finding that corroborates the findings of Mather et al. (1992) described above.

Rather than arguing against a role of configural processing, the data suggest the existence of two entirely different sources for the inversion effect observed with biological motion stimuli.

One of them may very well be related to the inversion effect observed with faces and some other object classes: Diagnostic features inherent to the subtle changes in the configuration of a structured visual stimulus can be derived better when the stimulus is seen in a familiar orientation. The second cause, however, has nothing to do with configural organization. We suggested (Troje & Westhoff, 2006) that it is due instead to our visual system's ability to exploit inherent knowledge about the dynamics of physical events in the visual world and particularly about effects of gravity and inertial forces on the biomechanics of animate motion. Inversion effects would then result from the orientation specificity of gravitational acceleration. In fact, it is known from work on the visual perception of both inanimate and biological events that the visual system makes very specific assumptions about the gravity-defined relation between spatial and temporal scales and that these assumptions are used to derive one from the other (for a more detailed discussion of this literature, see below).

We (Troje & Westhoff, 2006) therefore hypothesized that the visual system also exploits its sensitivity to gravity-defined dynamics to detect the gravity-dependent movements typical for the limbs of terrestrial animals in locomotion. Unless an animal is actively disguising itself (as may be the case in a stalking predator), locomotion is optimized for energy efficiency (Alexander, 2003), thus resulting in invariants such as the Froude number (Alexander, 1989), which relates stride frequency and leg length for animal locomotion in a similar way as the length and the period of a swinging pendulum are related under normal constant gravity conditions.

The rhythmic, either ballistic or pendular motion typical for the limbs of animals is very unlikely to be produced by inanimate objects. The inversion effect observed for perceiving the direction of scrambled biological motion might thus reflect a mechanism that takes advantage of

the characteristics of this motion to detect the presence of terrestrial articulated animals independently of their particular shape and nature. The underlying visual filter would be orientation-dependent due to the directional nature of gravity. The upright trajectories of animal feet would trigger this mechanism and result in the feet becoming salient—which is not the case if they are presented upside-down. According to this view, direction-specific information required to solve the direction discrimination task is only revealed from the upright, gravity-consistent stimulus. The role of the proposed mechanism for direction discrimination as measured in Troje and Westhoff’s experiment, however, is only secondary. The primary role is its diagnostic value for the quick and efficient detection of potentially interesting or dangerous animals: Predators, prey, or conspecifics.

It has since been suggested (Johnson, 2006; Troje, 2008) that this mechanism is evolutionarily old, potentially innate in humans, should work well in the visual periphery, is preattentive in nature, and is implemented in older, probably subcortical parts of the human brain. In the remainder of this chapter, we will review and discuss evidence for these hypotheses based on data collected over the last few years both in our own lab, as well as by other research groups.

## Perception of Direction and Animacy

Among the propositions put forth by us and others (Johnson, 2006; Troje & Westhoff, 2006) was a suggestion that the local cues contained in biological motion not only serve the retrieval of directional information but may function as general signals to cue the presence of other animals in the visual environment. This suggestion predicts that the presence or absence of the relevant local cues should not only be reflected in the observer’s ability to perceive direction, but also influence the perceived animacy of these displays. Indeed, we have demonstrated that, like the

perception of direction, the perception of animacy from scrambled displays is also orientation-specific (Chang & Troje, 2008). In one experiment, we presented observers with coherently organized and scrambled versions of point-light displays depicting a variety of animals at upright and inverted orientations. Observers were asked to rate each display on a rating scale according to how alive the depicted entity appeared to be. Our findings are shown in Figure 6-2. Coherent walkers were generally perceived to be more animate than scrambled walkers. This is not surprising as the perception of animate entities surely must also rely on the presence of a recognizable form. Critically, however, the perceived animacy of the display decreased with inversion, not only for the coherent displays, but also for the scrambled displays. The orientation specificity of animacy perception from scrambled displays is thus congruent with the orientation dependency observed for discerning direction of motion from these same displays.

In a subsequent experiment, we compared direction discrimination and animacy perception more directly by showing the same subjects identical stimuli in two tasks: one requiring subjects to indicate perceived direction and one requiring them to rate perceived animacy of the display. Results revealed that direction discrimination accuracies of observers were directly correlated with animacy ratings. These latter findings suggest that the two paradigms are controlled by similar perceptual mechanisms. In the same series of experiments (Chang & Troje, 2008), we also varied the duration of the displays within the range of 200 ms to 1,000 ms. Interestingly, we found no effect of display time, neither on direction discrimination performances nor on the animacy ratings. For upright scrambled biological motion, a duration of 200 ms, which is just a fraction of a full gait cycle, seems to be more than enough to convey both direction and animacy. In contrast, for inverted scrambled biological motion, even display times of 1,000 ms are not sufficient to provide the visual system with the same information.

## The Role of Gravitational Acceleration

What are the characteristics of the underlying mechanism? We know that the cues to direction and the associated inversion effects are carried exclusively by the motion of the feet (Troje & Westhoff, 2006). We also know that only 200 ms of display time is enough (Chang & Troje, 2008). So, what is special about the foot motion? What carries the inversion effect? It must be something that imposes vertical asymmetry even on very small fragments of the trajectory of the moving feet.

Gravity is a likely candidate. Locomotion is optimized to minimize energy expenditure and is in large parts shaped by gravitational force -- with the work done by leg muscles reduced by a pendular transfer between potential and kinetic energy (Cavagna, Saibene, & Margaria, 1963; Cavagna, Heglund, & Taylor, 1977). The notion that observers may use inherent knowledge about the effects of gravity for visual perception has been implied in the past, both for the perception of inanimate dynamic events (e.g., Hecht, Kaiser, & Banks, 1996; McConnell, Muchisky, & Bingham, 1998; Pittenger, 1985; Stappers & Waller, 1993) and for the perception of biological motion (e.g., Jokisch & Troje, 2003; Runeson & Frykholm, 1981; Shipley, 2003). For example, Pittenger (1985) showed that observers could estimate the length of a swinging pendulum if information about its period is provided. In another study, Stappers and Waller (1993) presented observers with animations of a fountain on a video background and asked observers to adjust the magnitude of gravity to make the display appear natural. The authors found that observers were well able to match gravitational acceleration with apparent depth. In biological motion perception, Runeson and Frykholm (1981) found that observers were able to judge a box's mass from a point-light display of an actor lifting a box and reasoned that observers exploited the relationship between postural changes and the acceleration of the box

using dynamic principles. Jokisch and Troje (2003) provided further evidence to suggest that the visual system exploits a fixed relationship between spatial and temporal parameters that exists in a gravity-based environment by showing that observers, when presented with point-light displays of dogs walking with varying frequencies, judged the size of the dog to be smaller at high stride frequencies than at low frequencies. In another study, Shipley (2003) presented a point-light figure walking on his hands in upright and inverted orientations and found that walker detection was better for the upright display that contained familiar, gravity-consistent dynamics despite the unfamiliar orientation of the body. These studies suggest that observers use knowledge about the direction and effects of gravity when interpreting dynamic events, including biological motion.

The use of gravity-based interpretations of local biological motion implies that the critical cue contained in the foot motion lies with its vertical acceleration. A critical role for vertical acceleration is indeed supported by our recent empirical work (Chang & Troje, 2009a). Motivated by our earlier findings that walking direction can be discriminated from scrambled displays within exposure times corresponding to only a fraction of a full gait cycle (Chang & Troje, 2008), we first asked whether all fragments of the gait cycle contain equally salient directional information and contribute equally to the associated inversion effect. We created novel stimuli derived from single fragments of foot motions and asked observers to complete a direction discrimination task. We derived two different sets of displays. In one set, we presented multiple copies of the same fragment, while in the second set we combined copies of two fragments sampled 180 degrees in phase apart from each other in the gait cycle, thus representing the two feet moving in counterphase.

Figure 6-3 shows the foot motion of a right-facing human walker and illustrates the construction of our stimuli. Each display consisted either of several copies of a particular

fragment (Figure 6-3C) or of pairs of fragments (Figure 6-3D) of a specified temporal length specified from a foot of a walker in sagittal view. We found that these displays permit the recognition of the walker's facing direction and also carry an inversion effect—but only if derived from paired fragments (see also Casile & Giese, 2005; Fleischer & Giese, Chapter 22, this volume). More specifically, our results revealed that observers presented with stimuli derived from individual fragments of the gait cycle merely based direction judgments upon the horizontal displacement of the individual dot elements (Figure 6-4A). For example, when presented with a stimulus that contained only fragment 7, derived from a right-facing human walker foot as shown in Figure 6-3, observers invariantly reported leftward direction—reflecting the physical displacement of the single fragment. Averaged over all segments, and thus across a full gait cycle, the perceived displacements average out and no net displacement is reported. Accordingly, we found that average performance rates for these stimuli were at chance level for the single fragment stimuli.

For the paired fragment stimuli, the results were very different. We found discrimination performances for the paired stimuli to be above chance level and we also found that performance was subject to a strong inversion effect (Figure 6-4B). Critically, these results suggest that retrieving direction from scrambled biological motion displays involves more than processing the local spatio-temporal cues inherent in the isolated motion of the foot. Rather, the foot's elemental cues must be evaluated with reference to the motions of other elements, or at least to other parts of the same element's trajectory. Why is the additional reference point necessary?

A key to answering this question is the distinction between the direction as conveyed by the explicit translatory displacement of a walker, which we call *extrinsic* motion, and direction conveyed by implicit cues contained in the deformation of the body, which we call *intrinsic*

motion. This distinction can be understood by considering earlier demonstrations (Johansson, 1974; Rubin, 1927) on how the kinematics of the display of a rolling wheel are decomposed into extrinsic translatory motion and deviant circular or intrinsic motion. Presented alone, the point representing the rim of the wheel is perceived as moving along a cycloidal path. When combined with a point placed to represent the axle however, observers perceive a rolling wheel. This phenomenon is a result of the translatory motion of the axle point being used as a reference frame for the rotating components.

The point-light walkers we used in our experiments do not show any extrinsic motion. Like most other researchers in the field, we subtracted translatory displacement so that the walker appears to be walking stationary on a treadmill. The reason for doing this is to concentrate on the more interesting aspects: the information contained in the intrinsic motion of the walker. When presenting only isolated segments that do not reveal their origin as fragments of intrinsic walker motion, the dominant stimulus is most likely the explicit and coherent displacement of the individual segments. Their origin as fragments of intrinsic motion can only be revealed when a reference relative to which their motion can be assessed is provided. Only the addition of the second fragment in our fragmented foot stimuli allows a switch from the perception of solely extrinsic motion of the individual dots to the perception of intrinsic motion, which contains cues about the walker's facing direction.

Still, there remains a question as to what accounts for the inversion effect observed for perceiving local biological motion. Inversion affects only vertical components of the trajectories. Thus, inversion effects must be due to vertical asymmetries in the swing phase segments of the gait cycle (i.e., Figure 6-3B, fragments 1-4), rather than to the (vertically symmetric) stance phase segments. Asymmetries may be contained in the polarity of vertical velocity, vertical

acceleration, or even higher order derivatives. As noted earlier, there are many reasons to believe that the important cue lies with the vertical acceleration exhibited by the swing phase fragments. We therefore investigated the relationship between acceleration and the local inversion effect by correlating the presence of vertical acceleration in individual fragments to the degree of the experienced inversion effect. In order to increase the variance in the amount of acceleration, we extended the number of our fragmented foot displays by including stimulus samples from a running human and from foot motions of other animals (a cat and a pigeon) (Figure 6-5A). Some of the results from this experiment are shown in Figure 6-5B. Here, the inversion effect is quantified by a difference score obtained by computing the difference in discrimination accuracies between the upright and inverted versions for each stimulus. We then compared this inversion score for each stimulus to the absolute difference in vertical acceleration between its two constituent fragments. The figure shows that the size of the inversion effect was correlated with the amount of vertical acceleration contained in our stimuli (Chang & Troje, 2009a).

Even stronger evidence supporting a role of acceleration for the local inversion effect can be found in a final experiment in which we compared behavioral discrimination ability for stimuli derived from the naturally accelerating human walker foot and stimuli that were identical but had accelerations removed along the trajectory paths (i.e., the dots moved with constant speed along its trajectory). The data corresponding to this experiment are shown in Figure 6-6. Critically, we showed that the inversion effect disappeared for the constant-speed stimuli (Chang & Troje, 2009a).

Along with the finding that the inversion effect is exclusively carried by the motion of the feet (Troje & Westhoff, 2006, see also Figure 6-1), the data suggest that the local biological motion inversion effect is, in fact, carried by the vertical accelerations in the foot motion.

However, it remains unclear as to how acceleration contributes to achieving the facing direction of a display. Consider that foot motions carry asymmetries not only along the vertical dimension (as discussed above to be relevant to the inversion effect), but also along the horizontal dimension. Only horizontal asymmetries can contain information about the facing direction of the walker. Whatever this information is, it remains the same if the stimulus is inverted (i.e., flipped about the horizontal axis). What seems to change, however, is the visual system's ability to make use of it. We suggest that vertical acceleration plays a role in flagging a stimulus or event as being animate: Animate motion is characterized by a mass being propelled away from the ground and then pulled back by gravity. We suggest that only if an event is deemed to be animate does our visual system attribute a facing direction to it. Thus, vertical asymmetry carried by the acceleration may determine the validity of a trajectory for further analysis as an animate agent. Once the stimulus has been validated and categorized in that way, horizontal asymmetries are used to determine the facing direction of the walker. What the adequate stimulus feature really is remains to be investigated.

Thus far, we have shown that the same local cues that convey directional information from biological motion also drive the perceived animacy of the stimulus. Furthermore, we have made some progress toward identifying the specific cues that are critical—notably, the role of acceleration in the foot motion. Returning to the propositions first put forth by Troje and Westhoff (2006) and later by Troje (2008), we can now consider more generally the role of the local motion mechanisms as compared to those that mediate global structure-from-motion.

## Innate Components of Biological Motion Perception?

In light of the propositions set forth previously as to the roles of the different mechanisms for biological motion perception (Troje, 2008; Troje & Westhoff, 2006), we have proposed that the

global and local mechanisms in particular may act in an analogous manner as mechanisms for face recognition suggested by Morton and Johnson (1991), termed *conspec* and *conlern* (Chang & Troje, 2009b; Johnson, 2006). In brief, Morton and Johnson (1991) proposed that two systems underlie the development of face perception in infants. They suggested that newborn infants possess an innate mechanism, termed *conspec*, which is based on a very coarse, template-based, innate visual filter that responds to the general, average appearance of a face. The system is presumed to guide attention toward faces, thus ensuring that the developing visual system receives sufficient exposure to an important stimulus class and secures the intake of as many examples of it as possible in order to eventually learn how to retrieve all the sophisticated information they carry. This learning itself involves a second mechanism, termed *conlern*, and requires individual experience and maturation to extract information about identity, emotion, and other characteristics of human faces. We propose that local and global mechanisms for biological motion perception act in an analogous manner. According to this view, the local motion signals characterized above drive an early mechanism that may be evolutionarily old and potentially innate. This mechanism may serve as a general detection system that directs attention to an animate agent in the visual environment. This agent can then be further investigated via a second mechanism, which in turn retrieves the dynamically changing shape of the articulated body of the agent and from it derives its identity, actions, and intentions.

Several lines of research appear to be compatible with the existence of an early, potentially innate mechanism for biological motion perception. An interesting study in this context suggests that such mechanisms might be present in other animals as well, implying that they are evolutionarily old. Using imprinting behavior in chickens, Vallortigara and colleagues showed that newly hatched, visually naïve chicks respond to point-light animations. In their first

study (Vallortigara, Regolini, & Marconato, 2005), the authors found that chicks prefer to approach displays of biological motion more than they do randomly moving dots or a rigidly rotating object. This preference was not specific to a walking hen, but existed similarly for stimuli of a walking cat and also for scrambled versions of either of the two animals. In a follow-up study (Vallortigara & Regolin, 2006), the authors showed that newly hatched chicks respond to the facing direction of a coherent point-light hen, but only if presented right-side up. Together, these results suggest that the chicks respond to some feature present in the local biological motion of individual dots, and that, whatever this feature is, it is only effective if presented upright. We consider it likely that these newly hatched, entirely naïve chicks in these studies responded to the same gravity-based invariant that also indicates animacy and direction to humans.

In humans, there is evidence that at least some aspects of biological motion perception are in place very early in development. Simion, Regolin, and Bulf (2008) tested a group of 43 neonate infants with an average age of less than 2 days using a preferential looking paradigm. They presented these neonates with a choice between the point-light display of a hen and a display showing randomly moving dots and found that even at this very young age and with basically no prior visual experience, they already preferred the biological motion stimulus. The fact that neonates preferred the hen (they were not tested with human walkers) suggests that whatever they liked about the biological motion stimulus most likely had nothing to do with the specific shape (a hen, in this case) depicted by the stimulus. These results are in line with the assumption that the neonates oriented toward the gravity-defined signature of the feet and that this preference is either already present at birth or is at least triggered by minimal exposure during the first day of life.

## Biological Motion Perception in the Visual Periphery

If the proposed visual filters for the local motion of an agent's feet serve initial detection of the animal in the visual environment in order to trigger orientation responses and initiate further inspection, they should work well not just in foveal vision, but also in the visual periphery. This prediction is particularly interesting in the light of findings by Ikeda, Blake, and Watanabe (2005), who showed that, for some tasks, biological motion perception in extrafoveal areas of the visual field is extremely poor. In their experiments, observers were shown different point-light displays masked by scrambled versions of the target motion. Subjects had to indicate which of two subsequently presented intervals contained the coherent point-light action. The number of masking dots was varied, and detection thresholds were measured at different eccentricities from fixation and for different display sizes. Results showed that increasing the size of the display by any amount could not compensate for the decrease in sensitivity in the visual periphery.

The task in this experiment involved segregating a coherent figure from a background of dots based on the figure's motion-mediated structure. The local motions of target and background were identical and were thus not informative cues for solving the task. Under these conditions, biological motion perception seems to be poor in peripheral vision.

In the absence of a mask, however, peripheral perception of biological motion changes dramatically. Thompson et al. (2007) showed that observers could still discriminate the direction of coherent (i.e. unscrambled) walkers at 10 degrees eccentricity if the stimuli were presented unmasked. In their work, the authors used relatively large stimuli and, particularly for unmasked stimuli, the task was so easy that performance was close to ceiling in most conditions. Reaction times, however, showed an effect of eccentricity. In general, reaction times were a little longer for eccentrically presented stimuli. An interesting observation, however, was that this was not the

case for upright scrambled displays, which elicited the same response times at 10 degrees eccentricity as at fixation. The addition of a mask had very different effects on central and on peripheral presentation. While performance fell off quickly with increasing mask densities in the periphery, it remained constant and then decreased much more slowly for central vision. The authors thus concluded that while figure-ground segregation dramatically deteriorates in the visual periphery, biological motion perception is intact. Particularly, the ability to exploit the local components that convey direction from scrambled motion seems to work very well for peripheral vision.

In the above study, omitting the mask resulted in performances that were so close to ceiling that accuracy data were hard to interpret. Also, the sizes of the point-light displays were chosen to be such that they would not limit performance at the chosen eccentricities. In a more recent study, Gurnsey and colleagues (Gurnsey, Roddy, Ouhnana, & Troje, 2008) conducted a more systematic investigation of how the size of a point-light stimulus affects biological motion perception in the visual periphery. The authors adjusted the direction discrimination paradigm in order to avoid ceiling effects without having to employ a mask. In their study, observers had to identify the facing direction of a walker oriented only  $\pm 4$  degrees away from frontal view. Eccentricities up to 16 degrees were tested. Results indicated that the performances in the visual periphery matched the one for central vision if the size of the stimulus was increased by a factor that reflects the decrease in the number of neurons in visual cortex representing more peripheral areas of the visual field (i.e., the cortical magnification factor; see Levi, Klein, & Aitsebaomo, 1985).

In summary, the existing work on biological motion perception in the visual periphery clearly shows that sensitivity to biological motion is preserved. Conclusions about the relative

role of motion-mediated structure, on the one hand, and local motion cues, on the other hand, however, have to be drawn with care. While some of the existing results can be interpreted in favor of our theory that invariants in the local motion of the feet play a specific role in the visual periphery, additional experiments are required to test this prediction more explicitly.

## The Role of Attention

Another prediction that emerges from the idea that a visual filter sensitive to the motion of the feet of terrestrial, legged animals in locomotion is used to detect such animals in the visual environment is that this system does not require explicit attention, but rather attracts it. Thornton (Chapter 3, this volume) argues that this is not the case for biological motion tasks that aim at the recovery of motion-mediated shape. He and others, however, have provided a number of hints that this is, in fact, the case for other aspects of biological motion.

A very strong hint comes from a paper by Thornton, Rensink, and Shiffrar (2002). These authors studied the effect of different visual masks under dual-task conditions. The primary task was a walker direction discrimination task. In some blocks, a second task was introduced that observers were required to complete simultaneously. This second task was a change detection task that required concentrated attention from the observer. Two different masks were used. One of the masks consisted of scrambled walkers that therefore rendered any potential local cues in the walker display useless. The other mask consisted of stationary flickering dots. The main result of the experiment was a very prominent interaction between the mask that was being used and whether observers determined walking direction alone or simultaneously with completing the secondary task. While the choice of the mask had only moderate effects in the single-task condition, performance dropped from very good values for the flicker mask all the way to chance level for the scrambled mask. The study therefore showed that the stimulus that retained

accessibility to local cues was processed more passively and automatically and required less focused attention when compared to the case in which only global, structural cues were available to solve the task.

Another result that is in favor of the assumption that local biological motion is processed preattentively comes from a study by Bosbach, Prinz, and Kerzel (2004). These authors used stimulus–response compatibility effects to show that, even though the facing direction of a point-light display was entirely task-irrelevant, it still affected response times of key presses—either left or right—that observers were asked to execute in response to the color of the walker. Interestingly, these effects were also measured for scrambled walkers.

A more recent observation that speaks to the automatic, preattentive nature of local cues has been presented by Jiang and He (2007). They used an interocular noise suppression paradigm during which random dynamic noise is shown to one eye and a point-light display to the other. The time it takes for the point-light walker to become the dominant percept is taken as a measure for its salience. Interestingly, an upright scrambled walker display took less time to emerge from suppression as compared to an inverted scrambled walker against identical suppression noise. The authors interpreted this result as evidence for automatic, preattentive processing of the local, orientation-specific invariants still preserved in upright scrambled biological motion.

Some studies seem to contradict the above results since they failed to show that scrambled biological motion stimuli convey facing direction in an automatic, preattentive way. For instance, Thornton and Vuong (2004) tested the effect of coherent and scrambled walkers flanking a central biological motion display. Observers were asked to report the facing direction of the central walker, ignoring the additional flanking stimuli. Response times revealed

congruency effects for the coherent flankers, but not for the scrambled ones. Another study that did not show any effect of scrambled biological motion walkers was conducted by Fujimoto (2003). Here, it was shown that upright coherent biological motion induced the perception of motion direction on a stationary square-wave grating with periodic phase-reversals. Scrambled versions of their stimulus did not have that effect.

The lack of any effects of scrambled motion in these studies may be due to the particular stimulus that was used in these studies. Similar to many others studies on biological motion perception in the past, the above authors generated their walker stimuli by means of an algorithm that was published more than 30 years ago by Cutting (1978). The stimulus depicts the main joints of a walker in profile view. While it may well fulfill the purpose of studying perceptual organization of the dots into the coherent percept of a walker, it turns out that it does not contain (or at least to a much lesser degree) the local invariants that signal facing direction and animacy in more realistic walkers (Saunders et al., 2009).

## What role does shape play?

While the ability to derive facing direction from upright, scrambled biological motion is impressive, performance for these displays remains much lower than for unscrambled displays (e.g. Fig. 6-1). It seems straightforward to assume that the difference between the two conditions reflects the role of motion-mediated articulated shape. Only coherent motion allows the construction of the articulated shape of the body -- in a way replacing the missing sticks between the isolated dots. That shape, in turn, provides a wealth of information to determine facing direction (e.g. the knees point toward- and the elbows oppose the facing direction). It turns out that this interpretation might be incomplete, if not wrong. In a recent study (Hirai, Chang, Saunders, & Troje, 2011), we introduced a new stimulus that disrupts articulated

structure and deprives the display from any configural asymmetries that could have provided the observer with useful information. For these displays, performance on the facing direction discrimination task was almost comparable to that obtained for non-scrambled, coherent point light walkers. The new stimulus was derived by scrambling the dots only horizontally, and leaving their vertical position veridical. We also randomized the phase of the individual dots with respect to each other. Together, horizontal scrambling and phase scrambling guaranteed that the arrangement of the dots with respect to each other was devoid of any cue as to the facing direction of the walker. As with the fully scrambled walker, the only information about facing direction was contained in the local trajectories of individual dots.

Even though we presented these stimuli in a mask of 100 flickering (stationary, limited-lifetime) dots, observers derived the correct facing direction of the walker in more than 90% of the trials (Figure 6-7). This is much higher than the rate we obtained with fully scrambled walkers and is almost comparable to the performance obtained with unscrambled, coherent walkers in the earlier experiments (e.g., Figure 6-1).

To further investigate the cues being exploited in these displays, we created variations of the stimuli by inverting the local trajectory of each dot, or by manipulating the global organization of the walker by inverting it entirely or scrambling it randomly.

With the global layout of the walker held upright, local inversion resulted in a reversal of perceived direction with observers indicating the “correct” direction in only 25% of the trials. In the other 75% of the cases they perceived the walker to be facing the opposite direction.

Scrambling the vertical layout reproduced the results that we obtained in our previous experiments (Troje & Westhoff, 2006): locally upright trajectories still provided a strong sense of direction, while inverting them produced performances around chance level. Finally, inverting

the vertical layout of the dots resulted in 58% correct responses – still significant, but only moderately above chance level. In this last condition, local inversion had no effect.

What do these results tell us about the use of shape information for facing direction? The only “configural” information contained in the stimuli used in this experiment was encoded in the vertical location of the individual dots. The differences in performance between the fully scrambled and the globally upright condition in this experiment therefore cannot be attributed to the presence of horizontal asymmetries in global, motion-mediated shape. Rather than providing additional information about facing direction, the vertical arrangement of the dots seem to provide validity to the cues present in the directional information contained in the local trajectories.

In some way, we encountered a similar mechanism before. Inversion of the scrambled stimuli, that is, mirror flipping a structure about a horizontal axis as done by Troje and Westhoff (2006), does not change the amount or nature of the horizontal asymmetries that discriminate between the two facing directions. What it alters is whether or not the visual system is able (or willing?) to rely on this information. One cue as to the validity of local information seems to be congruency with gravitational direction: if the vertical acceleration of a dot is consistent with gravity pointing down, the horizontal asymmetries are used to determine facing direction. If they are incompatible with expectations about the direction of gravity, the same cues are not used.

In this last experiment, we discovered a second “validating” cue: vertical layout. If the dots that carry diagnostic horizontal asymmetries are located below the rest of the display, they are interpreted to indicate facing direction. Conversely, if these dots are located at the top of the display, they are ignored. Randomly scrambled biological motion contains trials in which the diagnostic dots happen to be located toward the bottom and other trials in which they fall in the

top part of the display. Performance observed with fully scrambled biological motion is therefore expected to be intermediate between the globally upright and the globally inverted condition – a pattern clearly reflected in our data (Figure 6-7).

Although we have speculated on cues that seem to validate local motion information, what is the actual cue to facing direction? The fact that perceived direction reverses to some degree in the conditions in which trajectories are locally inverted suggests that facing direction is determined by whether the feet follow a clockwise or a counter-clockwise trajectory. We termed this the “rolling wheel” cue: A point on the rim of a wheel rolling on the ground from left to right follows a clockwise trajectory. If it rolls from right to left, it instead subscribes a counter-clockwise trajectory. The “rolling wheel” cue in itself contains directionality, but it seems that the visual system only exploits it in cases in which it can assume that it really comes from a foot: If it is close to the ground, and if vertical acceleration is consistent with the way feet normally give way to gravity. Shape plays a role, but it is not necessarily the motion-mediated articulated structure of the body and the horizontal asymmetries contained in a sagittal projection, but rather the vertical arrangements that modify the validity of local cues.

## Conclusion

The pendular and/or ballistic movements of the extremities of an animal in locomotion contain information about the direction in which an animal is facing and convey the percept of animacy. Both the ability to derive facing direction and the perception of animacy depend on the orientation of the stimulus and disappear when it is turned upside-down. Based on these findings, we have suggested that the visual invariants responsible for these effects play a role in detecting biological motion in the visual environment independently of its particular shape and identity. Such a detection system may serve two main functions: First, it alerts the observer to a

potentially dangerous or otherwise demanding situation. The presence of another living being might require immediate action, and the other agent should at least be inspected to the point at which the observer can decide whether it is a potential predator, prey, or conspecific. Second, the developing visual system would benefit from a mechanism that helps the organism to orient toward a visual stimulus it has to learn so much about. Animate motion contains a wealth of information above and beyond signaling the mere presence of another agent. To become an expert on retrieving this information, the developing visual system has to take in as many examples as possible in order to learn the complex statistics of this stimulus class.

A number of predictions derive from this theory. In this chapter, we discussed a few of them in light of recent research. Based on our own work, we showed that the same invariant responsible for retrieving facing direction from scrambled motion also contributes to the percept of animacy in these stimuli. Furthermore, we confirmed that the observed orientation specificity is in fact based on signatures reflecting the impact of gravitational acceleration on the moving body. Evidence from other research groups supports the notion that the early detection mechanism is innate in humans. The fact that the same invariants also control preferences in the context of imprinting behavior in chickens indicates that it is evolutionarily old and that it may be present in other species as well. Work on biological motion perception in the visual periphery shows that the ability to process direction from local motion invariants is preserved for peripheral vision, while further inspection of the stimulus based on motion-mediated structure requires central vision. Finally, we discussed a number of studies in support of the prediction that the local aspects of biological motion are processed passively and automatically, but pointed out that current evidence that these invariants capture attention and control orientation responses is still relatively weak.

We did not discuss here the extensive work that has been done to localize areas of the brain involved with biological motion processing. However, we want to make a few comments here in that respect as they relate to the specific mechanism discussed in this chapter.

Most of the work on functional human brain anatomy with respect to biological motion perception looks at biological motion as representing one single phenomenon. Observers are presented with biological motion stimuli, and the resulting blood oxygen level-dependent (BOLD) MRI response is contrasted with the one obtained when presenting “nonbiological” control stimuli. One control stimulus that is often used for this purpose is scrambled biological motion. The contrast between coherent and scrambled biological motion perfectly represents the visual system’s ability to perceptually organize a point-light display and derive the underlying articulated structure from it. However, both coherent and scrambled biological motion contain to the same extent the signatures we discussed in this chapter – namely the invariants contained in the local motion of individual dots. This contrast is, therefore, unsuitable to identify parts of the brain that respond to these invariants. The appropriate contrast would be the contrast between upright and inverted scrambled biological motion.

Which brain areas would we expect to be responsible for analyzing local biological motion? If the preferences observed in neonate human infants and in chickens for upright “foot” motion really reflect homologous developments, we would not expect them to be localized in the neocortex at all. If they are old enough to be shared with animals that never evolved a neocortex, they would rather be expected to be implemented in older, subcortical parts of the visual system, probably the tectopulvinar system. In birds, this is the primary visual system, with the optic tectum (located in the midbrain) being one of the most prominent structures. In mammals, the structure homologous to the optic tectum, the superior colliculus, is known to drive orientation

responses in many animals, and it is responsible for directing eye movements in humans. It might, therefore, be a likely candidate for the implementation of the early detection of biological motion via gravity-driven visual invariants obtained from the motion of the feet of animals in locomotion.

In conclusion, we want to emphasize that the proposed mechanism is, of course, specific to terrestrial, legged animals. It would not respond to all sorts of other animal classes, such as snakes, flying birds, or fish. It would also not respond to animals so small and light that air resistance dominates gravity-based effects (insects and other invertebrates, maybe even small vertebrates). In that sense, the proposed mechanism is not a general “life detector.” However, it covers a very prominent part of the animal kingdom, particularly one that constitutes predatory pressure to mid-sized terrestrial animals such as ourselves.

## References

- Aggarwal, J. K., & Cai, Q. (1999). Human motion analysis: A review. *Computer Vision and Image Understanding*, 73, 428–440.
- Alexander, R. M. (1989). Optimization and gaits in the locomotion of vertebrates. *Physiological Reviews*, 69(4), 1199–1227.
- Alexander, R. M. (2003). *Principles of animal locomotion*. Princeton, NJ: Princeton University Press.
- Baker, C. L., Jr., & Braddick, O. J. (1985). Temporal properties of the short-range process in apparent motion. *Perception*, 14(2), 181–192.
- Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics*, 23, 145–152.
- Beintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Science USA*, 99(8), 5661–5663.
- Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*, 5(4), 221–225.
- Bosbach, S., Prinz, W., & Kerzel, D. (2004). A Simon effect with stationary moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1), 39–55.
- Casile, A., & Giese, M. A. (2005). Critical features for the recognition of biological motion. *Journal of Vision*, 5, 348–360.
- Cavagna, G. A., Saibene, F. P., & Margaria, R. (1963). External work in walking. *Journal of Applied Physiology*, 18, 1-9.
- Cavagna, G. A., Heglund, N. C., & Taylor, C. R. (1977). Mechanical work in terrestrial locomotion: two basic mechanisms for minimizing energy expenditure. *American*

- Journal of Physiology – Regulatory, Integrative and Comparative Physiology*, 233, 243-261.
- Chang, D. H., & Troje, N. F. (2008). Perception of animacy and direction from local biological motion signals. *Journal of Vision*, 8(5), 3 1–10.
- Chang, D. H., & Troje, N. F. (2009a). Acceleration carries the local inversion effect in biological motion perception. *Journal of Vision*, 9(1), 19 11–17.
- Chang, D. H., & Troje, N. F. (2009b). Characterizing global and local mechanisms in biological motion perception. *Journal of Vision*, 9(5), 8 1–10.
- Cutting, J. E. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behavior Research Methods & Instrumentation*, 10, 91–94.
- Cutting, J. E. (1981). Coding theory adapted to gait perception. *Journal of Experimental Psychology: Human Perception and Performance*, 7(1), 71–87.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5), 353–356.
- Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception & Psychophysics*, 44(4), 339–347.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115(2), 107–117.
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, 22(1), 15–22.
- Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727–738.

- Fleischer, F., & Giese, M. A. (2012). Computational mechanisms of the visual processing of action stimuli. In K. Johnson & M. Shiffrar (Eds.), *People watching: Social, perceptual, and neurophysiological studies of body perception* (Chapter 22). New York: Oxford University Press.
- Fujimoto, K. (2003). Motion induction from biological motion. *Perception, 32*, 1273–1277.
- Gurnsey, R., Roddy, G., Ouhana, M., & Troje, N. F. (2008). Stimulus magnification equates identification and discrimination of biological motion across the visual field. *Vision Research, 48*(28), 2827–2834.
- Hecht, H., Kaiser, M. K., & Banks, M. S. (1996). Gravitational acceleration as a cue for absolute size and distance? *Perception & Psychophysics, 58*(7), 1066–1075.
- Hirai, M., Chang, D. H. F., Saunders, D. R., & Troje, N. F. (2011). Body configuration modulates the usage of local cues to direction in biological motion perception. *Psychological Science 22*, 1543-1549
- Hiris, E., Humphrey, D., & Stout, A. (2005). Temporal properties in masking biological motion. *Perception & Psychophysiology, 67*(3), 435–443.
- Ikeda, H., Blake, R., & Watanabe, K. (2005). Eccentric perception of biological motion is unscalably poor. *Vision Research, 45*(15), 1935–1943.
- Jiang, Y., & He, S. (2007). Isolating the neural encoding of the local motion component in biological motion. *Journal of Vision, 7*(9), 551a.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*(2), 201–211.
- Johansson, G. (1974). Vector analysis in visual perception of rolling motion. A quantitative approach. *Psychologische Forschung, 36*(4), 311–319.

- Johnson, M. H. (2006). Biological motion: A perceptual life detector? *Current Biology*, *16*(10), R376–R377.
- Jokisch, D., & Troje, N. F. (2003). Biological motion as a cue for the perception of size. *Journal of Vision*, *3*(4), 252–264.
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*(6), 575–580.
- Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. *Journal of Neuroscience* *26*(11), 2894–2906.
- Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception*, *30*(1), 73–83.
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*(7), 963–977.
- Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society B: Biological Sciences*, *249*(1325), 149–155.
- Mather, G., & West, S. (1993). Recognition of animal locomotion from dynamic point-light displays. *Perception*, *22*(7), 759–766.
- Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260.
- McConnell, D. S., Muchisky, M. M., & Bingham, G. P. (1998). The use of time and trajectory forms as visual information about spatial scale in events. *Perception & Psychophysics*, *60*(7), 1175–1187.
- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, *98*(2), 164–181.

- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, *395*(6705), 894–896.
- Pittenger, J. B. (1985). Estimation of pendulum length from information in motion. *Perception*, *14*(3), 247–256.
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51–B61.
- Rubin, E. (1927). Visuell wahrgenommene wirkliche Bewegungen [Visually perceived genuine motions]. *Zeitschrift für Psychologie*, *103*, 384–392.
- Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(4), 733–740.
- Saunders, D. R., Suchan, J., & Troje, N. F. (2009). Off on the wrong foot: Local features in biological motion. *Perception*, *38*(4), 522–532.
- Shipley, T. F. (2003). The effect of object and event orientation on perception of biological motion. *Psychological Science*, *14*(4), 377–380.
- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceeding of the National Academy of Science USA*, *105*(2), 809–813.
- Stappers, P. J., & Waller, P. E. (1993). Using the free fall of objects under gravity for visual depth estimation. *Bulletin of the Psychonomic Society*, *31*(2), 125–127.
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, *13*(3), 283–286.
- Thompson, J. C. (2012). The how, when, and why of configural processing in the perception of human movement. In K. Johnson & M. Shiffrar (Eds.), *People watching: Social*,

- perceptual, and neurophysiological studies of body perception* (Chapter 18). New York: Oxford University Press.
- Thompson, B., Hansen, B. C., Hess, R. F., & Troje, N. F. (2007). Peripheral vision: Good for biological motion, bad for signal noise segregation? *Journal of Vision*, 7(10), 1–7.
- Thornton, I. M. (2012). Top-down versus bottom-up processing of biological motion. In K. Johnson & M. Shiffrar (Eds.), *People watching: Social, perceptual, and neurophysiological studies of body perception* (Chapter 3). New York: Oxford University Press.
- Thornton, I. M., Rensink, R. A., & Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception*, 31(7), 837–853.
- Thornton, I. M., & Vuong, Q. C. (2004). Incidental processing of biological motion. *Current Biology*, 14(12), 1084–1089.
- Troje, N. F. (2008). Biological motion perception. In Basbaum, A. I. (Ed.), *The senses: A comprehensive references* (pp. 231–238). Oxford: Elsevier.
- Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: Evidence for a “life detector”? *Current Biology*, 16(8), 821–824.
- Vallortigara, G., & Regolin, L. (2006). Gravity bias in the interpretation of biological motion by inexperienced chicks. *Current Biology*, 16(8), R279–R280.
- Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biology*, 3(7), e208.
- Webb, J. A., & Aggarwal, J. K. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence*, 19(1), 107–130.

## Figure captions

Figure 6-1 Results of a direction discrimination experiment. Stimuli were shown either coherent or scrambled. In addition, stimuli were presented entirely upright or inverted, or with only parts of the display inverted. Specifically, the experiment contained a condition in which the whole walker was upright, but the trajectory of the two dots representing the feet was inverted. For instance, for the coherent version of the “feet-only” condition, the feet were still at the bottom of the display and everything else was in place, too, but the local motion trajectory of the feet was flipped upside-down. In contrast, in the coherent version of the “all but feet” condition, the whole walker was presented upside-down (head at the bottom, feet at the top of the display), but the trajectory of the feet was vertically flipped with respect to the rest of the body and thus upright. Figure adapted from Troje and Westhoff (2006).

Figure 6-2 Mean ratings assigned to coherent and scrambled, upright and inverted point-light walkers by observers asked to indicate on a Likert scale the degree as to whether each stimulus was “alive.” Error bars represent  $\pm 1$  standard error of the mean. Figure adapted from Chang and Troje (2008).

Figure 6-3 The construction of motion displays derived from individual fragments of the foot motion. The full trajectory of the foot of a right-facing human walker is shown in (A). Starting positions for each fragment specified at equal temporal intervals are indicated by the filled circles along the path. The trajectories of the individual fragments extended from the starting positions and lasted for 150 ms (B). A stimulus was derived from multiple copies of a single fragment as shown in (C), or a particular pair of fragments, as depicted in (D), for an example based on fragments 2 and 7. Figure adapted from Chang and Troje (2009a).

Figure 6-4 Results of a direction discrimination experiment in which only fragments of foot motion were shown. Presented in (A) are data from an experiment in which each display consisted of 10 identical segments (Figure 6-3c). For a better comparison with the following experiment, the average proportion of correct responses over two segments (the same pairs as used below) is shown. Presented in (B) are data from an experiment in which each stimulus consisted of multiple samples of two fragments (Figure 6-3d). Figure adapted from Chang and Troje (2009a).

Figure 6-5 Correlation between the inversion effect and the vertical acceleration contained in the stimulus fragments. The inversion effect is quantified by a difference score obtained by computing the difference in discrimination accuracies between the upright and inverted versions for each stimulus. Two outliers from this dataset (marked by asterisks) were excluded from the analysis. Figure adapted from Chang and Troje (2009a).

Figure 6-6 Behavioral performances on a direction discrimination task for natural stimuli derived from the human walker foot and constant speed stimuli that were identical to the natural stimuli but had accelerations removed along the trajectories. Error bars represent  $\pm 1$  standard error of the mean. Figure adapted from Chang and Troje (2009a).

Figure 6-7 Performance in a direction discrimination experiment in which all stimuli were horizontally scrambled in order to deprive them of any structural cues indicating their facing direction. The global organization was either normal (upright), inverted, or scrambled. Local trajectories were presented either upright or inverted. Figure adapted from Hirai et al. 2011.

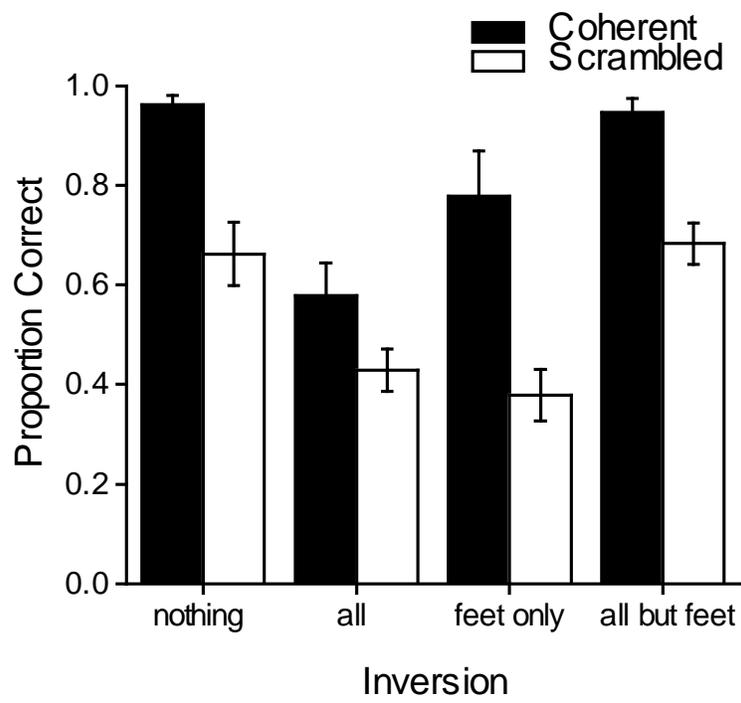


Figure 1

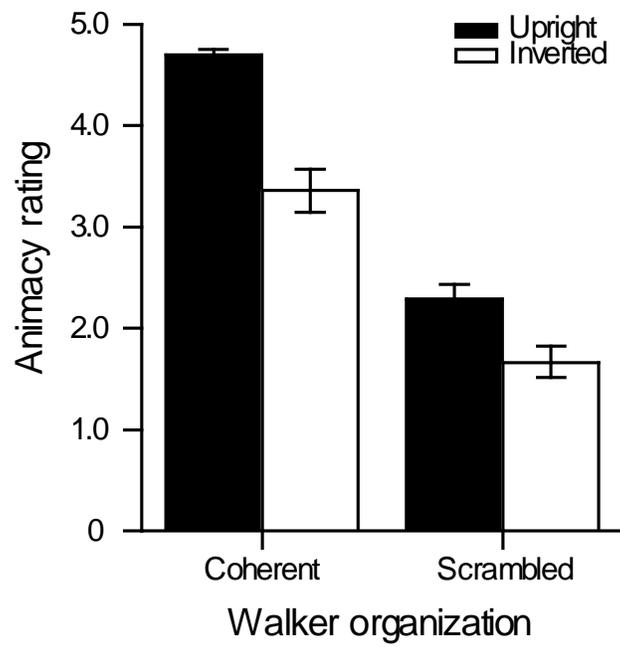


Figure 2

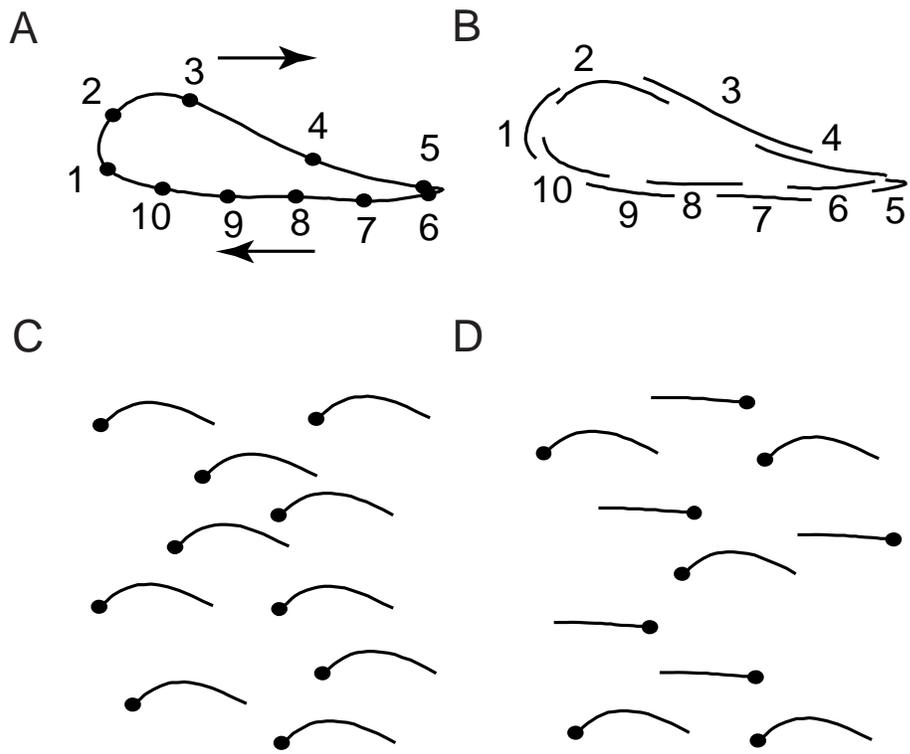


Figure 3

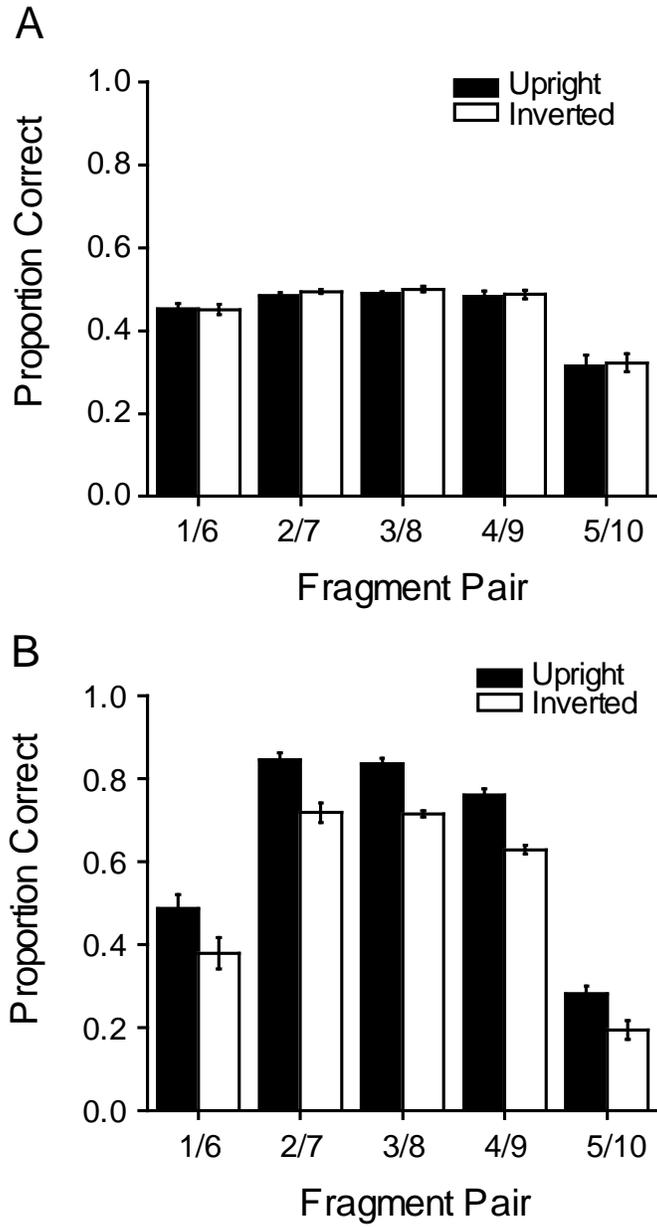
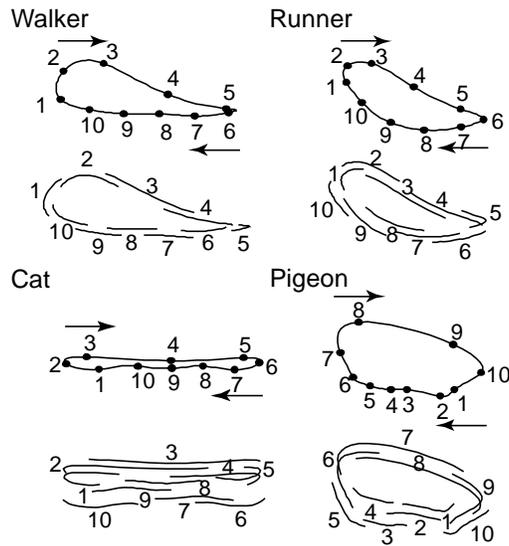


Figure 4

A



B

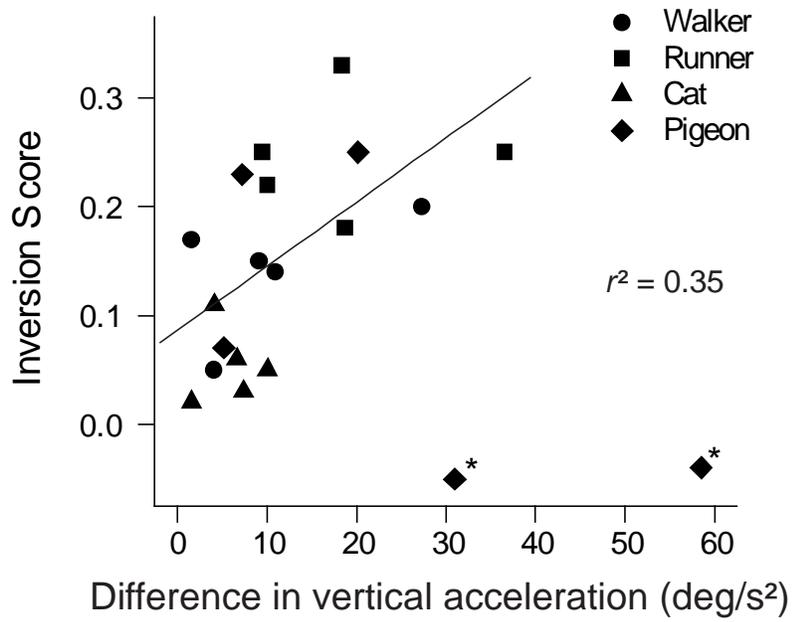


Figure 5

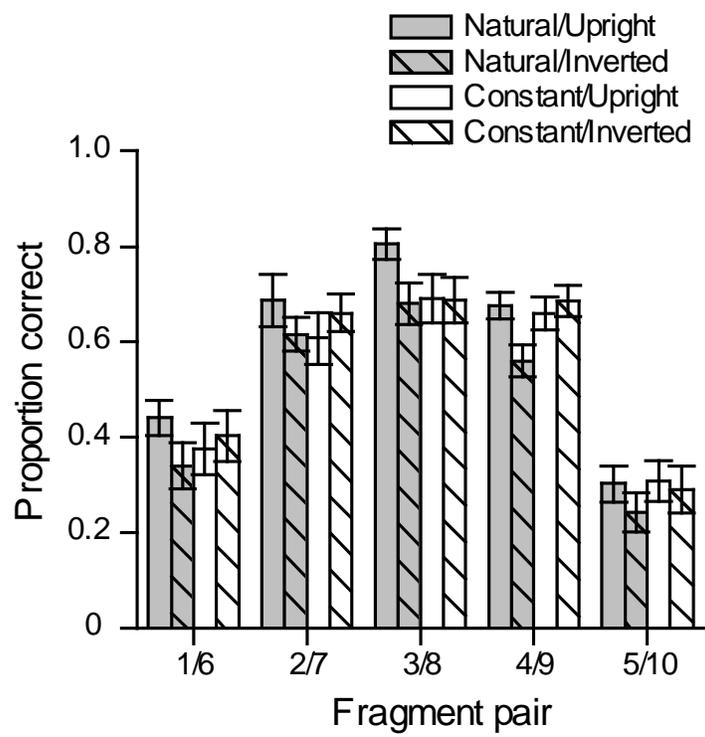


Figure 6

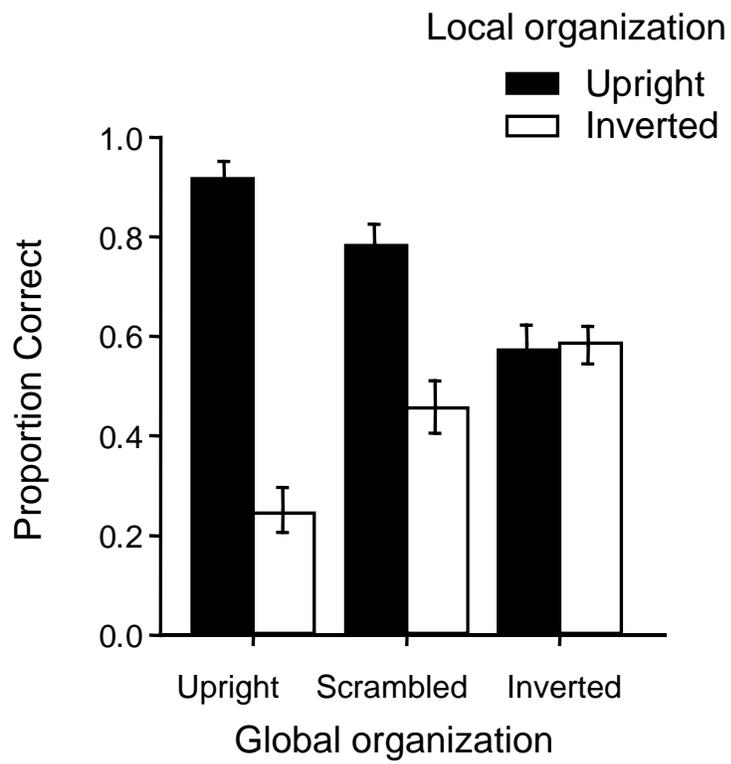


Figure 7