Report

The Inversion Effect in Biological Motion Perception: Evidence for a "Life Detector"?

Nikolaus F. Troje[1,2,*] and Cord Westhoff[2]
[1] Department of Psychology
Queen's University
Kingston, Ontario, K7L3N6
[2] Department of Psychology
Ruhr-Universität Bochum
44780 Bochum
Germany

Summary

If biological-motion point-light displays are presented upside down, adequate perception is strongly impaired [1, 2]. Reminiscent of the inversion effect in face recognition, it has been suggested that the inversion effect in biological motion is due to impaired configural processing in a highly trained expert system [3–5]. Here, we present data that are incompatible with this view. We show that observers can readily retrieve information about direction from scrambled point-light displays of humans and animals. Even though all configural information is entirely disrupted, perception of these displays is still subject to a significant inversion effect. Inverting only parts of the display reveals that the information about direction, as well as the associated inversion effect, is entirely carried by the local motion of the feet. We interpret our findings in terms of a visual filter that is tuned to the characteristic motion of the limbs of an animal in locomotion and hypothesize that this mechanism serves as a general detection system for the presence of articulated terrestrial animals.

Results and Discussion

A few light dots placed strategically on a moving human or animal body are instantaneously organized into the coherent percept of a living creature [6]. Among the very early findings in the context of biological motion perception is the observation that perception of biological motion is strongly impaired if the point-light displays are shown upside down [1–3, 7]. This inversion effect seems to be relatively independent of the experimental task and affects detection of a masked walker [2, 8, 9], direction discrimination [8], sex classification [10], and action recognition [4, 11].

The inversion effect for biological motion perception is reminiscent of a similar observation in human face recognition. If faces are inverted, recognition is severely impaired (for a review, see [12]). The inversion effect is absent only in experiments that employ artificial facial stimuli differing only in the appearance of local features while keeping their configuration constant [13–15]. The

*Correspondence: troje@post.queensu.ca

face-inversion effect therefore seems to be carried by configural processing (for an alternative view, see [16]).

It seems obvious to assume that a similar mechanism is responsible for the inversion effect in biological motion. Even though the motion of a human body is generally nonrigid, many of its parts move rigidly. Configural invariants are particularly inherent to the articulated structure of the body and the constant distances between connected joints. The fact that the inversion effect persists in detection tasks using scrambled motion masks that render all local information useless argues for a role of configural information [8, 17]. It is therefore not surprising that the inversion effect in biological motion has been attributed to an impairment of configural processing [3–5].

In this study, we do not want to challenge the contribution of configural processing to the inversion effect in biological motion, but we want to demonstrate that there is a second, completely independent, and more prominent cause. In the first experiment, we will directly compare the effect of inversion to the effect of spatially scrambling biological motion. In scrambled point-light displays, local motion is kept intact but the single-dot trajectories are randomly displaced within the restricted area of the display, entirely disrupting the shape of the figure. Scrambled motion contains no configural information. If the inversion effect is due to configural information processing, no inversion effect should be observed with scrambled biological motion.

Here, we show that this is not the case, and we discuss an alternative explanation of the inversion effect in biological motion. We use a task in which adult human participants are asked to report the perceived direction of a stationary walking point-light figure. In order to show that the obtained results are not restricted to human walking, we also use point-light displays of a walking cat and a walking pigeon. In the first experiment, we show that upright, spatially scrambled motion still contains information about the direction of a walker, but only if the walker is presented upright. In Experiment 2, we show that temporal scrambling also has very little effect on direction perception and the associated inversion effect. On the basis of the conclusion that relevant information is contained in local dot trajectories, we investigate the specific role of the feet in Experiment 3.

In the first experiment, we used four conditions arranged in a 2 × 2 factorial design: Point-light displays could be either upright or inverted, and they either had their veridical configuration or were spatially scrambled. The mean accuracy for unscrambled stimuli (both upright and inverted) was 0.83 and for scrambled stimuli 0.64, corresponding to a reduction in correct response rate of 0.19. Inversion affected accuracy by reducing the average performance rate from 0.88 to 0.59 (difference: 0.29). A two-way repeated-measures ANOVA revealed that both main effects were highly significant [scrambling: $F_{(1,9)} = 86.1$, $p < 0.001$; inversion: $F_{(1,9)} = 38.3$, $p < 0.001$]. The interaction between the two factors

was not significant [F(1,9) = 1.1], reflecting the fact that the performance decrease caused by inversion is similar for unscrambled (0.33) and for scrambled (0.24) motion (Tukey's Honest Significant Difference is 0.188, p < 0.01). The participant's sex and type of creature (human, pigeon, or cat) had no effect.

The results of this first experiment are not compatible with the hypothesis that the inversion effect in biological motion is due to impaired configural processing. However, even though the spatial configuration of the dots is not essential, the temporal pattern among them might be of importance. The second experiment was designed to investigate whether this is the case. Here, we added conditions in which, in addition to spatial scrambling, the relative phase between the different dots was randomly offset (Figure 1A, conditions e and f). Note that after this manipulation, all dots still move with the same fundamental frequency. Even though the phases between the dots are no longer veridical, the phase relations among them stay constant throughout a single trial, and the pattern still resembles a rhythmic gait. Therefore, we also added conditions in which the speeds of the individual dots were multiplied by constant factors picked randomly within a range between 0.5 and 2 (conditions g and h).

The performance exhibited in the first four conditions (conditions a–d) confirms the results of Experiment 1. A separate analysis of the accuracy data of these first four conditions shows that both scrambling and inversion affect performance [F(1.9) = 14.3, p < 0.005, and F(1,9) = 67.2, p < 0.001, respectively] and that inversion similarly affects intact and scrambled displays (no interaction between the two factors). Recognition of upright, scrambled biological motion is better than recognition of inverted scrambled motion (Tukey's HSD test, p < 0.01).

In a second analysis, we compared the different conditions that employ stimuli with scrambled motion (conditions c – h) by means of a two-way ANOVA with factors Inversion and Scrambling. Here, the second factor has the three levels *spatially scrambled*, *phase scrambled*, and *random frequency*. Inversion has a highly significant effect [F(1,9) = 122, p < 0.01]. No main effect was found for factor Scrambling, and there was no interaction between the two factors either.

We also looked at effects of observer's sex and the type of creature on response accuracy by running an ANOVA involving all conditions. Neither of these two factors showed any effect or an interaction with the other factors.

Apparently, a considerable part of the inversion effect is carried by information from local dot trajectories rather than by spatial or temporal configurations. Inversion means basically mirror flipping about a horizontal axis and therefore only affects the vertical component of a trajectory. Hence, we expect that the inversion effect is primarily carried by dots that show significant vertical motion. In Experiment 3, we therefore looked at the specific role of the feet, which contribute the largest part of the overall variance to the vertical movement of all dots. In addition to the conditions used in Experiment 1, we added more levels to the factor Inversion. In the first one, only the feet were inverted, and in the second, all dots except the ones representing the feet were inverted (Figure 1B, conditions i–m).
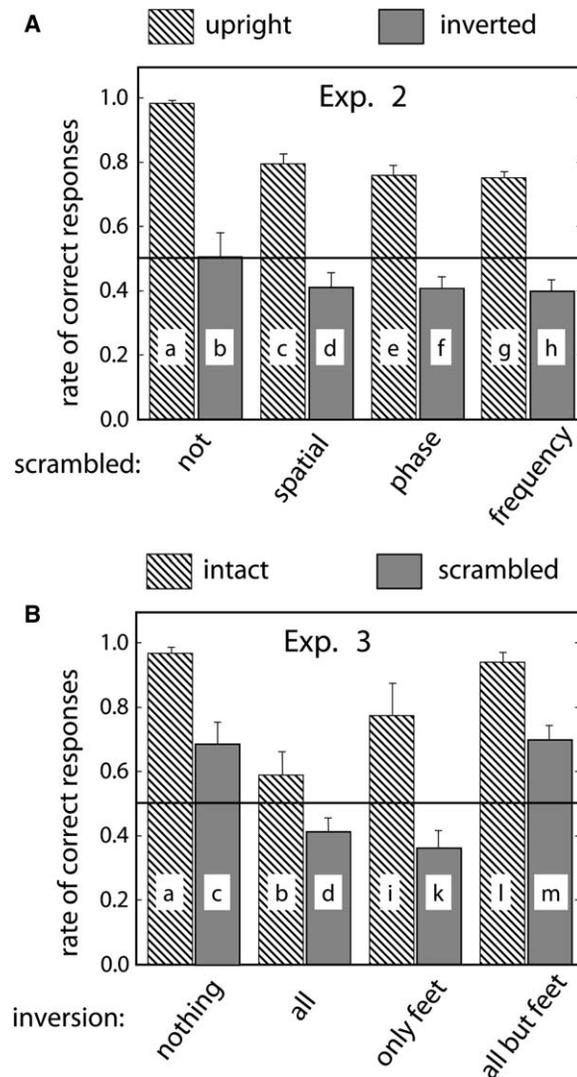

Figure 1. Average Accuracies for Experiments 2 and 3

Error bars indicate standard errors of the mean. Tukey's HSD for p < 0.01 are 0.265 for Experiment 2 and 0.185 for Experiment 3. Lowercase letters refer to the different conditions as described in the text.

The performance exhibited in the first four conditions confirmed the results of Experiment 1 (Figure 1C). A 2-way ANOVA with factors Scrambling and Inversion (the latter with all four levels *nothing*, *all*, *only feet*, *all but feet*) shows main effects of both factors [F(3,27) = 46.6, p > 0.001, and F(1,9) = 17.6, p < 0.001, respectively] and no interaction between them. The interesting point about the other conditions is that inverting only the feet has a much stronger effect than inverting all points except the feet. Whereas the difference between the role of the feet and the rest of the body is less pronounced for unscrambled stimuli (in fact, the difference is smaller than Tukey's HSD), it becomes highly significant for the scrambled stimuli (p < 0.01). A separate one-way ANOVA of the conditions with scrambled stimuli shows a strong effect of factor Inversion [F(3,27) = 10.2, p < 0.001] and reveals significantly higher performance for the conditions in which the feet were not inverted (*nothing* and *all but feet*) than for the conditions in which they

were inverted (*all* and *only feet*) (Tukey's HSD test, p < 0.01). No differences between the two conditions *nothing* and *all but feet* and the conditions *all* and *only feet* were found.

These results challenge the view that the inversion effect in biological motion is a single, unitary phenomenon. Rather, there are at least two different inversion effects at work. One of them is probably due to the fact that the inverted, articulated shape of a body is harder to match to an experience-based template than a normal upright shape. This "shape inversion effect" is probably closely related to the well-known face-inversion effect and also to the more recently reported body-inversion effect [5]. Here, motion would only play a role in revealing the articulation of the body from the point-light displays and could as well be replaced by the sticks in a static stick figure.

In addition to this shape-based inversion effect, there is clearly a second one that is completely independent of the arrangement of the dots into a coherent structure and depends only on the motion of the feet. The characteristic ballistic-velocity profile, which is created when an articulated, terrestrial animal in locomotion propels its limbs away from the ground and then—playing gravity as efficiently as possible—lets them fall back down, has the potential to provide a reliable cue for the presence and the location of an animal in the visual environment. In a world of constant gravity, spatial and temporal parameters are highly dependent, therefore introducing a redundancy that attributes a highly specific signature to ballistic movements. Given that such movements occur only if energy is employed and dissipated, it is hard to imagine how something other than a living creature can generate similar motion patterns—at least outside of our modern, technological world.

A number of previous studies have shown that the human visual system does in fact employ dynamic models about the physical environment. In particular, the relation between spatial and temporal parameters defined by gravitational acceleration can be used as a cue to the size of animate and inanimate objects, given the time course of events that involve ballistic or pendular components [18–20]. Observers can derive dynamic information from the kinematics of body movements, such as the weight of a box being lifted from the ground [21–23] or the target of a dart being thrown [24].

Direct evidence that dynamic, gravity-dependent models are used to interpret biological motion and are related to the inversion effect comes from a recent study in which the author showed that the point-light display of a gymnast walking on his hands is readily recognized, whereas recognition is impaired if the same display is presented upside down [11]. Support for the idea of a local visual filter that operates independently of structure and form comes from work with infants and animals. When 3- to 5-month-old infants were habituated to a translating point-light walker, a significant amount of response recovery was found when the walker was turned upside down. This was not the case when the figure of the walker was spatially scrambled [25]. Recently, it was shown that newly hatched chicks show a preference for upright biological motion over rigidly moving objects and other control stimuli [26]. The particular shape of the point-light display played no role as long as the displays were shown upright rather than inverted [27].

Shaping this visual filter and tuning it to the invariants of animal locomotion might be achieved through individual learning or by means of evolutionary adaptation. However, the results obtained from newly hatched chicks and from infants argue for an innate and possibly evolutionary old mechanism that the human visual system shares with other animals. The observation that it is relatively easy to get close to wild animals in a car, a canoe, or a similar vehicle might be due to the absence of the typical movement of feet. Similarly, the creeping movement of a hunting cat can be interpreted in terms of disguising the ballistic component in its locomotion.

Our findings about the role of the feet as a cue to the direction of motion of scrambled point-light displays support the notion of such a general "life detector." At this point, however, this idea remains speculative, and further experiments using tasks that focus more directly on the detection of animacy are required to consolidate this view.

Experimental Procedures

Participants
Data were obtained from altogether 15 male and 15 female students from Ruhr-University in Bochum, Germany. All of them received course credit for their participation. Each individual participated only in one experiment so that five male and five female observers participated in each of the three experiments. None of them had participated in any other study using point-light displays before.

Stimuli
Stimuli were based on three different point-light sequences. The first one showed a human figure that represented the average walker computed from a database consisting of motion-captured data from 50 women and 50 men. For details about the data acquisition and about the algorithm used to average the 100 walkers, see [28]. A set of 11 markers was used, which represented the head, one shoulder, one hip, the two elbows, two wrists, two knees, and two ankles. The second point-light sequence showed a walking pigeon. Data were recorded from a pigeon that was trained to walk back and forth between two feeders. The pigeon was supplied with 11 retro-reflective markers, which were captured by means of an optical motion capture system (Vicon, Oxford Metrics). Constant translation was subtracted from the data, and a third-order Fourier series was fitted to the data to smooth them and loop them into a repeatable walking cycle. The third sequence showed a walking cat. The data are based on a high-speed (200 fps) video sequence showing a cat walking on a treadmill. Fourteen feature points were manually sampled from single frames. As with the pigeon sequence, data were approximated with a third-order Fourier series to obtain a generic walking cycle.

All three animals (including the human walker) presented stationary walking (as on a treadmill) and were shown in sagittal view. The displays were run at veridical speed. Gait frequencies were 0.93 Hz for the human, 1.6 Hz for the pigeon, and 1.7 Hz for the cat. The display size on the screen was 3.1 × 5.7 degrees of visual angle for the human, 4.4 × 4.4 degrees for the pigeon, and 5.7 × 3.1 degrees for the cat.

All point-light displays were embedded in a mask of 100 randomly positioned stationary dots with a limited lifetime of 170 ms. The size of the mask on the screen subtended an area of 9.6 × 9.6 degrees.

All other stimuli were derived from these veridical, upright point-light displays by inverting the whole display or only parts of it, and by scrambling them spatially and temporally. Inversion was achieved by either mirror flipping the whole display about a horizontal axis (conditions b, d, f, h) or by mirror flipping only selected dot trajectories about a horizontal axis passing through the center of gravity of this trajectory (conditions i, k, l, m). In these conditions,

either only the feet or all the dots except the feet were inverted. The number of dots specifying the motion of the feet was different in the three different creatures. In the human figure, only the two ankle markers were used. In the pigeon, four markers were considered part of the feet (the two front and the two back toes), and in the cat, six markers (the ankle and toe on each front leg and one toe marker on each hind leg) were considered to belong to the feet.

Spatial scrambling was done by placing each trajectory at a random position within a 4.4 × 4.4 degree square in the center of the display area (conditions c–m). Scrambling the phase was implemented by replacing each dot's veridical phase with a phase picked randomly between 0 and 360 degrees. Scrambling the frequency was done by multiplying the veridical speed of the walker with a number between 0.5 and 2, picked randomly according to a uniform distribution on a logarithmic scale.

Design and Procedure
Participants were seated at a viewing distance of 65 cm to a 17 inch CRT computer screen. They were told that they would be seeing a number of different animals as point-light displays masked with a clutter of flickering dots. We instructed them not to worry about the particular nature of the animals and simply to respond to the direction in which the animal seems to be moving. Responses had to be given by pressing the left or right arrow key. After instructions were given, participants completed 15 practice trials. The first four practice trials showed unmasked, intact walkers (cat in both directions, pigeon, human). The following 11 trials showed different versions of masked stimuli including scrambled and inverted versions.

Two-factor designs were used in all three experiments. The first factor encoded Inversion and the second Scrambling. For each of the possible combinations, we ran 24 trials, counterbalancing for the nature of the walker (human, cat, pigeon) and the direction of walking (left, right) and randomizing the order of all trials within the experimental session.

In Experiment 1, both factors had two levels, and therefore the session consisted of a total of 96 trials. Experiment 2 employed four different levels of factor Scrambling and a total of 192 trials. In Experiment 3, the factor Scrambling had again only two levels but the factor Inversion had four, again resulting in 192 trials.

Each display remained on the screen until a response was given and was followed by an intertrial interval of 2 s. Running Experiment 1 took only about 6 min. Experiments 2 and 3 lasted twice as long.

Supplemental Data
Supplemental Data include an animation that shows the stimuli used in Experiment 1 (also identical with conditions a to d in Experiments 2 and 3) and are available with this article online at: http://www.current-biology.com/cgi/content/full/16/8/821/DC1/.

References

1. Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. Perception *13*, 283–286.
2. Pavlova, M., and Sokolov, A. (2000). Orientation specificity in biological motion perception. Percept. Psychophys. *62*, 889–899.
3. Proffitt, D.R., and Bertenthal, B.I. (1990). Converging operations revisited: Assessing what infants perceive using discrimination measures. Percept. Psychophys. *47*, 1–11.
4. Dittrich, W.H. (1993). Action categories and the perception of biological motion. Perception *22*, 15–22.
5. Reed, C.L., Stone, V.E., Bozova, S., and Tanaka, J. (2003). The body-inversion effect. Psychol. Sci. *14*, 302–308.
6. Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Percept. Psychophys. *14*, 201–211.
7. Ikeda, H., Blake, R., and Watanabe, K. (2005). Eccentric perception of biological motion is unscalably poor. Vision Res. *45*, 1935–1943.
8. Bertenthal, B.I., and Pinto, J. (1994). Global processing of biological motions. Psychol. Sci. *5*, 221–225.
9. Troje, N.F. (2003). Reference frames for orientation anisotropies in face recognition and biological-motion perception. Perception *32*, 201–210.
10. Barclay, C.D., Cutting, J.E., and Kozlowski, L.T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. Percept. Psychophys. *23*, 145–152.
11. Shipley, T.F. (2003). The effect of object and event orientation on perception of biological motion. Psychol. Sci. *14*, 377–380.
12. Farah, M.J., Tanaka, J.W., and Drain, H.M. (1995). What causes the face inversion effect? J. Exp. Psychol. Hum. Percept. Perform. *21*, 628–634.
13. Freire, A., Lee, K., and Symons, L.A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. Perception *29*, 159–170.
14. Leder, H., and Bruce, V. (1998). Local and relational aspects of face distinctiveness. Q. J. Exp. Psychol. *51*, 449–473.
15. Maurer, D., Grand, R.L., and Mondloch, C.J. (2002). The many faces of configural processing. Trends Cogn. Sci. *6*, 255–260.
16. Sekuler, A.B., Gaspar, C.M., Gold, J.M., and Bennett, P.J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. Curr. Biol. *14*, 391–396.
17. Shiffrar, M., Lichtey, L., and Chatterjee, S.H. (1997). The perception of biological motion across apertures. Percept. Psychophys. *59*, 51–59.
18. Hecht, H., Kaiser, M.K., and Banks, M.S. (1996). Gravitational acceleration as a cue for absolute size and distance? Percept. Psychophys. *58*, 1066–1075.
19. Jokisch, D., and Troje, N.F. (2003). Biological motion as a cue for the perception of size. J. Vis. *3*, 252–264.
20. McConnell, D.S., Muchisky, M.M., and Bingham, G.P. (1998). The use of time and trajectory forms as visual information about spatial scale in events. Percept. Psychophys. *60*, 1175–1187.
21. Runeson, S., and Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. J. Exp. Psychol. Gen. *112*, 585–615.
22. Bingham, G.P. (1987). Kinematic form and scaling: Further investigations on the visual perception of lifted weight. J. Exp. Psychol. Hum. Percept. Perform. *13*, 155–177.
23. Bingham, G.P. (1993). Scaling judgements of lifting weights: Lifter size and the role of the standard. Ecol. Psychol. *5*, 31–64.
24. Knoblich, G., and Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. Psychol. Sci. *12*, 467–472.
25. Bertenthal, B.I., Proffitt, D.R., and Cutting, J.E. (1984). Infant sensitivity to figural coherence in biomechanical motions. J. Exp. Child Psychol. *37*, 213–230.
26. Vallortigara, G., Regolin, L., and Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. PLoS Biol. *3*, e208.
27. Vallortigara, G., and Regolin, L. (2006). Natural-born physicists: A gravity bias in the interpretation of biological motion in visually-inexperienced chicks. Curr. Biol. *16*, this issue, R279–R280.
28. Troje, N.F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. J. Vis. *2*, 371–387.