

## Research Report

### THE EFFECT OF OBJECT AND EVENT ORIENTATION ON PERCEPTION OF BIOLOGICAL MOTION

Thomas F. Shipley

Temple University

**Abstract**—*Detection and recognition of point-light walking is reduced when the display is inverted, or turned upside down. This indicates that past experience influences biological motion perception. The effect could be the result of either presenting the human form in a novel orientation or presenting the event of walking in a novel orientation, as the two are confounded in the case of walking on feet. This study teased apart the effects of object and event orientation by examining detection accuracy for upright and inverted displays of a point-light figure walking on his hands. Detection of this walker was greater in the upright display, which had a familiar event orientation and an unfamiliar object orientation, than in the inverted display, which had a familiar object orientation and an unfamiliar event orientation. This finding supports accounts of event perception and recognition that are based on spatiotemporal patterns of motion associated with the dynamics of an event.*

In point-light displays of human actions, in which only points marking the location of the joints are visible, both events and actors are recognizable. A swirling set of 26 points can quickly resolve itself into a couple dancing a polka. The phenomenal ease of this perceptual achievement belies the sparse structural details in these displays. Each static frame may appear as a constellation of points, bearing no relationship to a human, but when even a few frames are presented in sequence, a robust form is evident (Johansson, 1973). Interest in this phenomenon may reflect the contrast between the richness of what can be seen and the poverty of the static stimulus. The displays provide sufficient information to distinguish characteristics of the actor, including his or her identity and sex (Cutting & Kozlowski, 1977; Kozlowski & Cutting, 1977), and of the action (e.g., observers can identify how far an invisible object has been thrown; Runeson & Frykholm, 1983).

Explanations of how observers extract form and action from these displays fall into two classes: *event-from-form* and *event-from-dynamics* theories. Event-from-form explanations suggest that visual processes first extract form, or parts of the form, and then determine the action. According to this perspective, in a point-light walker display, motions of elements serve to segment the form; recognition of the form (e.g., a human) allows identification of the event from the pose (e.g., dancing). There is neurological and behavioral evidence that form information is sufficient for event perception. Patients classified as “motion blind” are nevertheless able to recognize events in human point-light displays (Vaina, LeMay, Bienfang, Choi, & Nakayama, 1990). Furthermore, Beintema and Lappe (2002) have shown that location information is sufficient to discriminate the direction of walking when motion is made uninformative by varying, from one frame to the next, the number and location of the dots on a point-light walker.

Address correspondence to Thomas F. Shipley, 1701 N. 13th St., Department of Psychology, Temple University, Philadelphia, PA 19122; e-mail: tshipley@temple.edu.

They proposed that recognition is achieved with a dynamic form template that can integrate form information over time, while allowing for postural changes. This idea has proven computationally tractable; a number of models that use the location of a form over time to recognize an event have been developed (for a review, see Shah & Jain, 1999). However, simple template matches are unlikely to provide a full account of human recognition. Hill and Pollick (2000) have shown that temporal deviation from average values in motion sequences is used to recognize actions. A recent model (Giese & Poggio, 2000) avoids templates, using instead the relative orientation of limb segments from a few prototypical image sequences.

The event-from-dynamics approach is based on spatiotemporal information for action. Runeson and Frykholm (1983) argued that the most useful information for action is information about the dynamics, the forces acting on objects. Evidence for the pickup of dynamic information from kinematics, the pattern of motion of objects, comes from studies showing that observers could determine the actual weight an actor was lifting in point-light displays when the actor was pretending to lift a lighter or heavier weight. Subjects could identify the actual weights because it is impossible to mimic the acceleration pattern and posture associated with lifting any particular weight. Additional support comes from recent work by Troje (2002), who employed dynamically and structurally normalized displays and demonstrated that gender classification is accomplished primarily with dynamic information. A kinematic-specification-of-dynamics account would suggest that recognition of walking requires recognition of the forces present when an object walks. The motions of walking limbs will approximate the motions of nested pendulums; during walking, energy is stored as gravitational potential energy (in contrast, during running, energy is stored in the stretching of connective tissues and muscles).

In both types of theories, experiences with specific events are necessary for perception. Research on orientation effects in biological motion perception (e.g., Bertenthal, 1993; Johansson, 1973; and Pinto, 1997) has led to the same conclusion. A fully visible walking human turned upside down appears to have an abnormal gait—to “hop, dance, and shake” (Krüger, 1939, cited in Kohler, 1964, p. 95). Recognition of a human is reduced, although not eliminated, when a point-light walker is shown upside down (Pavlova & Sokolov, 2000; Sumi, 1984). Detection of point-light walkers in noise is reduced by turning the walker upside down (Bertenthal & Pinto, 1994; Pavlova & Sokolov, 2000). Inversion also reduces discrimination, for observers 5 months and older, between normal motions and phase-shifted motions, in which the temporal coordination between points has been randomized (e.g., the elbow might be swinging forward as the wrist is swinging backward; Bertenthal, 1993). Finally, Grossman and Blake (2001) found that upright point-light displays activated a region of the human posterior superior temporal sulcus approximately twice as much as the same displays shown upside down.

The two classes of accounts for event perception offer contrasting explanations for these effects. According to event-from-form models, an inverted form would not be detected as well as the original form be-

## Event Orientation and Perception of Biological Motion

cause the pieces would not appear in their familiar relative locations (the spatial fit between image and template, or prototype, would be poor), so the effects are the result of difficulty recognizing a person in an unfamiliar orientation. Alternatively, according to event-from-dynamics models, the effects may reflect difficulty recognizing the action because the dynamic relations specified by the kinematics are unfamiliar. In inverted displays, the normal pattern of acceleration present in the limbs as they are pulled away from and then fall back toward earth is not present.

The present experiment tested the relative importance of form and dynamics familiarity by using displays of an upside-down walker, without altering gravitational relations. In these displays, an actor walked on his hands. If observers are sensitive to the dynamics of action, inverting a person walking on his or her hands should interfere with perception. Alternatively, because the inversion presents a human in its familiar orientation, if form recognition is critical, perception should be facilitated.

A signal detection task was employed because it offers an objective measure of grouping. Detecting a walker in noise requires more than a single point; observers use the relations among multiple points (Bertenthal & Pinto, 1994). Four types of displays were used: walking on feet shown right side up (actor and gravity in a familiar orientation), walking on feet shown upside down (actor and gravity in an unfamiliar orientation), walking on hands shown right side up (actor in an unfamiliar orientation and gravity in a familiar orientation), and walking on hands shown upside down (actor in a familiar orientation and gravity in an unfamiliar orientation).

## METHOD

### Subjects

Twenty-three Temple University undergraduates participated in partial fulfillment of requirements for an introductory psychology course.

### Stimuli

The stimuli were animated point-light displays based on videotape recordings of a human walking 2.5 m on his hands or feet. So that the general position of the limbs would be equated, in both types of walking his hands were extended above his head. A walkway was built with a low ceiling so the walker could touch the floor and ceiling simultaneously. When the walker walked on his feet, his hands traversed the ceiling as if they were walking: As the body moved, the arms rotated until a hand needed to be lifted; that arm then swung to bring the hand to its new location. Similarly, when the walker walked on his hands, his feet traversed the ceiling as if they were walking. Four sequences of each type of walking were recorded. The average duration of the sequences was 2.7 s. The average numbers of steps walked on feet and on hands were 2.5 and 3.5, respectively.

The video sequences were converted into point-light displays by projecting each frame onto a digitizing tablet and recording the  $x$ ,  $y$  coordinates of all visible joints. Each walker was defined by a set of 13 points representing the shoulders, elbows, wrists, hips, knees, ankles, and head. Because of occlusion by another part of the body, not all points were visible on any given frame. Noise was added to displays in the form of additional points that moved with the same motions and disappeared at the same intervals as the walker's points, but

were displaced in random directions and phases. The spatial displacement averaged 10% of the average vertical extent of the walker, and ranged up to 20%. Each noise point's motion phase (i.e., when the point moved relative to the motions of the other masking points) was also randomized so that the phase relations were uniformly distributed. Additional random-phase points were added to the target-absent trials so that their density matched that of the target-present trials. For the upside-down displays, the noise was based on the upside-down target.

Two levels of masking were employed (Cohen, 2001). Either one or two noise points were added for each signal point present in the display.

### Procedure

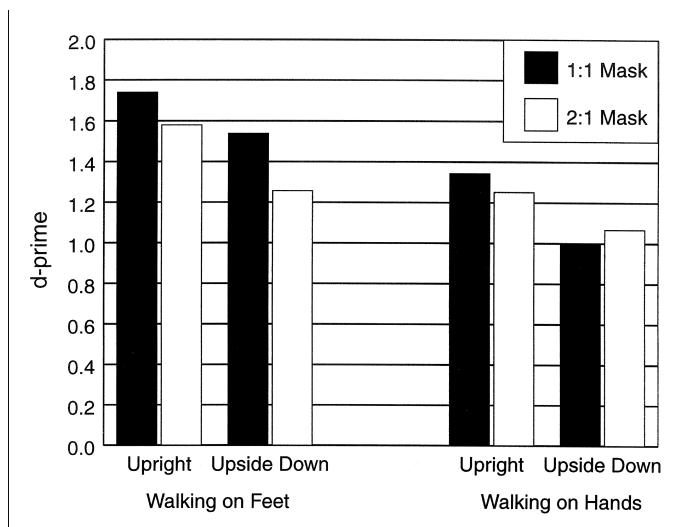
Subjects were first shown each mode of locomotion once without noise for an average of 21 s (eight repetitions of the sequence). Subjects were then told their task was to determine whether or not a moving form was present in displays that were like those just seen but had additional moving points. Each test trial was shown for an average of 8 s (three repetitions of the sequence). Orientation was blocked for each subject. Half the subjects were shown the displays in the normal orientation first, and half saw the upside-down displays first. For each orientation, subjects saw 96 trials (2 modes of walking  $\times$  4 examples of each mode  $\times$  3 repetitions of each example  $\times$  2 levels of masking  $\times$  2 levels of trial type: present/absent) presented in a random order.

## RESULTS

Accuracies for the two modes of walking with each orientation are shown in Figure 1. Turning both types of walker upside down reduced detection. The effect of inversion on detecting walking on feet replicates previous findings (Bertenthal & Pinto, 1994). The effect of inversion on detecting walking on hands indicates that the orientation of gravity, not form, is important for detection. There was no reliable effect of the order of presentation, so the data were entered into a three-way analysis of variance with orientation, mode of walking, and masking density as factors. There was a main effect of orientation, with upright walkers more detectable than upside-down walkers,  $F(1, 22) = 7.01, p < .02$ . The average  $d'$ s for upright and upside-down displays were 1.48 and 1.21, respectively. There was also a main effect of mode of walking, with walking on feet more detectable than walking on hands,  $F(1, 22) = 10.60, p < .01$ . The average  $d'$ s for walking on feet and hands were 1.53 and 1.16, respectively. The effect of masking density was not significant,  $F(1, 22) = 2.43, p > .1$ . It is not clear why density had a small effect in this study.

None of the higher-order interactions were significant, all  $F$ s(1, 22)  $< 1.4, p$ s  $> .2$ . Most important, accuracy did not increase when the point-light walker on hands was turned upside down (so that the head was up). In fact, subjects were reliably less accurate at detecting walking on hands when the display was turned upside down than when it was upright,  $F(1, 22) = 4.81, p < .05$ .

The important finding here involves the effect of orientation on detection of a point-light figure walking on hands. Nevertheless, one may wonder why overall accuracy was lower for walking on hands than for walking on feet. One possibility is that the subjects were less familiar with this action. Alternatively, the quality of the two actions may have differed. For the actor, walking on hands, a less frequent action, may not have been as smooth and coordinated as walking on feet.



**Fig. 1.** Mean accuracy detecting upright and inverted point-light figures walking on hands (right) or feet (left). Results are shown separately for the 1:1 and 2:1 levels of masking.

There is no accepted model of how the information available in these displays is used to detect human form and action, so there is no broadly accepted way to compare the inherent quality, or detectability, of different actions (see Troje, 2002, for an analysis of this issue).

## DISCUSSION

When a display of point-light walking on feet is turned upside down, both the orientation of the human and the dynamic relations are unfamiliar. By placing the familiarity of the figure's orientation and the dynamic relations in competition, this experiment showed that unfamiliarity of the dynamic relations is an important factor in the reduced grouping and recognition associated with upside-down displays. The results support approaches to event perception that include dynamic information, but do not mean that form information is not involved.

Form information may serve to create an initial representation of form. In the case of moving objects, dynamic information may serve to modify the representation. Recent reports suggest form representations are not necessarily static; priming work by Kourtzi and Shiffrar (2001) indicates that perceiving nonrigid motion involves a continuous transformation of a representation of the object's form. Such transformations may also serve as the basis for remembering events. The idea that motion information and perhaps dynamic information are used for segmentation is supported by studies in which static noise prevented biological motion recognition in a motion-blind patient (McLeod, Dittrich, Driver, Perrett, & Zihl, 1996) and dorsal extrastriate activity increased during search for biological motion (Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001).

Critical questions remain: How is dynamic information detected, and how is it used to transform object representations? In the domain of human actions, the answer to both questions may reside in the observer's ability to act. The visual system may use the complex pattern-generation ability of the motor cortex to help segment and group the

motions produced by animate objects (e.g., Thornton, Pinto, & Shiffrar, 1998). Evidence for a motor theory of biological motion perception, analogous to the motor theories of speech perception, comes from research relating perceptual reports to activity in the motor cortex (Decety et al., 1997; Stevens, Fonlupt, Shiffrar, & Decety, 2000; Vaina et al., 2001), and psychophysical findings that the inversion effect is present for point-light animals only if they have humanlike gaits (Cohen, 2001; Pinto, Zhao, & Shiffrar, 1997; Shipley & Cohen, 1999). The same brain structures that generate action may serve to modify representations of another person's action.

Finally, although an approach based on dynamic representations involving the motor system may be useful for human actions, and even some inanimate events that share dynamics with human action, perceiving complex, long-duration events may require different approaches. Recognition of large-scale events may require segmenting the whole into small-scale event parts (Newtson, 1973; Zacks, Tversky, & Iyer, 2001) and spanning large temporal gaps with content-limited memory mechanisms (Verfaillie, De Troy, & Van Rensbergen, 1994). Event perception, so conceived, is a perceptual cycle in which perception is a complex, evolving interplay among memory, attention, and information pickup (Neisser, 1976).

**Acknowledgments**—This research was supported in part by National Eye Institute Grant RO1 EY13518. The author thanks Maggie Shiffrar for helpful discussions about this experiment, and Robert Weisberg, James Enns, Allison Sekuler, and an anonymous reviewer for helpful comments on an earlier draft of the manuscript.

## REFERENCES

- Beintema, J.A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences, USA*, *99*, 5661–5663.
- Bertenthal, B.I. (1993). Infants' perception of biomechanical motions: Intrinsic image and knowledge-based constraints. In C. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 175–214). Hillsdale, NJ: Erlbaum.
- Bertenthal, B.I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*, *5*, 221–225.
- Cohen, L.R. (2001). *The role of experience in the perception of biological motion*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.
- Cutting, J.E., & Kozlowski, L.T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, *9*, 353–356.
- Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain*, *120*, 1763–1777.
- Giese, M.A., & Poggio, T. (2000). Morphable models for the analysis and synthesis of complex motion pattern. *International Journal of Computer Vision*, *38*(1), 59–73.
- Grossman, E.D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research*, *41*, 1475–1482.
- Hill, H., & Pollick, F.E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science*, *11*, 223–228.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.
- Kohler, I. (1964). The formation and transformation of the perceptual world. *Psychological Issues*, *3*(4, Monograph No. 12).
- Kourtzi, Z., & Shiffrar, M. (2001). Visual representation of malleable and rigid objects that deform as they rotate. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 335–355.
- Kozlowski, L.T., & Cutting, J.E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*, 575–580.
- McLeod, P., Dittrich, W., Driver, J., Perrett, D., & Zihl, J. (1996). Preserved and impaired detection of structure from motion by a "motion blind" patient. *Visual Cognition*, *3*, 363–391.
- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W.H. Freeman.
- Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*(1), 28–38.
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*, 889–899.

## Event Orientation and Perception of Biological Motion

- Pinto, J. (1997). Developmental changes in infants' perception of point-light displays of human gait. (Doctoral dissertation, University of Virginia, 1996). *Dissertation Abstracts International*, 57(8-B), 5365.
- Pinto, J., Zhao, Z., & Shiffrar, M. (1997). What is "biological motion"? Part 2: Generalization to non-human forms. *Investigative Ophthalmology and Visual Science*, 38, 357.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112, 585–615.
- Shah, M., & Jain, R. (1999). *Motion-based recognition*. Norwell, MA: Kluwer Academic.
- Shipley, T.F., & Cohen, L.R. (1999). The role of experience in perception of point-light walker displays. *Abstracts of the Psychonomic Society*, 4, 4.
- Stevens, J.A., Fonlupt, P., Shiffrar, M., & Decety, J. (2000). New aspects of motion perception: Selective neural encoding of apparent human movements. *NeuroReport*, 11(1), 109–115.
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, 13, 283–286.
- Thornton, I.M., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, 15, 535–552.
- Troje, N.F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371–387.
- Vaina, L.M., LeMay, M., Bienfang, D.C., Choi, A.Y., & Nakayama, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. *Visual Neuroscience*, 5, 353–369.
- Vaina, L.M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J.W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences, USA*, 98, 11656–11661.
- Verfaillie, K., De Troy, A., & Van Rensbergen, J. (1994). Transsaccadic integration of biological motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 649–670.
- Zacks, J.M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, 130, 29–58.

(RECEIVED 6/24/02; REVISION ACCEPTED 10/22/02)