Disorganizing biological motion

Amelia R. Hunt

Department of Psychology, Harvard University, Cambridge, MA, USA

Fred Halper

Department of Psychology, Harvard University, Cambridge, MA, USA, & Psychology Department, Essex County College, Newark, NJ, USA

The rapid and seemingly effortless organization of visually impoverished point-light displays of humans walking is often held up as a compelling example of the perception of form from motion. Here we show that motion information is not sufficient for the impression of a human walker to be extracted from a point-light display. We manipulated the 13 small dots out of which the typical point-light walker is constructed. Attempts to use size, color, or shape changes to disrupt walker perception had only modest impact on its robustness. But when all the local elements of the walker were replaced with complex unique objects, perception of the walker was severely disrupted. Of seventy-seven naive observers presented with this array for one full gait cycle, none perceived a human walker, even though movement paths were unchanged. We conclude that the spontaneous perception of a human walking is not an inevitable consequence of the motion of the points in a point-light walker but depends on the points themselves being relatively simple and uniform.

Keywords: biological motion, attention, perceptual organization, object recognition


Introduction

In 1973, Johansson filmed a person walking in the dark, illuminated only by lights affixed to the joints, and discovered that naive subjects easily recognize the film clip as a human form. Recognition occurred even at exposure durations as short as 200 ms (Johansson, 1976). The swift and spontaneous recognition of a walking person in a constellation of moving dots has since become a well known phenomenon that exemplifies our ability to extract complex information from impoverished visual input (for recent reviews, see Blake & Shiffrar, 2007; Thornton, 2006). For example, Kozlowski and Cutting (1977) found that observers can accurately discriminate gender based on point-light displays. Other studies have uncovered the wide range of actions in addition to walking that can be perceived from point lights (e.g., Dittrich, 1993) and the recognition of intention or mood from gestures shown only in point lights (e.g., Clarke, Bradshaw, Field, Hampson, & Rose, 2005; Runeson & Frykholm, 1983). Together with work demonstrating the presence of the ability at a young age (Fox & McDaniels, 1982; and for a review, see Pinto, 2006) and the obligatory processing of task-irrelevant point-light walkers (Thornton & Vuong, 2004), there is strong evidence that a rich representation of human form can be extracted from point-lights relatively quickly and automatically. Indeed, a recent influential neural model of biological motion perception (Giese & Poggio, 2003) is based entirely on feed-forward processes and explicitly states that attentional, top-down processes are unnecessary.

Although there seem to be few constraints on the ability to recognize point-light walkers, one notable exception is inversion. A naive subject is less likely to perceive a walker displayed in an inverted state, even when the subject is aware that the display is inverted (Pavlova & Sokolov, 2000). More subtle discriminations like gender (Barclay, Cutting, & Kozlowski, 1978) or detection of an inverted walker in a noisy display (Bertenthal & Pinto, 1994) are also impaired. The inversion effect suggests that global configuration information is important for perception of the walker (but see Troje & Westhoff, 2006). Thus, it is not surprising, though perhaps not fully appreciated, that Johansson’s early Gestaltist roots may have guided his subsequent invention of point-light walkers. He did tell William Epstein in an interview (1994) of his affinity for common fate, Max Wertheimer’s only nonstatic principle. This came about as a result of the influence of his Gestalt mentor, David Katz.

Another important influence in perception of the point-light walker is attention. Thornton, Rensink, and Shiffrar (2002) and Cavanagh, Labianca, and Thornton (2001) demonstrated that focused attention is needed to group local moving elements into a global interpretation. In the 2001 study by Cavanagh and colleagues, for instance, observers fixated the center of an array while searching the periphery for a form walking to the right among
forms walking to the left (all forms remained stationary, as though walking on a treadmill). They found that reaction time to detect the left-walking form increased by about 200 ms for each additional rightward walker added to the display. This result suggests that attention had to be directed to each walker before a decision was made, consistent with evidence from visual search experiments demonstrating that focal attention is needed to group local elements into a global form (Enns & Kingstone, 1994). However, the ~200-ms per-item increase in RT observed by Cavanagh et al. (2001) was considerably shorter than the over 900-ms search slope required to detect a simpler yet less familiar configuration of moving dots. This result suggests that attention is required for grouping the dots into a global configuration, but that familiarity with the stimulus makes this grouping process more efficient.

This prior work demonstrating the importance of both global orientation and focused attention in putatively “effortless” biological motion perception led us to examine conditions that might lead to a disruption in the detection of a walker. In particular, the Gestalt principles of grouping based on similarity and phenomenal identity for a subset of the dots might provide conditions where the perception of local groups or motion information could overpower the global form, resulting in an inability to perceive a walker even among subjects who are familiar with the stimulus configuration. We explored this idea by replacing the “point-lights” (that is, simple and uniform dots that make up the typical point-light walker) with a variety of items. If the overall motion pattern alone is sufficient for the perception of a human form in the point-light array, then manipulation of the individual elements should have little impact, if any, on the degree to which the configuration appears to be a human walking.

Despite the vast literature on point-light walkers (the 1973 paper by Johansson has been cited over 700 times to date), few studies have manipulated the dot elements themselves to measure the impact on the walker. In one such study, Tadin, Lappin, Blake, and Grossman (2002) replaced the points of a point-light walker with Gabor patches. Visual comparison across the patches was facilitated when the patches were in a walker configuration relative to when the items were in a configuration that was harder to organize into a coherent percept. This demonstration of the influence of global organization on local processing represents an interesting complement to the current study because we examine the effect of local elements on the global organization of the walker. Another example of a study manipulating the point-light elements is Barclay et al. (1978), who showed that blurring and magnification of individual elements reduced gender recognition to chance probability, although walker perception itself remained robust. Based on this result, we could predict that manipulation of the size and shape of walker elements may degrade, though perhaps not eliminate, perception of the walker.

### Ratings of modified point-light walkers

Twelve observers were shown a walker movie for a full gait cycle, comprised of 42 frames at 12 frames per second. The 13 dots in each walker were replaced by a variety of elements, described below. The walker frames were altered using Adobe Flash MX and played back to the subjects using Flash Player. Viewing distance was not controlled. In all cases, the walker was shown on a white background. In none of the displays were the temporal properties or motion vectors of the standard walker altered. After each presentation, observers were asked to “Rate the extent to which the display appears to be a human walking,” on a scale of 1 to 7, with 1 “not at all” and 7 “completely.” Observers were recruited from the Vision Lab at Harvard University and were all familiar with biological motion displays. Each completed ratings of 10 walkers in a randomly generated order. The walkers were oriented either to the left or to the right and walked in place. The direction was randomly selected.

The walkers tested, and their ratings are listed in Table 1, which links to demonstrations of each walker. Our standard baseline display (Walker 1) was similar to the standard Cutting walker (1978), except that the dots were black on a white background. Observers rated this display as very similar to a human walking (mean rating was 6.8 out of a possible 7). We attempted to form strong local groups based on color to interfere with the overall organization of dots into a global form (Walker 2). We also attempted to interfere with the interpretation of the global motion by adding apparent local motion between the items (Walker 3, which has an annulus moving from point to point) and motion within the items (Walker 8, which introduces looming motion to the dots). We replaced the dots with unique objects in Walker 10. This

#### Table 1

<table>
<thead>
<tr>
<th>Walker</th>
<th>Mean rating out of 7 (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Standard</td>
<td>6.8 (0.13)</td>
</tr>
<tr>
<td>2. Bicolor dots</td>
<td>6.0 (0.30)</td>
</tr>
<tr>
<td>3. With annulus</td>
<td>5.5 (0.40)</td>
</tr>
<tr>
<td>4. Same object</td>
<td>5.3 (0.30)</td>
</tr>
<tr>
<td>5. Large dots</td>
<td>5.1 (0.43)</td>
</tr>
<tr>
<td>6. Unique letters</td>
<td>4.9 (0.39)</td>
</tr>
<tr>
<td>7. Same letter</td>
<td>4.8 (0.41)</td>
</tr>
<tr>
<td>8. Looming dots</td>
<td>4.8 (0.44)</td>
</tr>
<tr>
<td>9. Unique faces</td>
<td>2.8 (0.47)</td>
</tr>
<tr>
<td>10. Unique objects</td>
<td>2.6 (0.56)</td>
</tr>
</tbody>
</table>

Table 1. Ratings of the degree to which the displays resemble a person. Lines indicate conditions across which ratings do not differ significantly (p > .05). Each walker name is hyperlinked to an animated walker; click on it for a demonstration.
array is comprised of full-color photographs of objects, such as a computer mouse, an artichoke, and a spatula (see Figure 1). We also replaced the dots with full-color photographs of faces (Walker 9), letters (Walker 6), uniform objects (a clock, Walker 4), and uniform letters (an A, Walker 7). As a control, we also tested the standard walker with dots enlarged (Walker 5) to match in average size of the objects in Walker 10. The manipulation of the elements has a significant effect on ratings [repeated-measures ANOVA $F(9,81) = 16.93$]. A post hoc analysis using Tukey/Kramer critical values reveals 3 groups of walkers that differ from each other at an alpha level of .05 (within each group there are no significant differences). The first group contains Walkers 1–3 and shows that manipulations of color (Walker 2) and motion (Walker 3) do not lower ratings significantly relative to the standard walker (Walker 1). The second group contains Walkers 3–8, all of which differ significantly from the standard walker, but not from each other. It is important to note that this group contains the walker with black dots enlarged (Walker 5). The other walkers within this group do not lower ratings relative to the standard significantly more than a simple enlargement of the dots. Included in this group are uniform and unique letter walkers (Walkers 6 and 7), the walker with complex but uniform clocks (Walker 4), and the walker with looming motion (Walker 8). The third and final group contains only Walkers 9 and 10, which are the walkers with the elements replaced by faces and by unique objects. These walkers differ significantly from both the standard with small dots (#1) and with enlarged dots (#5).

In summary, simple enlargement of the dots has a significant impact on the ratings of the walker relative to the standard. Over and above that effect, replacing the elements with complex and unique objects or with faces produces a further drop in ratings of the resemblance of the display to a person walking.

Naive observer descriptions

Our goal in this experiment was to determine if manipulations of the dot elements alone could prevent subjects from making the spontaneous inference that they are observing a human walking. To that end, we collected free reports from 45 naive observers, who viewed the configuration that was rated as looking the least like a walker by our first set of observers (Walker 10, comprised of unique objects, see Figure 1 for a static representation, or click here for an animated demonstration). Eight of the observers viewed the stimulus posted on a local Web site and filled out a Web form, and 37 additional observers, who were enrolled in a psychology class at Harvard College, viewed it as a group and wrote their observations on paper. Both groups were first shown the walker for a full gait cycle at 12 frames per second and then responded to two questions: (1) “Describe your memory of what you saw in as much detail as possible;” and (2) “Have you seen a similar sort of display before?”

All 45 observers reported “objects” or “items” in their observations. Most subjects (32/45) described the motion of the objects, often in detail, making some attempt to interpret the motion pattern (for example, “moving in concert, much like a mobile;” “like a Chinese swinging bell;” “in a rotating column-like shape”). Despite these clear attempts to organize and describe the motion globally, none of the 45 subjects who viewed the display described a human form. When the 37 observers who viewed the item with the experimenter present were shown the display a second time and directly asked if they could see a person in the display, most confirmed that they were able to recognize the human form from the motion once they knew it was there.

A possible concern is that the object walker was presented at a rate that was slower than an average
walking speed. Unfamiliar walking speeds can impair discrimination of the identity of a point-light walker (Jacobs, Pinto, & Shiffrar, 2004). To be confident in our conclusions, we therefore recruited a second group of 32 naive subjects who were enrolled in a psychology class at Essex County College. This group of subjects viewed the object walker for a full gait cycle (2 seconds) at 20 frames per second instead of 12 (click here for a demonstration). They responded to the same two questions as the first group of subjects. The results were the same as before: none of the 32 new subjects reported seeing a walker in the display. Again, most of the subjects (21/32) mentioned the motion of the objects in their description of the display (e.g., “moving around in no special pattern;” “rotating in a circular motion”).

Out of a total of 77 subjects exposed to the walker, not one person reported the presence of a walker. We are therefore confident that spontaneous recognition of the walker was eliminated when the point-lights were replaced with objects.

Conclusions

We have demonstrated that simple and uniform local elements are a necessary condition for a human form to be perceived in a point-light display. Observers rated a walker made from dots as looking very much like a person walking. A walker made up of uniform but complex objects (clocks) did not disrupt walker perception relative to black dots of a similar size, suggesting complexity alone is not sufficient for disruption to occur. Similarly, a walker constructed from unique letters did not result in lower ratings relative to a walker constructed from uniform letters, suggesting unique identity alone is also not sufficient for disruption. However, when we replaced the dots with unique objects or faces, which were both complex and unique, the configuration was rated as looking very little like a person. In the second experiment, 77 naive observers were shown only the walker constructed from complex and unique objects, and none reported seeing a human form. The finding suggests that the presence of simple, uniform elements is critical to the perception of a person in a point-light display.

The results place an important constraint on theories of how form information is organized in a biological motion display. Human form perception is not an automatic consequence of the mere presence of consistent motion information. The results instead support an interpretation of biological motion perception as an active, top-down process, through which the visual system puts together the most “pregnant” and conceptually simplest organization. This conclusion presents a challenge to any model of biological recognition based in completely feed-forward processes. A recent model of biological motion recognition laid out by Giese and Poggio (2003; and further developed by Casile & Giese, 2005), for example, asserts that biological motion recognition does not depend on attention or top-down processes, though they note that top-down factors may be necessary for biological motion when the stimuli are cluttered or noisy obscuring the motion of the stimulus. In our walker made from objects, motion is unchanged, and recognizing the human form in the configuration is relatively easy once the observer is aware of this interpretation. A naive observer, however, may never conclude that they are observing a human form even after prolonged exposure. This finding therefore represents compelling evidence that some top-down mechanism imposes an interpretation of the elements as a representation of a human form.

Our interpretation of the current finding is that individual elements that are complex and unique prevent the observer from forming a perceptual hypothesis about the global motion of the stimulus. This occurs because limited processing resources must be shared between identifying and tracking the individual objects at a local level and grouping the individual moving elements into a human form at a global level. When the elements are complex but uniform (such as the walker made entirely out of clocks) or simple but unique (such as the walker made from letters), processing them is no longer so demanding as to prevent observers from linking the motion information from spatially distinct locations to form the percept of a walking person. But when the local elements are both complex and unique, local processing wins out, and the global information can only be organized when the human character of the global form is known. This interpretation suggests that organization of biological motion depends on allocation of limited processing resources to the global motion information and is therefore consistent with previous evidence demonstrating a role for serial, attention-based processes in biological motion perception (e.g., Cavanagh et al., 2001; Thornton et al., 2002).

Several promising avenues for future research follow from this work. It would be interesting, for example, to further explore what kinds of elements will prevent naive observers from recognizing the motion pattern, as this could shed further light on the top-down mechanism underlying biological motion. It is also possible to better understand the role of attention by manipulating it directly; for example, we would predict that attending to local elements would interfere with a global walker discrimination task, and attending to the global elements would interfere with local discrimination. There may also be some asymmetry in the interference effects, similar to those observed in other hierarchical stimuli (e.g., Navon, 1977). Finally, changes in brain activity when a subject goes from seeing a collection of moving objects to seeing a walking person could provide insight into the neural mechanisms subserving biological motion perception.
Acknowledgments

This research was supported by a Natural Sciences and Engineering Council of Canada (NSERC) fellowship (A.R.H.). The authors are grateful to Patrick Cavanagh and his “Vision on the Fly” class for contributing feedback to this project, and to Brad Newman for Flash consulting.

Commercial relationships: none.
Corresponding author: Amelia Hunt.
Email: ahunt@wjh.harvard.edu.
Address: Harvard University Vision Sciences Lab, 33 Kirkland St., 7th Floor, Cambridge, MA 02138, USA.

References


