Detection of biological and nonbiological motion

Eric Hiris

Department of Psychology, St. Mary’s College of Maryland, St. Mary’s City, MD, USA

Often it is claimed that humans are particularly sensitive to biological motion. Here, sensitivity as a detection advantage for biological over nonbiological motion is examined. Previous studies comparing biological motion to nonbiological motion have not used appropriate masks or have not taken into account the underlying form present in biological motion. The studies reported here compare the detection of biological motion to nonbiological motion with and without form. Target animation sequences represented a walking human, an unstructured translation and rotation, and a structured translation and rotation. Both the number of mask dots and the size of the target varied across trials. The results show that biological motion is easier to detect than unstructured nonbiological motion but is not easier to detect than structured nonbiological motion. The results cannot be explained by learning over the course of data collection. Additional analyses show that mask density explains masking of unstructured nonbiological motion but is not easier to detect than structured nonbiological motion. The results cannot be explained by learning over the course of data collection. Additional analyses show that mask density explains masking of different size target areas and is not specific to detection tasks. These data show that humans are not better at detecting biological motion compared to nonbiological motion in a mask. Any differences in detection performance between biological motion and nonbiological motion may be in part because biological motion always contains an underlying form.

Keywords: biological motion, form, masking, motion perception, point-light motion


Introduction

Biological motion is the movement of living organisms and commonly has been studied by displaying points of light representing the major joints of the body. Johansson (1973, 1976) demonstrated that biological motion could be perceived accurately from such point-light motion. In the biological motion literature, it is common for papers to begin with a statement that humans are particularly sensitive to point-light biological motion. For example, “numerous laboratory studies have demonstrated that the human visual system is highly sensitive to biological motion even when visible cues are reduced to only a few moving points of light” (Thornton & Vuong, 2004). “Human perception of biological movement … is amazingly robust” (Casile & Giese, 2005). “Human observers are particularly sensitive to human movement” (Pinto & Shiffrar, 1999). A wealth of research shows that human observers are sensitive to the information contained within biological motion displays, including the sex of the actor (Barclay, Cutting, & Kozlowski, 1978; Hill & Johnston, 2001; Mather & Murdoch, 1994), the emotion portrayed (Atkinson, Dittrich, Gemmell, & Young, 2004; Brownlow, Dixon, Egbert, & Radcliffe, 1997; Dittrich, Troschianko, Lea, & Morgan, 1996; Pollick, Paterson, Bruderlin, & Sanford, 2001), the identity of the actor (Loulia, Prasad, Harber, & Shiffrar, 2005; Stevenage, Nixon, & Vince, 1999; Troje, Westhoff, & Lavrov, 2005), and the intentions of the actor (Runeson, Frykholm, 1983; Troje, 2002). Also, it is clear that human observers can perceive biological motion with very few points of light (Johansson, 1973, 1976; Troje & Westhoff, 2006), with limited local motion information (Beintema, Georg, & Lappe, 2006; Beintema & Lappe, 2002), and/or degraded by masks (Cutting, Moore, & Morrison, 1988). However, little research has been performed addressing the relative sensitivity of observers to biological motion compared to nonbiological motion. The purpose of the current studies is to compare biological motion and nonbiological motion primarily based on one type of “sensitivity,” detection performance. To do so, several factors about biological and nonbiological motion must be taken into account.

Johansson (1973) claimed that by using points of light, one could “analyze human motion patterns without the interference of pictorial information.” Despite this claim, form may play an important role in perceiving point-light biological motion. Although it is true that most naive participants cannot recognize biological motion from a single static frame, nonnaive participants perform better than chance when presented with a single frame of a point-light motion sequence (Cutting et al., 1988; Hiris & Cramer, 2005). Biological motion may in fact be similar to other types of structure from motion displays. In structure from motion displays, the global form cannot be identified before the motion begins (Wallach & O’Connell, 1953). Under low light conditions, the perception of both biological and structure from motion were impaired, although coherent motion was not (Grossman & Blake, 1999). Also, it has been shown that the perception of biological motion and structure from motion can be impaired even when motion perception is impaired (Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990). Conversely, biological motion and structure from motion can be impaired although motion perception is normal (Cowey & Vaina, 2000). Together, these studies suggest that form
information is available in point-light displays and may play a role in perceiving point-light biological motion.

The role of form in perceiving biological motion is supported by the work of Beintema et al. (2006) and Beintema and Lappe (2002). They removed local motion from their sequential position point-light motion displays by using dots that changed location between the joints on every frame. Here local motion is no longer a reliable cue, yet detection of biological “motion” was not impaired. Their data suggest that biological motion might be processed as form information over time rather than from motion per se. In fact, models of biological motion perception indicate that recognition of biological motion from form over time is possible (Lange, Georg, & Lappe, 2006). The role of form in perceiving biological motion is supported further by recent fMRI research showing that sequential position displays activate form-processing areas more strongly than motion-processing areas (Michels, Lappe, & Vaina, 2005). Other fMRI research clearly indicates that even normal point-light displays activate form-processing areas and contribute to biological motion perception (see, for example, Peelen, Wiggott, & Downing, 2006; Peuskens, Vanrie, Verfaillie, & Orban, 2005; Thompson, Clarke, Stewart, & Puce, 2005).

Two studies have specifically compared biological motion to nonbiological motion in terms of detection performance. Neri, Morrone, and Burr (1998) showed that detection and discrimination performance were the same for translational motion but differ for biological motion. However, they used random motion masks for their conditions. Although these masks might be appropriate for their translational motion condition, random motion masks are not particularly effective in masking biological motion (Cutting et al., 1988). In fact, biological motion presented within a random motion mask can be processed relatively automatically (Thorton, Rensink, & Shiffrar, 2002). Neri et al. also failed to take into account that, in their study, biological motion had an underlying form whereas translation did not. In the second part of their study, Neri et al. showed that temporal summation occurred over a longer period for biological motion than translation. Also, temporal summation depended on the speed of motion for biological motion but not for translation. However, careful examination of Neri et al.’s Figure 3 also suggests two things: (1) The observer was equally sensitive to biological and translational motion at the longest display durations; and (2) extended presentation times are necessary for biological motion to be detected as easily as translational motion. Neither indicate that humans are better at detecting biological motion than nonbiological motion.

Poom and Olsson (2002) attempted to replicate and extend the findings of Neri et al. (1998). Poom and Olsson included conditions with rigid three-dimensional shapes in their study in an attempt to address how the lack of form in nonbiological motion might have influenced the results of Neri et al. Poom and Olsson found that discrimination performance depended on the speed of translation and could be either better or worse than biological motion. The inclusion of a three-dimensional form did not appear to make a difference in discrimination or temporal summation. However, the three-dimensional shapes were a spheroid and a single frame of biological motion rotating about its vertical center. Both of these structures have difficulties associated with them. A spheroid is not recognizable from a single frame of an animation sequence although a point-light display is recognizable. Also, a single frame of biological motion rigidly rotating does not contain any natural movements. In addition, Poom and Olsson used random motion masks in their studies, with the associated problems described above.

It is clear the biological and the nonbiological motion do differ in some respects. It has been recently shown that eccentric viewing disturbs biological motion perception (Ikeda, Blake, & Watanabe, 2005) much more than it disturbs simple motion detection and velocity perception (Johnston & Wright, 1986; Wright, 1987). This, however, suggests a disadvantage for biological motion compared to nonbiological motion processing in the periphery.

To summarize, previous research often claims that humans are particularly sensitive to biological motion. “Sensitivity” can mean anything from the ability to detect the information within biological motion to the ability to detect a degraded stimulus well. The current studies are focused on the latter issue by comparing performance on biological motion to nonbiological motion. Previous research in this area generally has failed to take into account form information in biological motion and/or has used masks that were less than optimal for biological motion. The goal of this study was to determine if humans are better at detecting biological motion compared to nonbiological motion. Given the importance of form in biological motion, biological motion was compared to nonbiological motion with and without an underlying form. Also, given the evidence that masks more similar to the biological motion target in terms of motion (Cutting et al., 1988; Hiris, Humphrey, & Stout, 2005) or form (Bertenthal & Pinto, 1994) result in a more effective mask, all masks for one set of observations were made by scrambling the target motions present in the display, thereby equating the effectiveness of the mask. Also, previous research has focused on manipulating the number of mask dots, whereas little is known about the explanatory power of the density of the mask dots relative to the target. To address this issue, targets of various sizes were presented within a constant-sized mask area to determine whether mask density predicted detection performance.

## Experiments

The current study compares biological motion to two types of nonbiological motion, translation and rotation. The nonbiological target motions were displayed either with or
without an underlying form present. Stimuli were presented with masks and detection performance was compared across conditions. The two main questions were whether observers were more sensitive to biological motion and whether adding form to nonbiological motion increased performance. Bellefeuille and Faubert (1998) had similar questions but approached them differently. They showed that the addition of biological motion to motion-defined animal forms did not increase performance. This suggests that biological motion does not add anything to performance beyond the form in their displays. However, does the opposite addition, adding form to nonbiological motion, increase performance? If form makes it easier to detect both biological and nonbiological motion, then one would expect biological motion to be easier to detect than nonbiological motion without form. However, biological motion and nonbiological motion with form should be equally easy to detect.

**Methods**

**Observers**

Four observers with normal or corrected-to-normal vision participated. Three observers were naïve and were paid US $135 each for their participation. The fourth observer was the author.

**Stimuli**

All stimuli consisted of 11 moving target dots representing biological, translational, or rotational motion. The target motions were centered on the screen except for a random 1.2° horizontal (H) and vertical (V) offset from trial to trial. The viewing distance was always 57 cm. All target stimuli were presented within a 26.6° H × 33.8° V mask area with a variable number of mask dots. There were 22, 44, 88, 176, 352, 704, or 1,408 mask dots randomly plotted within the mask area. This range of mask dots allowed observers to easily perceive large targets (given 22 mask dots) and made it difficult to perceive small targets (given 1,408 mask dots). Each dot in the displays subtended about 0.12° in visual angle. Each display consisted of 60 frames of motion. For biological motion, these 60 frames represented one-step cycle of walking. For translation, the 60 frames showed a rigid translation of 3.4°, 6.7°, or 13.2° for small, medium, and large stimuli, respectively. For rotation, these 60 frames showed a rotation of 90°. The displays were presented at a rate of 42.5 frames per second on a 22-in. Mitsubishi monitor (1,600 × 1,200 at 85 Hz) resulting in a 1.4-s long presentation.

There were two sets of three basic types of target motions used in this experiment. The first set of target motions consisted of translating biological motion based on Cutting (1978), a rigidly translating cloud of dots quasi-randomly placed on each trial within a rectangular area (unstructured translation), and rigidly translating set of dots defining the edges of a rectangle by plotting the dots at the same location on the edge of the target rectangle on each trial (structured translation). Unstructured translation dots were constrained so that at least one dot was positioned on each edge of the rectangular target motion area. Small, medium, and large stimuli translated 3.4°, 6.7°, or 13.2°, respectively, from right to left on each trial. Three sizes of each type of display were presented, small (3.05° H × 6.05° V), medium (6.1° H × 12.2° V), and large (12.2° H × 24.2° V). Three target motion sizes were used within a constant size mask area as a way of replicating the results within the experiment as well as a way of investigating the role of mask density in masking the target motion.

The target motions from the first set were embedded within random noise masks where each mask dot was assigned to move in a random direction with each direction being equally represented. Figures 1A, 1B, and 1C gives schematic representations of the first set of stimuli. Point-light biological motion displays began at a random point within their step-cycle from trial to trial.

The second set of target motions consisted of treadmill biological motion (translating biological motion with the common horizontal component of motion removed), a rigidly rotating cloud of dots randomly placed on each trial within a square area (unstructured rotation), and rigidly rotating set of dots defining the edges of a square by plotting the dots at the same point on the edge of the target square on each trial (structured rotation). Unstructured rotation dots were constrained so that at least one dot was positioned on each edge of the square target motion area. Rotation targets always rotated about the center of the square. The presentation could begin at any point during the step cycle or rotation. Three sizes of each type of display were presented. For biological motion, these sizes were small (3.05° H × 6.05° V), medium (6.1° H × 12.2° V), and large (12.2° H × 24.2° V). For unstructured and structured rotation, these sizes were small (4.3° H × 4.3° V), medium (8.6° H × 8.6° V), and large (17.2° H × 17.2° V). These values were chosen so that biological motion and rotation displays covered the same area for small, medium, and large displays. Note that for both rotation and translation, “unstructured” displays do have some structure (a rigid plane) although presumably this form is less detectable than the “structured” displays that have an identifiable shape (a square or a rectangle).

In this target motion set, point-light walkers were embedded within scrambled walker masks where each mask dot was assigned to move in the same way as one of the dots representing the walker. These assignments were made so that there were an equal number of mask dots that matched each walker dot. The unstructured and structured rotation stimuli were embedded within scrambled rotation masks where each mask dot was assigned to move in the same way as one of the dots representing the target rotation. Figures 1D, 1E, and 1F provide schematic representations of the second set of stimuli. Movies 1, 2, and 3 provide dynamic representations of these stimuli.
Figure 1. Schematics of the stimuli. Red dots indicate target dots and black dots indicate masking dots. All dots were white presented on a black background (with no walker outline) in the actual displays. All mask areas were 33.8° wide × 26.6° high. (A) Translating biological motion with a random motion mask. (B) Unstructured translation with a random motion mask. (C) Structured translation (rectangle) with a random motion mask. (D) Treadmill biological motion with a scrambled motion mask. (E) Unstructured rotation with a scrambled motion mask. (F) Structured rotation (square) with a scrambled motion mask.
In the second set of target dots, scrambled target masks were used in an attempt to equate the effectiveness of the mask dots. Cutting et al. (1988) showed that the more similar the motion of the mask dots are to the target dots, the more effective the mask is for biological motion targets. Presumably, the same principle applies for other motion targets as well. If random motion masks were used, one would expect differences in how effective the mask was for different targets depending on how similar the target motion is to the motions in the mask. By using scrambled target masks for all conditions, one can be certain that the local motion of the target and mask match perfectly. Therefore, for all conditions in the second set of target dots there were no local motion differences between the mask and the target.

In the experiment, a target was present on half of the trials. On trials where the target was absent, 11 additional mask dots were presented. To ensure that dot density was not a cue to the presence or the absence of the target, these mask dots were located where the target dots would have been present on the first frame of motion. These mask dots on target absent trials were then assigned motions just like the other mask dots for that trial.

Results and discussion

Two sets of comparisons were made. One set of comparisons was between translating biological motion, unstructured translation, and structured translation. For these displays, a random motion mask was used for two reasons: (1) Scrambled masks would have made the unstructured translation impossible to detect; (2) this allowed one set of comparisons to approximate the stimuli used by others (Neri et al., 1998; Poom & Olsson, 2002). The second set of comparisons was between treadmill biological motion, unstructured rotation, and structured rotation. In the second set of comparisons, masks made from scrambled target motions were always used.

Biological motion versus translation

All data were analyzed by repeated measures ANOVA (with Greenhouse–Geisser correction for sphericity) with target type (translating biological motion, unstructured translation, and structured translation) target size, and number of mask dots as the independent variables and sensitivity (d') as the dependent variable. The same independent variables were used to perform an analysis with criterion as the dependent variable.

The analysis of variance on sensitivity showed that all main effects and interactions were significant except for the two-way interaction between target type and target size. The results are shown in the top panels of Figure 2. Not surprisingly, as the number of mask dots increases, sensitivity decreases, $F(1.593, 4.78) = 263.738, p < .001$, partial $\eta^2 = .989$. Also, as target size increases, sensitivity decreases, $F(1.084, 3.252) = 485.953, p < .001$, partial $\eta^2 = .994$. This result will be addressed in detail in the section on mask density below. The significant interactions via key presses. Data from the practice sessions were not analyzed and observers were encouraged to ask any questions they had about their task or the procedures. Over several months, observers completed 30 runs of 2 sets of 210 trials (with a break between sets) for a total 12,600 trials per observer (6 target types, 3 target sizes, 7 mask densities, 100 trials per combination). No more than two data collection sessions were performed on any given day. The order of presentation was quasi-randomized within each set of trials. A given run of trials always consisted of only one of the target types (e.g., treadmill biological motion) presented at all 3 target sizes and all 7 mask densities. For these 21 conditions, 10 trials were performed for a given set (with two sets in a given run). Observers completed data collection on four of the target types (translating biological motion, treadmill biological motion, unstructured translation, and unstructured rotation) before beginning data collection on the remaining two target types (structured translation and structured rotation).

Procedure

Before each data collection, session observers were shown the small, the medium, and the large targets three times each with no masking dots to familiarize them with different size targets. Before each trial, a fixation point of one of three sizes appeared. The observers were told that the size of the fixation point indicated the size of the target for that trial. Observers then participated in 6 sets (one for each stimulus) of 42 practice trials (3 target sizes $\times$ 7 mask sizes $\times$ 2 repetitions). Observers judged whether the target was present within the mask dots and entered their responses

Movie 3. An example of a structured rotation target within a scrambled structured rotation mask.
can be explained by the different shapes of the curves in the top panels of Figure 2. Significant interactions including the target type variable may be in part attributed to the ceiling effect for translating biological motion in all but the most dense masks.

Most important for the present purposes, observers were most sensitive to translating biological motion, then structured translation, then unstructured translation, $F(1.097, 3.292) = 191.927, p \leq .001$, partial eta$^2 = .985$; verified by Bonferroni-corrected post hoc comparisons. These results do not fully support the prediction that the addition of form to nonbiological motion would make sensitivity equivalent to biological motion. However, detection of structured translation is substantially more similar to translating biological motion than unstructured translation. Also recall that the mask for all target motions was random motion for this set of target motions. A random motion mask is an excellent mask for translation, but is not a very effective mask for biological motion (Cutting et al., 1988). This difference between the quality of the mask for the target motion may explain why structured translation did not result in the same performance as translating biological motion. In the next section of results, data from the second set of target motions, in which all masks were created from scrambled target motions, will be examined.

The analysis of variance on criterion showed a significant three-way interaction between target type, target size, and mask, $F(2.544, 7.633) = 5.099, p < .05$, partial eta$^2 = .630$, and a significant main effect of target size, $F(1.016, 3.049) = 25.054, p < .05$, partial eta$^2 = .893$. As the bottom panels of Figure 2 show, criterion became more conservative as the size or the number of mask dots increased. In other words, the participants became less willing to say the target was present when the target became more difficult to see. Note that none of the pairwise comparisons for criterion between translating biological motion versus unstructured or structured translation were significant using a Bonferroni correction. The three-way interaction is explained by observers having different performance with the large unstructured translation condition compared to the large stimuli in the other two conditions.

### Biological motion versus rotation

In the second analyses of sensitivity and criterion, the target type variable consisted of treadmill biological motion, unstructured rotation, and structured rotation. The analysis
of variance on sensitivity showed that all main effects and interactions were significant except for the two-way interaction between target type and target size and the two-way interaction between target type and number of mask dots. In general, the results are similar to those in the previous section of results. The top panels of Figure 3 show that as target size increased, sensitivity decreased, $F(1.061, 3.183) = 1.297.256, p < .001$, partial eta$^2 = .998$, and as the number of mask dots increased, sensitivity decreased, $F(1.556, 4.669) = 769.548, p < .001$, partial eta$^2 = .996$. The significant interactions can be explained by the different shapes of the curves for unstructured rotation in the top middle panel of Figure 3.

Like the previous analysis, there was a significant effect of target type, $F(1.638, 4.193) = 34.188, p < .005$, partial eta$^2 = .919$. However, unlike the previous analysis, there was no evidence that detecting treadmill biological motion was different from detecting structured rotation (see the top right panel of Figure 3). All other pairwise comparisons were significant (as verified by Bonferroni-corrected post hoc comparisons). These results fully support the predictions; with an underlying form, nonbiological motion can be detected as easily as biological motion.

The analysis of variance on criterion showed that all main effects were significant, target type: $F(1.189, 3.566) = 10.631, p < .05$, partial eta$^2 = .780$; target size: $F(1.028, 3.083) = 12.780, p < .05$, partial eta$^2 = .810$; number of mask dots: $F(1.464, 4.392) = 14.189, p < .05$, partial eta$^2 = .825$. Also, the interaction between target type and target size, $F(1.763, 5.29) = 8.917, p < .05$, partial eta$^2 = .748$, and the interaction between target type and number of mask dots, $F(1.926, 5.779) = 6.158, p < .05$, partial eta$^2 = .672$, were significant. In general, this criterion analysis leads to conclusions similar to those reached for the criterion analysis in the previous section. Criterion became more conservative as the size or the number of mask dots increased. In other words, the participants became less willing to say the target was present when the target became more difficult to see. Also, the observers tended to have a more conservative criterion in the structured rotation condition compared to the unstructured rotation condition (as verified by post hoc comparisons with a Bonferroni correction).

The analyses of data from both sets of target motions suggest that random motion masks do not mask biological motion well. This results in ceiling effects in some conditions (see upper left panel of Figure 2). Also, these

![Figure 3](https://jov.arvojournals.org/pdfaccess.ashx?url=data/journals/jov/933520/) Results from treadmill biological motion, unstructured rotation, and structured rotation. The top row of panels plots sensitivity for the three target types whereas the bottom row plots criterion. In the upper middle and the upper right panels, the open symbols with dotted lines replot data from the treadmill biological motion condition for ease of comparison.
data suggest that there is no difference in sensitivity between biological motion and nonbiological motion with an underlying form. This provides no evidence that humans are particularly sensitive to biological motion with respect to detection performance and underscores the importance of considering the role of form in perceiving motion.

**Can the results be explained by learning?**

As described earlier, observers in the experiments above completed data collection on the translating biological motion, treadmill biological motion, unstructured translation, and unstructured rotation target types before beginning data collection on the structured translation and the structured rotation target types. Because of this, it is possible that structured translation and structured rotation look more like translating biological motion and treadmill biological motion, respectively, because of learning rather than the role of form in detecting motion. To test for this possibility, three of the four observers who participated in the initial two blocks of trials ran in a third block of trials. Given that the structured rotation matched treadmill biological motion and that there were no concerns about the use of a random mask in these displays, this third block of trials contained unstructured rotation trials. If the results depicted in Figure 3 are due to learning, then this second set of unstructured rotation trials should result in the best performance. If the results depicted in Figure 3 are due to the inclusion of form in nonbiological motion, then one would expect the second set of unstructured rotation trials to be similar to the first set of unstructured rotation trials.

**Figure 4** shows the results for unstructured rotation, structured rotation, and the second set of unstructured rotation trials. Note that these results only include data from the three observers that participated in all conditions. There were significant effects of target size, $F(1.3, 2.601) = 9,722.554, p < .001$, partial $\eta^2 = .999$, number of mask dots, $F(1.767, 3.535) = 455.893, p < .001$, partial $\eta^2 = .996$, the interaction of target size and number of mask dots, $F(1.795, 3.589) = 22.033, p < .05$, partial $\eta^2 = .917$, and most importantly target type, $F(1.302, 2.605) = 23.851, p < .05$, partial $\eta^2 = .923$. Bonferroni-corrected pairwise comparison showed that only the difference between structured rotation and both unstructured rotation conditions approached a significant difference ($p < .075$). Although it is true that the second unstructured rotation condition resulted in slightly better performance than the

---

**Figure 4.** Results from unstructured rotation, structured rotation, and a second run of unstructured rotation. The top row of panels plots sensitivity for the three target types whereas the bottom row plots criterion. In the upper middle and the upper right panels, the open symbols with dotted lines replot data from the original unstructured rotation condition for ease of comparison.
first unstructured rotation condition (see the upper right panel of Figure 4), this difference did not approach a significant difference ($p > .35$). In any case, given that observers can learn to perceive arbitrary motion (Hiris, Krebeck, Edmonds, & Stout, 2005; Jastorff, Kourtzi, & Giese, 2006), it would be surprising if observers did not show some small gain in performance over time.

Criterion in this analysis only depended upon target size, $F(1.074, 2.148) = 19.561$, $p < .05$, partial $\eta^2 = .907$. Again, this indicated that as the target size increased, the criterion became more conservative.

**Does the simplicity of the task explain the results?**

The task used in the main experiment was a simple detection task. In such a simple task, participants can use various strategies to detect the presence or the absence of a target without processing the entire structure of the stimulus (Hiris, Krebeck, et al., 2005). Do the results change when a more perceptually demanding task is used?

A coherence task, where the observer must determine if the top and the bottom half of the walker face the same direction, is a perceptually more demanding task that requires processing at the least information about both the top and the bottom half of the display (Mather, Radford, & West, 1992). Such a task would be ideal to use to determine if a more perceptually demanding task would yield different results. However, it is impossible to create an equivalent coherence task for rotating or translating nonbiological motion (e.g., the top and the bottom half of the display rotating in opposite directions or not would be trivially easy to detect). Instead, a facing task was used. Here the observer must indicate whether biological motion faces left or right or whether rotation is clockwise or anticlockwise. This task was performed for biological motion, unstructured rotation, and structured rotation at the medium target size only.

For the facing task, a target was present on every trial. For these trials, the motion of half of the mask dots was based on one direction of walking/rotation whereas the motion of the other half was based on the other direction of walking/rotation. Without this change in how the mask was constructed, observers would have been able to indicate the correct direction even in the absence of target merely by examining the facing/rotation of the mask dots. Three observers participated; two had not participated in any of the other experiments, the other was the author. Other than these changes, all other factors were the same as in the main experiment.

![Figure 5](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933520/)

Figure 5. Results from the facing task for biological motion, unstructured rotation, and structured rotation. The top row of panels plots sensitivity for the three target types whereas the bottom row plots criterion. In the upper middle and the upper right panels, the open symbols with dotted lines replot data from the biological motion target type for ease of comparison.
Results are shown in Figure 5. An analysis of sensitivity showed no effects of target type or interaction between the target type and the number of mask dots. There was a significant main effect of the number of mask dots as expected, $F(1.95, 3.90) = 171.761, p < .001$, partial $\eta^2 = .988$. An analysis of criterion showed no significant effects of target type, number of mask dots, or their interaction. These results further support the idea that humans are not better at detecting biological motion compared to nonbiological motion. Note that due to the differences in the construction of the masks, data from this experiment should not be directly compared to the main experiment. It is interesting to note that in these data, whether the rotation displays contained structure does not seem to make a difference in detection performance. It is unclear whether this is due to the differences in how the mask was created or the lack of size uncertainty across trials (all targets were medium in size). In any case, there is no evidence in these data that detection of biological motion is better than nonbiological motion. These data do suggest that form may not always be a critical factor in detection performance depending on the exact nature of the mask or task.

Are scrambled target motion masks effective for nonbiological motion targets?

Scrambled target motion masks were used in the main experiment when comparing treadmill biological motion, unstructured rotation, and structured rotation. The results showed that structured rotation and treadmill biological motion were approximately equally detectable. It is well known that scrambled motion masks are very effective in masking biological motion, but do scrambled (un)structured rotation masks result in the same suboptimal detection of (un)structured rotation targets? To determine if this was the case, medium-sized versions of treadmill biological motion, unstructured rotation, and structured rotation conditions were presented, but within a random motion mask. The same three observers that participated in the facing task described in the previous section also participated in this experiment. Other than these changes, all other factors were the same as in the main experiment.

If scrambled motion masks are effective for (un)structured rotation, one would expect to find much better performance with a random motion mask (similar to the difference for biological motion). Data from the three participants in this experiment were compared to the data from the relevant

![Figure 6](https://jov.arvojournals.org/pdfaccess.ashx?url=data/journals/jov/933520/) Plots comparing treadmill biological motion, unstructured rotation, and structured rotation with random motion masks. The top row of panels plots sensitivity for the three target types whereas the bottom row plots criterion. In the upper panels, the open symbols with dotted lines replot data from the main experiment for the same target types with scrambled target motion masks.
conditions from the four participants in the main experiment. As Figure 6 shows, it is clearly the case that random motion masks are much less effective at masking all the target types compared to scrambled motion masks. A mixed factors ANOVA (mask type as a between-subjects variable, whereas target type and number of mask dots as within-subjects variables) showed that all main effects and interactions were significant ($p < .05$). The significant three-way interaction was the result of random motion mask largely resulting in a ceiling effect, but not for the rotation targets with the most dense mask. Note that differences between target types are not surprising given that the motion of the mask dots is the same for all the target types, despite the targets having different motions. The analysis gave the same result without the author’s data (whose data were in both the original and the current data). These data show that the results in the main experiment cannot be explained by scrambled target motion masks not being effective for (un)structured rotation targets.

**Does mask density explain performance?**

The manipulation of the target motion size within a constant-sized mask area allowed replication of the results at different levels of difficulty. The change in target size within a constant mask area also allows for estimating the effective masking area. For a small target size, it is unlikely that the dots near the edge of the larger mask area would influence the detection of the target near the center of the mask area. If mask dots are only effective near the target motion, how close must the mask dots be?

The simplest answer to this question may be that the mask density (that is, number of mask dots within the target area) determines the effectiveness of the mask if everything else is held constant. To determine if this was the case, sensitivity data from Figures 2 and 3 were plotted in terms of mask density. Analyses on sensitivity were performed on the three mask conditions that resulted in mask densities of 7.3, 14.6, and 29.1 mask dots within the target area (a small target with 352, 704, or 1,408 mask dots; a medium target with 88, 176, or 352 mask dots; and a large target with 22, 44, or 88 mask dots). Figure 7 plots all the data from each target size with respect to mask density within the target area.

The analysis of sensitivity in terms of mask density for translating biological motion, unstructured translation, and structured translation no longer showed a significant main effect of target size, nor any significant interactions involving target size. However, for the analysis of sensitivity in terms of mask density for treadmill biological motion,

![Figure 7. Data for the main experiment replotted as a function of mask density.](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933520/ on 11/08/2018)
unstructured rotation, and structured rotation, there was a significant main effect of target size, $F(1.656, 4.968) = 26.885, p < .005$, partial eta$^2 = .9$, although there were no significant interactions involving target size. In general, mask density within the target area does well in explaining sensitivity to the target. However, large targets might result in slightly better sensitivity and small targets results in slightly worse sensitivity than would be expected. This can be seen most clearly in the bottom middle panel of Figure 7.

Although it would be powerful to say that only mask dots within the target area are effective in masking the target motion, one cannot make that claim from these data. Any estimate of the effective masking area that would account for the 16:4:1 ratio of the target size (within the limits of the masking area) would be equally effective in matching the data from the three target sizes. As the analysis and Figure 7 show, however, mask density does seem to be the critical factor in explaining most of the variability in masking the target motions.

Does masking depend on task?

Mask density appears to be the critical factor for sensitivity in a detection task. Is this true for other tasks? In an additional study, treadmill biological motion was presented in both a detection task and a coherence task to determine whether the task influenced masking as well as to replicate the results from the treadmill biological motion condition in the main experiment.

Four observers with normal or corrected-to-normal vision participated. Three observers were naïve and paid US$50 each for their participation. The fourth observer was the author. Stimuli were similar to the treadmill walker and associated mask from the main experiment (see Figure 1D) with the following change: For the coherence task, the top and the bottom half of the walker could face different directions (with matching changes in the mask dot motions).

The detection task was performed in the same manner as the main experiment. Over several days, observers completed 5 runs of 2 sets of 210 trials (with a break between sets) for a total of 2,100 trials per observer (100 trials per condition).

The procedure for coherence task was similar to the procedures for the main experiment. However, in the coherence task, observers judged whether the walker was coherent or not (Mather et al., 1992). On half of the trials, the point-light walker was coherent; the top and the bottom half of the walker faced the same direction, either left or right. On the other half of the trials, the point-light walker was incoherent; the top and the bottom half of the walker faced in opposite directions, left and right. In the initial familiarization procedure for the coherence task, each of the possible point-light walkers (coherent facing left, coherent facing right, incoherent facing left, incoherent

![Figure 8](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933520/)

Figure 8. Sensitivity, criterion, and sensitivity in terms of mask density for treadmill biological motion detection and coherence tasks.
facing right) was shown once with no mask. Observers judged the coherence of the point-light walker and entered his or her response via key presses. Data from an initial practice set of 168 trials were not analyzed. There were 168 trials in each set in the coherence task (3 walker sizes × 7 mask sizes × 4 possible point-light walker displays × 2 repetitions). Over several days, observers completed 6 runs of 2 sets of 168 trials (with a break between sets) for a total of 2,016 trials per observer (96 trials per condition).

Data were analyzed in the same manner as previous data. This analysis on sensitivity showed significant differences between tasks, $F(1, 3) = 14.439, p < .05$, partial $\eta^2 = .828$, with the coherence task being more difficult than the detection task. There was also a significant difference with target size, $F(1,118, 3.354) = 388.422, p < .001$, partial $\eta^2 = .992$, number of mask dots, $F(2,870, 8.610) = 810.715, p < .001$, partial $\eta^2 = .996$, as well as an interaction between target size and number of mask dots, $F(1.424, 4.272) = 20.290, p < .05$, partial $\eta^2 = .871$. There was also a significant three-way interaction between all the independent variables, $F(2,017, 6.050) = 7.178, p < .05$, partial $\eta^2 = .705$, as shown in the left panels of Figure 8.

The analysis on criterion showed a significant effect of number of mask dots, $F(2,456, 7.369) = 12.633, p < .005$, partial $\eta^2 = .808$, and an interaction between the number of mask dots and the target type, $F(1.622, 4.867) = 12.068, p < .05$, partial $\eta^2 = .801$. The middle panels of Figure 8 show that for the coherence task, criterion remained neutral in all cases. However, for the detection task, criterion became more conservative when there were more mask dots. In other words, when the task was difficult due to the number of mask dots, observers were biased to say the target was absent in the detection task but had no bias for coherent or incoherent in the coherence task.

The differences in criterion for the detection and the coherence tasks suggest that observers did approach the two tasks differently. However, an analysis of sensitivity in terms of mask density for detection and coherence no longer showed a significant main effect of target size, nor any significant interactions involving target size (see the right panels of Figure 8). This again suggests that mask density is the critical factor in masking and does not depend upon task even when the tasks differ in difficulty.

**General Discussion**

Are humans better at detecting biological motion compared to nonbiological motion? The evidence presented here suggests several important points about this question. First, humans do appear to be better able to detect biological motion in scrambled target motion masks if nonbiological motion contains no underlying form. However, given that biological motion always has an underlying form, this is not a fair comparison. Second, there is no evidence that humans are better able to detect biological motion if nonbiological motion contains an underlying form and, in some cases, even if it does not have an underlying form. Unlike Bellefeuille and Faubert (1998), who found that adding biological motion to form information resulted in no increase in performance, the data here show that adding form to nonbiological motion increases detection performance so that it is similar to biological motion (although even unstructured rotation was similar to biological motion in the facing task). Third, as expected, given the results of Cutting et al. (1988), the type of mask used may play a role. Random motion masks (translating biological motion, unstructured translation, and structured translation conditions) prove to be less effective for biological motion than nonbiological motion. The use of a random motion mask in previous research may have played a critical role in the results they obtained (Neri et al., 1998; Poom & Olsson, 2002). Fourth, the density of the mask in the target area appears to be the critical variable in detection and the mask density appears to be the critical variable in a coherence discrimination task as well.

Although the evidence presented here does not suggest that humans are particularly sensitive to biological motion in terms of detection performance, biological motion is still special in many regards. As discussed in the introduction, biological motion carries a wealth of information, including information about sex, emotion, identity, and intentions. However, these results indicate only that humans are particularly sensitive to the information within biological motion displays, not that humans are necessarily particularly sensitive (in terms of detection) to biological motion compared to nonbiological motion. It seems suspect to say that humans are “particularly sensitive” to biological motion because we can detect such a wealth of information in the displays, when such information is largely not available in nonbiological motion. It is also unclear whether human observers are particularly efficient at using all of the information available in biological motion displays (Polllick, Lestou, Ryu, & Cho, 2002).

Presumably humans have more experience with biological motion than a spinning square or a translating rectangle, yet the data presented here suggest no better detection performance for biological motion. It is not clear why one would necessarily be better able to detect biological motion over other types of motion. Even threatening biological motion displays (for example, an approaching angry point-light “attacker”) might have their match in nonbiological motion displays (for example, an object moving on a collision course with the observer). Although it makes sense to discuss whether sensitivity to the information within biological motion is due to visual–motor interactions (Jacobs, Pinto, & Shiffrar, 2004; Viviani & Stucchi, 1992), motor training (Casile & Giese, 2006), and/or visual experience (Giese & Poggio, 2003), it seems premature for any such discussion regarding being able to detect biological motion itself. In fact, it has been shown that humans quickly can...
learn articulated nonbiological motions in the same orientation-dependent way as biological motion (Jastorff et al., 2006). Furthermore, observers recognize motion as human, even when the display is inconsistent with the skeleton of the human body (Casile & Giese, 2005). Biological motion might be processed in a nonspecific, nonspecialized manner or, at most, in a way that gives no special detection benefit for it (the feet, however, might be particularly important in classifying motion as biological; see Troje & Westhoff, 2006).

In a model of biological motion perception proposed by Giese and Poggio (2003) and updated by Peuskens et al. (2005), both motion and form pathways feed into areas responsible for processing biological motion. The current data suggest that the additional information available in biological motion displays (emotion, sex, identity, intentions) may not influence detection performance. Detection performance may rely on lower level (although not necessarily low level) motion and form information. Although the motion and form information may interact by motion helping to extract the form and form implying motion, this information can apparently be masked just as easily in biological motion and nonbiological motion displays.

The data presented here underscore the importance of including a comparable nonbiological motion task when studying biological motion. Given the role of form in biological motion, comparable nonbiological motion should include an underlying form (although data from the facing task suggest form might not always be necessary in nonbiological motion). This is particularly the case if one intends to make claims regarding the relative sensitivity to biological motion in human vision (recognizing that sensitivity to the information carried by biological motion is a separate issue not addressed in the current studies). Care also must be taken if masks are used to ensure that the masks are equally effective for the different target motions. Using the same mask for every target motion almost certainly will ensure that the masks will not be equally effective.

### Acknowledgments

I thank Hillary Leech, L. Kim Epting, and two anonymous reviewers for comments on drafts of this manuscript.

Commercial relationships: none.
Corresponding author: Eric Hiris.
Email: ejhiris@smcm.edu.
Address: 18952 E. Fisher Rd., St. Mary’s City, MD 20686.

### References


