Subconfigurations of the human form in the perception of biological motion displays

J. Pinto *, M. Shiffrar

Department of Psychology, Rutgers University, 101 Warren Street, Newark, NJ 07102 USA

Abstract

We report four experiments examining processes that contribute to the perception of point-light displays of human locomotion. In three experiments, we employed a simultaneous masking paradigm to examine the visual system's use of configural information in global analyses of biological motion displays. In the fourth experiment, we obtained descriptions of our stimulus displays from naive observers. Performance in both the detection and identification studies suggests that the visual system responded equivalently to figures exhibiting any organization of limbs that is consistent with the human form. Moreover, the subconfigurations best detected were also most likely to be described independently as depicting a human figure. Thus our findings provide evidence that the visual system can exploit characteristic subconfigurations of the human form in the perception of human locomotion. © 1999 Elsevier Science B.V. All rights reserved.

1. Background and motivation

Human observers are particularly sensitive to human movement. For example, adults can rapidly perceive a human form in a display of discrete elements

---

* Corresponding author. Tel.: +1-973-353-5754; fax: +1-973-353-1171; e-mail: pintoj@andromeda. rutgers.edu

(commonly referred to as "point-lights" and illustrated in Fig. 1a) moving as if attached to the major joints of an otherwise invisible person (Johansson, 1973). Though no explicit contours, textures, or colors indicate the presence of a human form, the visual system is able to extract the structure of a human body from the motion of these elements within a fraction of a second (Johansson, 1976).

Since Johansson's introduction of point-light displays into contemporary perceptual psychology over two decades ago, numerous researchers have attempted to extend general models of the visual perception of structure from motion to account for the perception of human movement from motion-carried information. Under such general accounts, the same visual processes are used to extract the structure of any object. All objects, and all object parts, are thus perceptually equivalent. Indeed, models based on hierarchical vector analysis (Cutting, 1981) or assumptions of pairwise rigidity among elements (Hoffman & Flinchbaugh, 1982; Ullman, 1984; Webb & Aggarwal, 1982) have had some success in accounting for the perception of the complex, jointed structure of a human body.

Nonetheless, these general structural accounts have not been able to mimic either the robustness or the limitations of human performance in the perception of point-light walkers (Proffitt & Bertenthal, 1988). Adult observers fail to accurately detect (Bertenthal & Pinto, 1994), organize (Shiffrar, Lichtey & Heptulla Chatterjee, 1997),
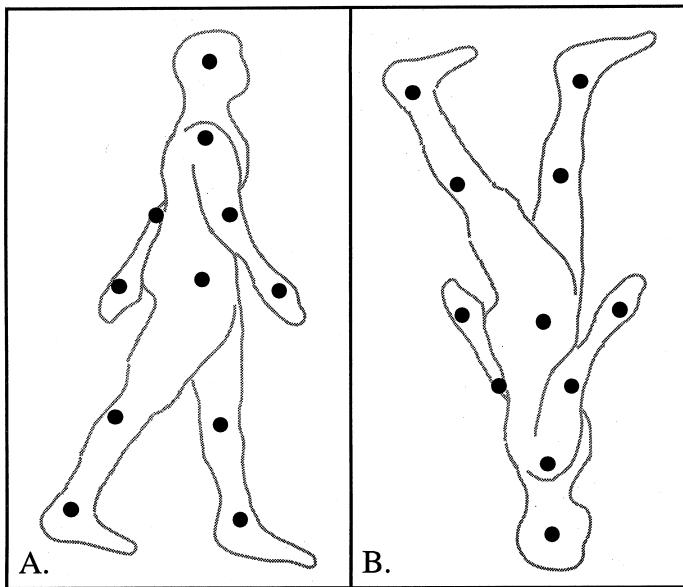

Fig. 1. Static illustrations of point-light displays of a walking human figure. (A) An upright figure in which all of the major joints, or points of articulation, of the human form are demarcated. In the actual displays, the gray outline of a person does not appear. It is included here to illustrate the structure of the figure that quickly becomes apparent when the elements move. (B) The same form rotated to an inverted orientation. Though structurally identical to the upright figures, inverted displays are rarely identified as depicting human form (e.g., Sumi (1984)).

or identify (Sumi, 1984) these displays when the figure is presented upside-down (Fig. 1b). Both upright and inverted displays of human locomotion possess the same hierarchical structure, the same rigid relationships, and the same oscillatory motion trajectories. Under accounts in which perception relies entirely on such properties, the visual analysis of upright and inverted displays should be identical. Human performance, however, is not. Models that posit only general analyses and organizing heuristics are insufficient to account for the orientation-specificity of the perception.

Increasing evidence suggests that the perceptual analyses underlying the perception of human movement may differ from other motion analyses (e.g., Shiffrar & Freyd, 1990; 1993; Viviani & Stucchi, 1992). The visual perception of human movement may benefit from a convergence of motion and form processes that does not occur during other perceptual analyses (McLeod, Dittrich, Driver, Perrett & Zihl, 1996; Oram & Perrett, 1994; Perrett, Harries, Mistlin & Chitty, 1990; Vaina, Lemay, Bienfang, Choi & Nakayama, 1990). Other studies suggest that the visual analysis of human movement differs from other analyses because it depends upon activity of the motor system (Bertenthal & Pinto, 1993; Decety et al., 1997; Stevens, Fonlupt, Shiffrar & Decety, 1999; Viviani, Baud-Bovy & Redolfi, 1997). To the extent that either or both of these approaches are correct, one can conclude that the recognition of a human actor executing some movement may very well involve a special process that taps domain-specific information about human motor activity.

If such a mechanism exists, can it be best described as a global or local process? Local processes are thought to occur in lower levels of the visual system and to be restricted to brief temporal intervals and small spatial neighborhoods. The results of these "local" analyses are then passed onto and processed by higher level or more "global" mechanisms that process information across larger spatio-temporal extents. While local and global are difficult to define as absolute terms, most studies of the visual perception of human movement have defined local analyses as the computations conducted on individual points (joints) or point pairs (limbs). Global analyses are conducted over larger areas and generally involve an entire point-light walker. For a discussion of local and global factors in visual completion, see Tse (1999) and Van Lier (1999).

Evidence supporting the global analysis of human movement comes from studies demonstrating that observers are able to extract human structure from displays in which visual noise renders local motion information organizationally ambiguous. This suggests that the visual system can exploit configural or global information in the absence of unambiguous local motion cues (Bertenthal & Pinto, 1994; Shiffrar et al., 1997). Evidence supporting the hypothesis that the visual analysis of human movement relies on local processes comes from a series of apparent motion experiments in which subjects viewed an animated walker within a mask (Mather, Radford & West, 1992). Since subjects could only discriminate leftward – from rightward-facing walkers under short range apparent motion conditions, their perception of the walkers' movements appears to have depended upon local motion analyses operating within small temporal windows. More recent research has suggested that neither local nor global processes alone can account for the visual perception of human

locomotion (Thornton et al., 1998). Instead, visual processes at both high and low levels of the visual system appear to make important contributions to the visual perception of human movement. This conclusion suggests that a new approach to the study of the visual perception of human movement may be warranted.

Instead of focusing on the local/global debate, the goal of the current series of experiments was to determine the minimal stimulus information necessary for the visual perception of human movement. In this way we hoped to identify the information that normally triggers processing by the mechanisms thought to underlie the visual perception of human locomotion.

Our approach is based on the assumption that the perception of human movement involves the integration of form and motion information. Neither form nor motion information alone defines the category of human movement. As Johansson's earliest demonstrations show, observers do not perceive human form in any single static frame of the animated figure even though the configuration of the elements is consistent with a human figure (Johansson, 1973). Conversely, motion alone does not convey human form and movement. When the elements comprising a point-light display are spatially scrambled so that the configuration of the elements is no longer consistent with a human figure, the impression of human or animal form is greatly diminished (Pinto, 1996). Thus, the visual system's capacity to extract the human form from motion relies on the integration of both form and motion cues.

In order to create varying exemplars of human locomotion, we manipulated two perceptual components that play a fundamental role in the production of human walking: dynamic symmetry among the limbs and the principal axis of organization. Dynamic symmetry refers to the equal and opposite motions of adjacent limbs (either contralaterally or ipsilaterally). During human locomotion, when one limb moves forward, it's neighboring limbs move backward, anti-phase to the first limb. This anti-phase patterning of limb movements is an invariant of human gait (Bernstein, 1967). In the human body, the principal axis of organization refers to the primary structure about which the limbs are organized, namely the torso. Principal axes play an important role in the recognition of objects generally (Ling & Sanocki, 1995). The structure of an object's principal axis also appears to distinguish between classes of animals (Marr, 1982; Pinto, 1996). Thus, the principal axis of organization and the dynamic symmetry of the limb movements are features likely a priori to contribute to the perception of human form and movement.

In addition, empirical evidence shows that the visual system exhibits sensitivity to both of these features in the context of biological motion displays. Previous investigations of the perception of point-light walker displays suggest that the visual system is sensitive to spatio-temporal phase relations in the moving figure. While naive observers in recognition studies readily identify a canonical figure as a person walking, they modify that description to interpret perturbations in spatio-temporal phase relations. They report seeing either a person swaggering or stumbling from side-to-side (Bertenthal & Davis, 1988). Interestingly, these perturbations in spatio-temporal phase do not eliminate observers' recognition of human form per se. Similarly, in detection studies, non-canonical phase relations diminish, but do not eliminate, observers' ability to detect a target figure (Bertenthal & Pinto, 1994). As

inter-limb phase relations depart from a canonical anti-phase patterning, observers detect the figure less accurately (Pinto & Bertenthal, 1992). Nonetheless, detection of such figures remains reliably above chance. Thus, while the visual system is sensitive to phase information, it tolerates a range of phase relations in the perception and interpretation of human form and movement.

The spatial organization of the elements or joints around the torso also contributes to the perception of biological motion displays. Figures in which the elements are positionally scrambled are not identified as human forms, suggesting that the spatial structure of the figure is essential (e.g., Cutting, 1981). Nonetheless, observers can recognize the human form in figures even when the spatial organization of the elements violates important bodily structures such as the unity of the principal axis, or torso (Johansson, 1973; Pinto, 1996).

Taken together, these findings suggest that neither dynamic symmetry nor the structure of the principal axis is necessary for the perception of human movement. Perturbations to either dynamic symmetry or the structure of the principal axis, individually, appear to modify observers' interpretation of the figure, but not to disorganize its perception or preclude its recognition. These findings suggest that dynamic symmetry and the structure of the principal axis are salient features of human locomotion and may contribute, in a probabilistic fashion, to the activation of the mechanisms underlying the perception of locomotion. This hypothesis was tested in Experiments 2 and 3 by presenting subjects with displays in which a human walker's limbs were organized to preserve or eliminate dynamic symmetry and the principal axis of organization. These stimulus manipulations give rise to subconfigurations that exhibit some, but not all, characteristics of human locomotion. Our investigation of the characteristic properties of human movement begins, however, with an investigation of the hypothesis that the trajectories of the wrist and ankle joints are sufficient for the perception of human locomotion (Mather et al., 1992).

2. General methodological approach

Our aim in these studies was to examine the visual system's use of configural information in the extraction of human structure from motion-carried information. Previous research (Bertenthal & Pinto, 1994; McLeod et al., 1996; Thornton et al., 1998; Vaina et al., 1990) suggests that the perception of biological motion displays involves global or configural processing mechanisms that do not rely exclusively on unambiguous local motion analyses. To examine the operation of those global mechanisms, we employed a simultaneous masking paradigm, adapted from Bertenthal and Pinto (1994) and Cutting, Moore and Morrison (1988). In this paradigm, observers view displays containing a point-light human figure that is masked by the addition of superimposed moving point-lights created from scrambled point-light walker elements. Insofar as the visual system relies on the analysis of local motion information in the perception of biological motion displays, these additional noise elements should interfere with the perception of biological form. However, observers perceive human form in the masked display, even though their ability to identify or

track any individual point is disrupted (Bertenthal & Pinto, 1994). Thus, the use of this masking paradigm enables us to focus on the visual system's capacities to extract configural information carried in the moving elements of the target figures.

In addition, we chose a control stimulus that provided us with a measure of the use of configural information in an unfamiliar, complex stimulus. As mentioned above, inverted point-light displays are rarely recognized as human forms, even though they possess all of the local and configural information available in the upright, easily recognized figure. Observers perceive structure in these displays, but the structures they perceive vary widely (Sumi, 1984). It is plausible that perception of the inverted figure employs only general structural analyses. Detection of an inverted figure may thus provide a measure of the effectiveness of general mechanisms in the perception of a complex, jointed figure, similar to the familiar human form. In the studies below, we compared observers' detection of inverted figures to their detection of experimental configurations in order to ascertain what, if any, processing differences are elicited by the features of the experimental configurations.

3. Overview of current studies

The studies reported below examine the operation of visual processes that analyze the configural information contained in point-light displays of human locomotion. Do such processes exploit general structural properties, consistent with general structural accounts, or do they exploit specific characteristics of the human form in motion? Experiment 1 examined the first hypothesis and found general structural accounts inadequate. Our subsequent studies examined the second hypothesis in greater depth. Specifically, we investigated what specific configural information might provide the minimum information necessary to evoke the impression of a human form? Experiment 2 tested a strong version of the issue by examining differential detection of a canonically-organized walking figure and a collection of limbs without the inter-limb organization typical of human locomotion. Experiment 3 continued our investigation by employing biological motion displays in which the availability of two salient properties, dynamic symmetry and the principal axis of organization, was manipulated to create a set of exemplars of human figures walking. In Experiments 2 and 3, we measured the detectability of the figures. In Experiment 4, we measured the identifiability of the same set of figures in a recognition paradigm. If, in the perception of biological motion displays, the visual system exploits information that is highly characteristic of human locomotion, then we should find a correspondence between the detectability of configural information, in a simultaneous masking paradigm, and the identification of figures, in a simple recognition paradigm.

4. Experiment 1 – Does the motion of limb extremities signal human locomotion?

Based on the failure of general structural analyses, Mather and his colleagues proposed that, in the perceptual organization and analysis of point-light figures, the

visual system exploits structure and motion that are highly characteristic of human locomotion (Mather et al., 1992). In their investigations, they presented observers with a set of point-light figures from which they removed elements at different levels of the body's hierarchical structure. Observers performed a direction discrimination task, reporting whether each figure faced to the left or to the right. Mather and his colleagues found that, contrary to predictions drawn from general structural accounts, elimination of the elements representing the ankles and wrists impaired observers' performance most. Mather et al. (1992) speculated *post hoc* that the extremities (i.e., the wrists and ankles) carry motion information characteristic of human movement and that the visual system exploits it.

The broad conclusion advanced by Mather and his colleagues – that general structural accounts do not accord with human perceptual performance – is consistent with previous empirical (Sumi, 1984) and analytic work (Proffitt & Bertenthal, 1988). Nonetheless, their more specific conclusion – that extremities provide characteristic information – is less well supported. Their study examined the perception of directional orientation. Because the ankle and wrist joints of a point-light walker are the only elements that show bilateral asymmetry suggestive of direction, their removal would certainly impair direction discrimination. Thus, one cannot conclude from these findings that these elements signal human movement, or participate in the perceptual organization of the figure.

Thus, in our initial study, we sought to ascertain whether directional orientation and figural coherence judgments rely on the same stimulus information. We repeated Mather et al.'s stimulus manipulation in a different task, a presence/absence judgment. The detection of a target figure requires that the visual system extract a coherent structure–any coherent structure. We reasoned that a presence/absence task would provide a measure of the visual system's sensitivity to structural information contained in our displays. If, as Mather et al. suggest, the extremities carry characteristic information and that information is essential to the perception of coherent structure, then our results should mirror their findings. If, on the other hand, our findings show a different pattern, we would suggest that their results might not generalize beyond judgments of a walker's directional orientation.

4.1. Method

Participants. Ten observers, all students or employees of Rutgers University in Newark, participated in this experiment. All observers had normal or corrected-to-normal vision. Four subjects had had some prior exposure to point-light displays of human movement, though none but the second author had substantial experience with point-light displays. All but the second author were naive to the purpose of the study. During the experiment, the second author was blind to the stimulus conditions in which she was tested. Three additional participants failed to follow instructions and as a result, their data were excluded.

Equipment. Stimuli were displayed on a Macintosh 21″ (40×30 cm) RGB monitor set to 16 color planes and a 1152×870 pixel resolution. The monitor operated with a refresh rate of 75 Hz. Stimulus generation and presentation were controlled by a

Macintosh Quadra 950 with a processor speed of 33 MHz. This same equipment was used in all of the experiments reported here.

Stimuli and Experimental Design. All of the stimulus displays were based on the structure and motion of a person walking, illustrated in Fig. 1a. Human gait was simulated as a hierarchy of nested pendula depicted with 11 discrete elements (Cutting, 1978). The translatory motion component was eliminated from each element's motion vector, creating a figure that appeared to be walking on a treadmill. At its fullest extent, the figure measured 6.9 cm in height and 2.9 cm in width. Each element measured approximately 0.2° visual angle (VA) in height and width. From the subjects' seat, approximately 43 cm from the monitor screen, the stimulus display subtended approximately 9.2° VA in height and 3.8° VA in width.

In this study, we presented four experimental stimuli, the whole upright human figure (described above) and three subconfigurations:

Figure missing extremities. (Fig. 2a): The extremities, specifically the ankle and wrist elements, are eliminated from the otherwise complete walking human figure. According to general structural accounts, these elements are inessential to the computation of structure even though they are salient features of a human body.
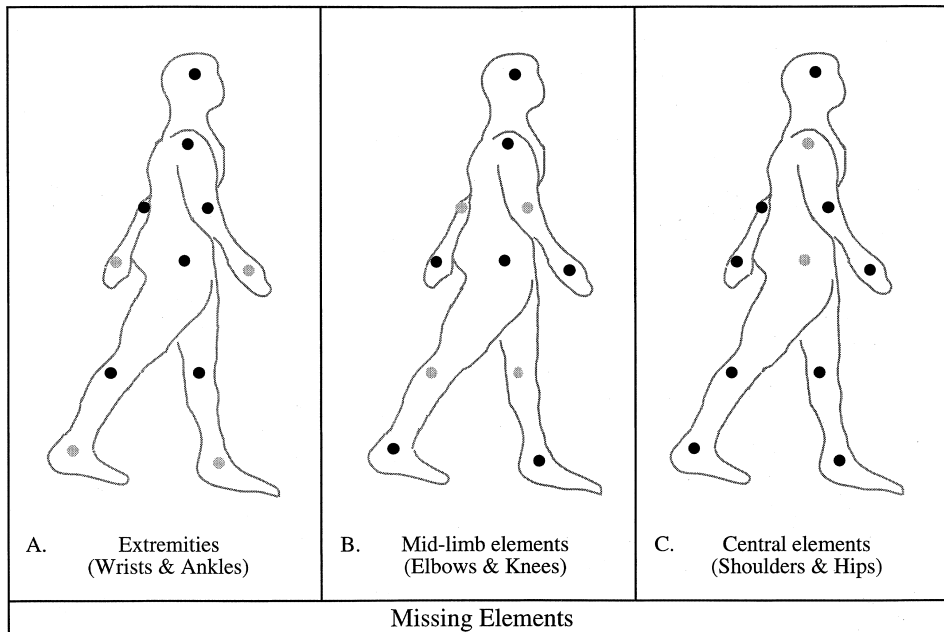

Fig. 2. Illustrations of the target figures presented in Experiment 1. Joints denoted in gray were omitted from the figure, though their trajectories were included in the computation of the locations of other joints. (A) The wrists and ankles were omitted, making the extremities of the limbs invisible. (B) The elbows and knees were omitted, leaving the joint angles internal to the limbs invisible. (C) The shoulder and hip elements were omitted, leaving the body's principal axis of organization invisible (Following Mather et al. (1992)).

Figure missing mid-limb elements (Fig. 2b): In the hierarchical structure of the human body, the knee and elbow joints represent the next level of organization. In this display, these elements do not appear. Omission of these elements eliminates the pair-wise rigid relations, among the limb joints, in the picture plane.

Figure missing central elements (Fig. 2c): In order to reduce the availability of information about the most inclusive or controlling joints, we eliminated explicit demarcation of the shoulder and hip joints, requiring the visual system to interpolate them from the motions of distal visible elements.

In each of these subconfigurations, the movements of the missing elements are computed. Computations for dependent elements of the hierarchy thus maintain their ordinary form. Only the explicit representation of the joint is eliminated.

In addition to the four experimental displays, we presented an inverted human figure, a whole form that was rotated 180° about its horizontal axis (Fig. 1b). We have included it in our stimulus set as a measure of the visual system's ability to detect an unfamiliar structure that shares the same basic symmetrical and repetitive organization as the recognizable, upright human form.

Each of these five configurations served as a target figure in a detection task. We employed a simultaneous masking paradigm to test observers' sensitivity to the figures. Observers were presented with the target figures superimposed with visual noise (Fig. 3c). Visual noise was created by copying the local motion
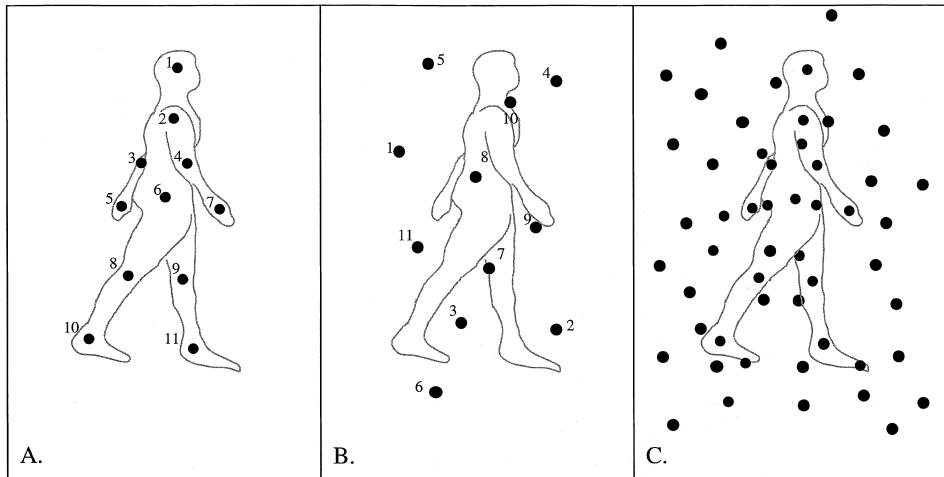

Fig. 3. An illustration of the construction of stimulus displays. Displays were created by superimposing visual noise on the target figure. (A) For reference, an annotated point-light figure. The numbers beside the elements identify a corresponding element in Panel B. (B) Visual noise was created by duplicating the motion vectors of each figural element in the point-light figure. The noise elements were then placed randomly in the display. For example, the element numbered 5 in panel A (the left wrist) is randomly located in panel B. (C) In the target-present display illustrated here, 6 sets of noise elements were superimposed on a target figure. Only the configuration among the elements distinguishes the target and noise elements. In target-absent displays (not shown), the target figure was replaced with an additional set of noise elements.

vectors of the target elements and locating them randomly in a 9×13 cm area in the center of the computer monitor (compare Figs. 3a and 3b). The target and noise elements were identical in size, color, motion, and shape (Cutting et al., 1988). Any elements eliminated from the target figure were also eliminated from the noise, in order to maintain a constant signal-to-noise ratio across experimental conditions. As a result, only configural information distinguished the target from the noise elements. In 50% of the trials, selected randomly, the target was presented with 44 noise elements. In the remaining trials, the target was replaced with 11 additional noise elements located randomly within an area the size of the target figure.

On each trial, the figure produced two full gait cycles, lasting approximately 3200 ms (2 gait cycles $\times$ 40 frames/gait cycle $\times$ 40 ms/frame). The figure's center was located randomly within a 2.3×3.3 cm area centered within the masked portion of the video display, in order to insure that no part of the target appeared unmasked as the limbs oscillated about the torso. The direction it faced (right or left) and the starting frame in the animation were randomly determined. Each stimulus served as the basis for a block of 100 trials. The 5 blocks of trials were intermixed with blocks designed for the other experiments in this series. This minimized potential order effects. For each observer, the order of presentation was determined randomly.

Procedure. To accustom observers to the displays and the task, we administered a two-phase training regimen. The construction of the stimulus displays presented during the training sequence was identical to that of the experimental stimulus displays (described above), except that a point-light figure of a "car" served as the target. In the first phase of training, the car figure was presented without visual noise for 2850 ms. Subsequently, the target was either re-located, or replaced with randomly located elements. Visual noise was superimposed on the target for 750 ms. Subjects judged whether the target remained after the onset of the visual noise. Twenty trials were presented in each block. Training was repeated until the subject achieved 85% accuracy on two consecutive blocks of trials. The second phase of training was identical to the first, except that the "car" figure was presented with visual noise for the entire duration of the trial or 3600 ms. In addition, following a procedure identical to that of the second training phase, observers who had not yet been exposed to point-light displays of a human form were trained to detect the presence of a point-light walker in visual noise.

Subsequent to training, observers were presented the five blocks of experimental trials. Subjects judged the presence or absence of human form, either fully or partially rendered. They were provided no other information about the structure of the stimulus displays. Thus, test performance reflects the relative import of a subconfiguration of elements in the apprehension of human form.

Observers were tested individually. To avoid fatigue, testing was distributed over two or more test sessions. Participants were seated comfortably, arms-length from the display monitor (approximately 43 cm), in a darkened room. They used the computer keyboard to initiate each trial, and to indicate whether or not they detected a target figure.

4.2. Results and discussion

For each subject in each condition, we computed d' (z(hit rate) – z(false alarm rate)), as a measure of sensitivity. Hit and false alarm rates of 0 or 1 were adjusted to eliminate infinite z values (Macmillan & Creelman, 1991). Single-sample t-tests, conducted for each stimulus condition, revealed that observers reliably detected the presence or absence of the target displays (all $t(9)$s > 2.4, all ps < 0.05). Thus, the structure of each target, including the inverted figure, was sufficiently accessible to systematically influence subjects' responses. Within each target condition, we conducted an analysis of variance examining the effect of prior exposure (or lack thereof) on detection performance. Prior exposure exerted no reliable effect on performance in any condition (all $F(1, 9)$s < 2.61, ns). Further, in a within-subject analysis, we found no evidence of a condition × prior exposure interaction ($F(4, 5) = 1.64$, ns).

Did observers' performance vary with the missing elements? To examine this question, we used a repeated measures analysis of variance to compare performance in the four experimental conditions (i.e., displays based on upright figures). As illustrated in Fig. 4, detection of the target figure varied among the stimulus conditions ($F(3, 7) = 11.9$, $p < 0.01$). Further analysis revealed that the decrement in detection was not uniform across subconfigurations. In planned comparisons, we examined the detection of each target figure relative to the detection of the whole. Contrary to Mather et al.'s findings, detection of the figure missing elements on the extremities did not differ from detection of the whole figure ($t(9) = 1.12$, ns). In
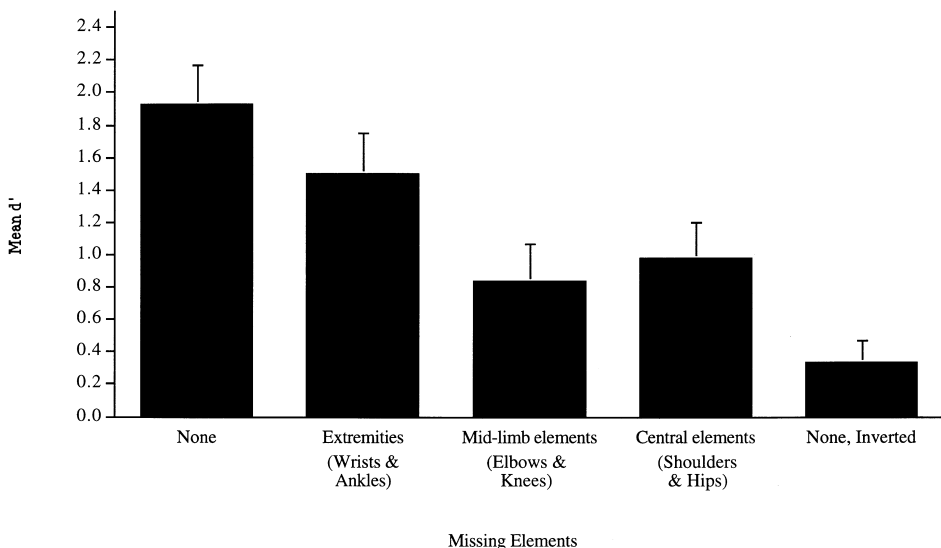

Fig. 4. Summary of results from Experiment 1. The mean d' is plotted as a function of stimulus configuration. Error bars represent the standard error.

contrast, omission of the most central elements in the figure (i.e., the shoulder and hip elements) did significantly diminish performance ($t(9) = 6.04$, $p < 0.001$). Omission of the mid-limb joints also impaired performance ($t(9) = 3.5$, $p < 0.01$). Thus, we did not reproduce Mather et al.'s findings in our task, suggesting that the extremities may not be essential to the detection of human structure, *per se*.

Our findings are also inconsistent, however, with those expected from available models of general structural analyses. Under a strong formulation of hierarchical vector analysis (Cutting, 1981), the visual analysis of a point-light figure begins with the computation of a single reference point to which all the visible elements bear some systematic relationship. The motions of the visible elements are then analyzed relative to that geometric reference. Visual analysis proceeds sequentially through levels of the hierarchical organization of the point-light display. Each of the three subconfigurations presented in Experiment 1 maintained a hierarchical structure amenable to such an analysis. Indeed, computationally, each subconfiguration is equally coherent. The fact that the omission of elements has differential effects on detection performance is thus itself notable and interesting. It suggests that figural coherence may not be the only basis for observers' detection performance in this study.

A similar analysis can be provided for accounts that rely on an assumption of pairwise rigidity to recover the structure of a point-light figure (Hoffman & Flinchbaugh, 1982; Ullman, 1984; Webb & Aggarwal, 1982). In these models, when the distance between a pair of visible elements remains constant as the elements move in a three-dimensional space, the visual system infers that the elements are rigidly connected. Successive applications of this rigidity assumption yield a whole structure. In the present study, each target subconfiguration maintains two of the three elements on each of the limbs of the whole figure. In principle, those elements could be interpreted as rigidly related. Thus, though the subconfigurations manifest fewer rigid parts than does the whole figure, some structural information is available. Indeed, observers detect structure, performing reliably above chance across all conditions. However, rigid relations alone can not account for variations in performance across conditions. Since each subconfiguration possesses the same number of rigid relations, rigidity alone should give rise to equal detection of each. Like accounts based on vector analysis, accounts based on pairwise rigidity alone do not accord with observers' detection performance.

In addition, consistent with our assessments of the operation of these general structural models, detection of all of the experimental figures surpassed detection of the inverted control figure (all $t(9)$s > 2.29, $p < 0.05$). Insofar as the detection of the inverted figure reflects the visual system's analysis of a figure untutored by experience or familiarity, the relative superiority of detection of the experimental subconfigurations suggests that the visual analysis of those subconfigurations did not rely exclusively on the same structural information manifest in the inverted figure. Quite possibly, the visual system exploited configural information associated with human or animal forms and maintained in the subconfigurations. Our next studies investigated this possibility directly.

5. Experiment 2 – Is the relative location of limbs essential to the perception of human form?

The organization of limbs comprising a human or animal form is thought to be an important characteristic in the recognition of a figure. However, the importance of limb organization has been assumed, rather than tested, in most research programs. It is plausible that the visual analysis of human form relies on the presence of configural units extending over relatively small spatial extents. Individual limbs, for example, might provide the visual system with sufficient information to signal the presence of a human, or animal, form. Bertenthal and Pinto (1994) provide indirect evidence in support of such a hypothesis. In a simultaneous masking paradigm, like that employed in the current studies, they masked point-light walker displays with visual noise resembling individual limbs. The limb-like noise elements reduced detection of the whole human figures more than did individual noise elements. Bertenthal and Pinto did not directly compare the detection of the whole figure and randomly-located limbs, so the role of the analysis of limb-like units in the visual analysis of the whole human figure remains unclear.

In our second study, therefore, we compared detection of the whole walker figure to detection of four limbs randomly-located in the display. The whole figure possessed four limbs located along a single principal axis of organization. If the visual analysis of point-light walker displays relies on the spatial inter-limb organization of the figure, then the whole figure should be detected more accurately than the randomly-organized limbs. On the other hand, if the visual analysis of human locomotion relies principally on the presence of limbs, then we should see comparable performance in the two conditions.

5.1. Method

The ten observers who participated in Experiment 1 also participated in this study. The procedure was identical to that employed in Experiment 1. Since the observers had participated in the previous study, we provided no training trials prior to testing.

So as not to increase observers' practice detecting the whole, canonical figure, we did not repeat that condition. Instead, we used the data, drawn from the same subjects, reported in Experiment 1. As a result, in this experiment, observers simply viewed a display composed of four limbs that were positioned at random so as to eliminate all of the canonical inter-limb spatial relations (Fig. 5a). In order to ensure that all of the limbs appeared within the same area occupied by the canonical walker, limbs were permitted to overlap as they moved. The resulting display provides a strong test of the hypothesis that limbs alone is sufficient to signal the presence of a human form.

With the exception of the target figure, the display design and presentation, and the testing procedure were identical to those described in Experiment 1. Observers
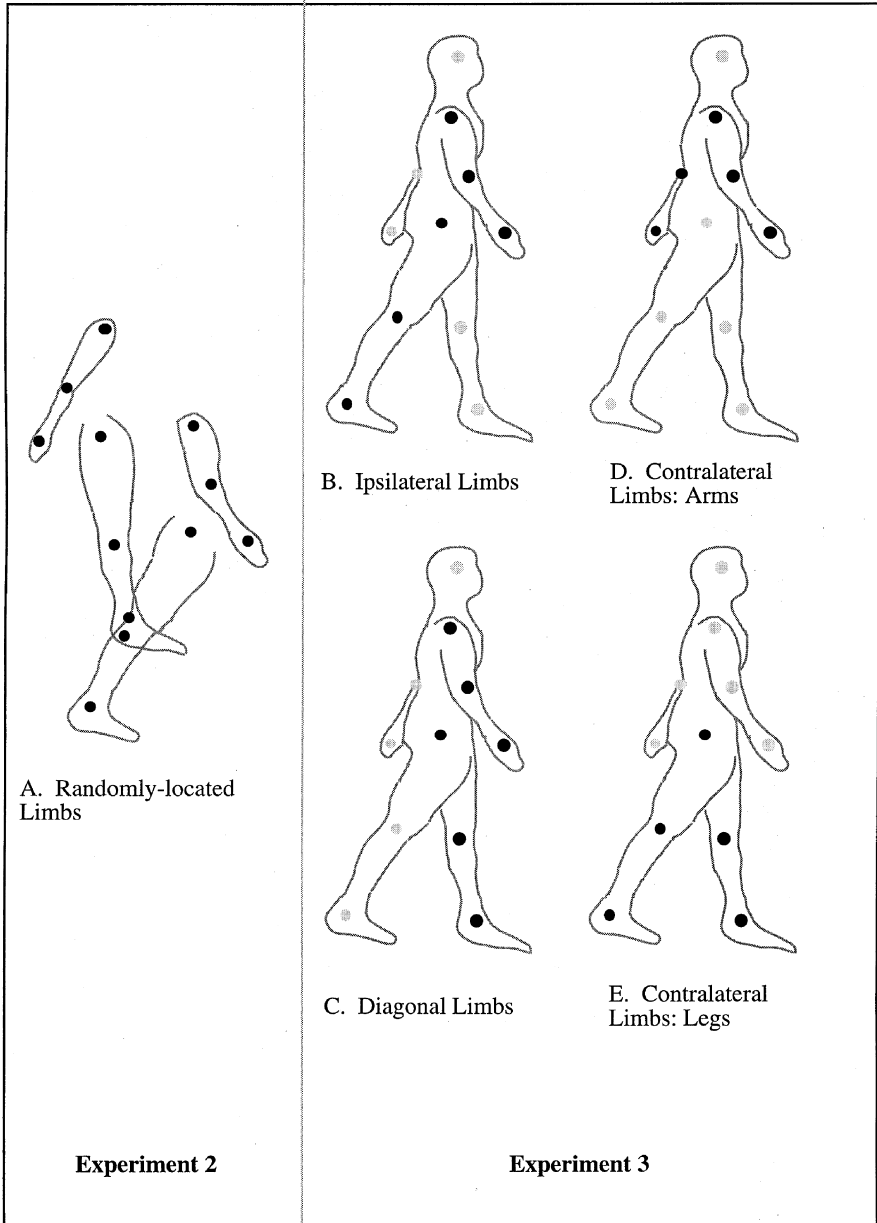
Fig. 5. The design of target figures for Experiments 2, 3 and 4. Black denotes a visible element; gray an invisible element. (A) Randomly-located limbs show none of the spatial inter-limb organization characteristic of human locomotion. Center column: Ipsilateral (B) and Diagonal (C) limb pairs exhibit an elongated principal axis about which the limbs were organized. Right column: Arms (D) and Legs (E) exhibit dynamic mirror symmetry among limbs. In Experiment 2, all four limbs were included in the display. In Experiment 3, target figures included only a single pair of limbs.

judged the presence or absence of a human form – presented in full or in part – on each of 100 trials.

5.2. Results and discussion

As in Experiment 1, we computed d' for each subject. The measure of detection of the whole upright figure from each of the ten participants in the current study was obtained in Experiment 1 and included in this dataset for comparison. An analysis of variance revealed no differential influence of prior exposure to point-light walkers or training prior to Experiment 1 on the detection of the random limb figure ($F(1, 8) =$ 2.89, *ns*).

A single-sample *t*-test revealed that observers reliably detected the presence and absence of the random limb displays ($t(9) = 4.1$, $p < 0.01$). The visual system's ability to detect these figures suggests that sufficient information is available in the limbs alone to discriminate between the target-present and target-absent displays. Nonetheless, detection of the whole figure and the random limb display differed ($F(1, 9) = 35.53$, $p < .001$). The whole figure was detected with greater frequency and accuracy than was the set of randomly-organized limbs.

What might account for the superior detection of the whole figure? The coherence of the whole figure, relative to the randomly-organized limbs, provides an obvious explanation. But coherence is insufficient to account for the entire pattern of results. As reported in Experiment 1, and several previous studies (Bertenthal & Pinto, 1994; Shiffrar et al., 1997; Sumi, 1984), detection of the whole figure diminishes substantially when the figure is rotated 180° in the picture plane. Thus, while greater coherence may contribute to the superior performance, coherence alone can not explain it.

The superior detection of the whole figure, consistent with the findings of previous research, may reflect the deployment of mechanisms specialized or attuned to the global characteristics of human locomotion. The existence and operation of such specialized mechanisms would yield superior detection performance with whole, canonical walkers. However, the stimulus information to which such a mechanism may be responding can not be clearly identified from these results. It is possible that the visual system signals the presence – or possible presence – of a human or animal form when limbs alone are present, but that further analysis by the observers disconfirms the initial signal. When examined more fully, the randomly-organized limbs may possess features that are *inconsistent* with the human form. An observer instructed to determine whether a human form is present may employ criteria that discount the random-limb displays even when the four limbs are present and detected.

A third possibility is that the superior detection of the whole upright figure actually does reflect the operation of visual system mechanisms that specifically signal the presence of human or animal form. If limbs alone are not sufficient to signal the presence of a human walker, what information is sufficient? Answers to this question are important for the specification of hypothesized specialized mechanisms. In the next study, we examined the possibility that characteristic

subconfigurations of human locomotion are wholly sufficient for perception of human movement.

6. Experiment 3 – What subconfigurations are characteristic of human locomotion?

In the studies that follow, we sought to ascertain directly whether diagnostic characteristics of human locomotion might be exploited in the perception of biological motion displays. As discussed in the general introduction, dynamic limb symmetry and the structural framework provided by the principal axis of organization appear to be salient properties of human gait. We thus varied the presence of these properties to create exemplars of human figures. We reasoned that any hypothetical, specialized analyses should be attuned to figures possessing one or both of these properties. More importantly, the absence of both properties should diminish the human-like characteristics of the figure and thus diminish the detectability of the figure.

By slightly altering the computation of the point-light figure, we created three displays: ipsilateral limbs, diagonal limbs, and contralateral limbs. As Fig. 5 illustrates, these subconfigurations maintain one, both or neither of the two organizing features under investigation. For example, the contralateral limbs oscillate anti-phase to one another but do not exhibit an explicit elongated principal axis. We chose to test the perception of several variants of human locomotion as a means of examining whether the relations among exemplars suggests equivalence of the sort one might expect of members of the same perceptual category. We tested observers' ability to detect these subconfigurations in visual noise.

6.1. Method

Eight of the observers who participated in Experiment 1 also participated in this study. The procedure was identical to that employed in Experiment 1. Since the observers had participated in Experiment 1, we provided no training trials prior to testing. Four point-light walker displays were constructed. As in Experiment 1, the stimulus displays were based on the structure and motion of a person walking. However, each display possessed two limbs that were varied, across conditions, to maintain or eliminate dynamic symmetry and the principal axis. The limbs always appeared in the same positions that they would have held if they had been part of a whole walker (given that, as before, the walker could have appeared anyway within the display window). Limb position was always canonical. As a result, upright legs always appeared in the lower portion of the display window while upright arms were presented in the upper portion of the display.

Ipsilateral limbs (Fig. 5b): Ipsilateral limbs – i.e., limbs on the same side of the body – manifest dynamic symmetry. As the right arm moves forward, for instance, the right leg moves backward. In addition, the limbs join the body at the shoulder and hip, demarcating the elongated torso that is the body's the principal axis.

Diagonal limbs (Fig. 5c): Diagonal limbs move in phase with one another. For example, as the right arm moves forward so too does the left leg. As a result, these

limbs do not exhibit dynamic symmetry. Rather, they are entrained in time and space. Each limb joins the body at the shoulder or hip thereby marking the principal axis.

Contralateral arms (Fig. 5d): Adjacent contralateral limbs (be they arms or legs) maintain the dynamic symmetry of opposing movements. They are not organized along an elongated principal axis. Rather, they emanate from a single location. In this case, only the arms of the walker were displayed as normally attached to the shoulders.

Contralateral legs (Fig. 5e): Same as above except that a pair of legs was shown emanating from the torso.

Each of the four subconfigurations served as the target for a block of 100 trials. To create visual noise, the individual motion vector of each element in the target subconfiguration was repeated four times, resulting in a signal-to-noise ratio of 0.25. As in Experiment 1, the noise elements were randomly positioned in the display to mask the target figure. Elements omitted from the target figure were also omitted from the visual noise. As before, subjects were instructed to report the presence or absence of the human form – whether complete or partial.

6.2. Results and discussion

As in Experiment 1, we computed d' for each subject in each condition. Performance in the two contralateral limb conditions (arms and legs) did not differ ($t(7) = 0.90$, *ns*). So, for ease of exposition, we created a composite score for contralateral limbs, averaging d' for the two conditions. Single-sample *t*-tests revealed that observers reliably detected the presence or absence of the target displays (all $t(7)$s > 5.1, all *p*s < 0.01). Measures of detection of the whole upright and whole inverted figures from each of the eight participants in the current study were obtained in Experiment 1 and those measures were included in this dataset for comparison. An analysis of variance conducted within each condition revealed no differences in performance as a function of differences in prior experience or training regimen (all $F(1, 6)$s < 2.11, *ns*). In addition, in a within-subjects analysis, we found no interaction between prior experience and the display conditions ($F(2, 5) = 0.24$, *ns*).

Did detection for the target figure vary across experimental displays? As Fig. 6 shows, sensitivity was greatest for the contralateral limb pairs and least for the diagonal limb pairs. Planned comparisons confirmed that detection of the whole figure and of the diagonal and ipsilateral limb pairs differed reliably ($t(7) = 3.49$, $t(7) = 3.26$, respectively, both $p < 0.05$). Detection of the whole and contralateral limb pairs, however, did not differ ($t(7) = 1.29$, *ns*). Nonetheless, detection performance among the four displays (whole walker, contralateral limbs, ipsilateral limbs, and diagonal limbs), examined with repeated measures analysis of variance, did not differ reliably ($F(3, 5) = 3.27$, *ns*), suggesting that the differences in detection are best understood as quantitative, rather than qualitative, differences. Nonetheless, detection of all subconfigurations differed from that of the inverted control stimulus (all $t(7)$s > 3.65, $p < 0.01$), suggesting that detection did not rely on general mechanisms alone.

Like the findings of Experiment 1, the results of the current study do not readily support any available account of the perception of biological motion displays in
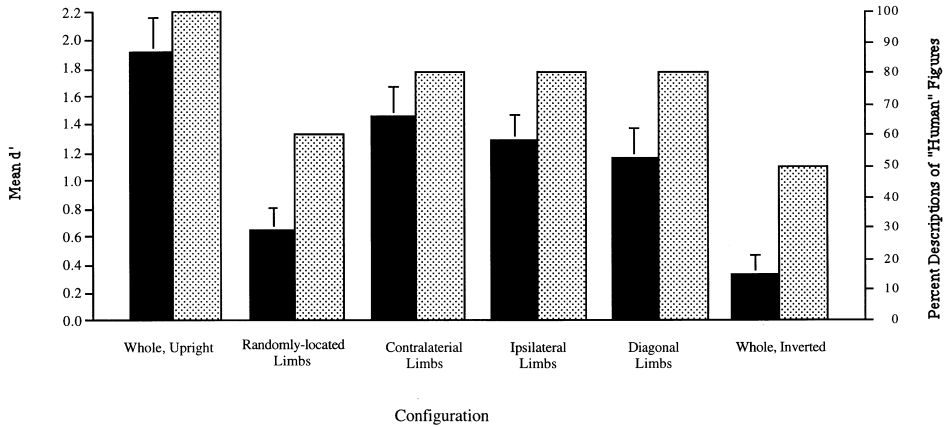
Fig. 6. Summary of results from Experiments 2, 3, and 4, organized by stimulus configuration. The black bars (associated with the left axis) represent the mean d' obtained in Experiments 1, 2, and 3. Error bars represent the standard error. The gray bars (associated with the right axis) represent the proportion of free response descriptions, obtained in Experiment 4, in which human form was mentioned.

which general structural information is systematically or uniformly exploited by the visual system. Nor does the pattern of detection performance we obtained reflect the systematic exploitation of structural information more specifically pertinent to human or animal forms, that is, either dynamic limb symmetry or a principal axis of organization. If visual detection required the dynamic symmetry of the figures, then we would expect that the ipsilateral and contralateral limbs, but not the diagonal limbs, would be detected. If the visual system relied on the elongated structure of the body's principal axis, then we would expect comparable performance in all but the contralateral limb condition. Neither of these patterns was obtained. Detection of the four subconfigurations did not differ ($F(2, 6) = 1.53$, ns). Moreover, the presence of both features yielded no statistically significant gain over and above the presence of either feature alone, in this task, suggesting that neither an additive nor a multiplicative combination of these properties can account for differences in the detection of the figures. However, while neither feature appears necessary, the absence of both reduced detection of the randomly-organized limbs to chance, confirming that the features are pertinent to the perception of the figures.

In this study, we found that, regardless of the specific composition of the limbs presented, when the organization of limbs was consistent with human form, the visual system appears to have treated the figures as equivalent. Such findings, suggestive of a categorization process, have been reported also in previous research (Dittrich, 1993; Verfaillie, 1993). Indeed, our findings suggest that the analysis of biological motion displays involves a multi-featured, probabilistic category structure in which exemplars exhibiting different features are treated as equivalent (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976; see Smith & Medin, 1981 for a review). Consistent with a graded category structure, the slight differences in detection may suggest that some exemplars are perceived as more representative of the

category than others (Rosch & Mervis, 1975; Rosch et al., 1976). Our findings therefore provide evidence that global analyses of biological motion displays exploit configural information indicative of human locomotion as a category.

In these studies, observers were instructed to judge the presence or absence of a human figure – wholly or partially represented in the display. Thus, the category membership of the figures was established semantically in advance of the experimental trials. Under such circumstances, the presence of either dynamic symmetry or an elongated principal axis, together with features we have not yet identified, may provide sufficient information for the visual detection of the target figures. As a result, this study may provide an overly liberal measure of the information required to evoke visual processes attuned to human locomotion. Unfortunately, while the masking paradigm affords us a means by which to isolate the global analysis of configural information in biological motion displays, it requires that we limit the interpretation of figures to a binomial classification, human or not human. Differences in the perceptibility of these subconfigurations may reflect the degree to which they appear characteristic of human locomotion, but they may also reflect more general global features such as figural coherence, redundancy, or symmetry. Thus, we sought converging evidence in a fourth study in which observers' responses were minimally limited by the experimental task.

7. Experiment 4 – The identification of human movement

The outcomes of Experiments 2 and 3 suggest that the visual analysis of biological motion displays exploits characteristic subconfigurations of human locomotion that can not be reduced to a single feature or a computationally combined set of features. However, in the paradigm we employed in those studies, observers' responses to the task or to the instructions may have influenced their performance. More specifically, the masking paradigm permits successful detection of a figure using minimally sufficient information, particularly since our instructions may have primed the observers to look for human characteristics. Thus, it is possible that the information necessary to signal human form is broader than our current findings suggest. Conversely, our procedure may have resulted in an underestimate the importance of simple features. Observers, instructed to judge the presence of a human form, could detect a figure but judge it as non-human, as a result of its perturbations or missing components. This concern is particularly relevant to the interpretation of low detection performance in the randomly-located limb and inverted walker conditions where the displays possess features inconsistent with human form.

To remedy this possible procedural shortcoming, we conducted another study, using a simple recognition paradigm. We presented naive observers the configurations employed in Experiments 2 and 3, without visual noise. Observers then provided free response descriptions of each configuration presented. We sought these responses in order to obtain a measure of the degree to which each subconfiguration appears characteristic of a human figure in motion, without potentially biasing instructions.

7.1. Method

Participants. Seventy students at Rutgers University participated in this study in order to fulfill a course requirement or to receive extra credit in a psychology course. All subjects had normal or corrected-to-normal vision. One additional observer was tested, but excluded due to experimenter error. Another observer was excluded because he had seen point-light displays of human locomotion previously.

Stimuli. The five subconfigurations presented in Experiments 2 and 3, as well as the whole, upright and inverted, walkers from Experiment 1, were included in this study. The figures were presented unmasked, for 3200 ms, or two gait cycles.

Design and Procedure. Observers were tested individually. Each observer was presented one configuration, selected randomly and asked to describe, in writing, what he or she saw. In order to avoid biasing the observer, the experimenter said nothing about the nature of the stimulus displays. If necessary, the observer was permitted a second presentation of the stimulus display. Each configuration was presented to 10 observers.

7.2. Results and discussion

Since we were most interested in ascertaining whether and which subconfigurations elicit the impression of human form, each response was classified as *human* or *non-human*, according to the entity participants described. Two assistants, naive to the purpose of the study, independently coded subjects' responses. Classification criteria and examples are provided in Table 1. The two raters agreed on 100% of the response classifications.

So as to avoid distributional assumptions, we used X^2 statistics to make pairwise comparisons between responses to the whole upright figure and responses to each of the other configurations. Each set of contralateral limbs (arms and legs) was ana-

Table 1
Classification criteria for Experiment 4

Classification	Criteria	Examples
Human	Any mention of a person	"A person walking." "...two people's legs walking." "A man or animal walking." "A guy with no head holding a helium balloon."
Non-human	Any other description, including descriptions that mention of an artifact which mimics a person	"[A] marionette" "A search light." "[Dots] go around in a circular motion." "A couple of pendulums swinging."

lyzed separately, since X^2 is very sensitive to sample size. Subsequent to the initial X^2 analysis, however, the two conditions were combined into a single condition, *contralateral limbs*.

As illustrated in Fig. 6 (gray bars), the diagonal, ipsilateral, and contralateral (both arms and legs) configurations elicited ''human'' responses nearly as often as did the whole, upright figure (all X^2 (1, $N = 10$) = 2.22, *ns*). Only responses to the randomly-located limbs and to the inverted control figure differed from responses to the upright whole figure (X^2 (1, $N = 10$) = 5.00 and X^2 (1, $N = 10$) = 6.67, respectively, both $p < .05$).

Like Experiments 2 and 3, the results of this identification study suggest that membership in the perceptual category of human form is graded. Configurations which provided some, but not all, of the information specifying a whole human body evoked the impression of human form. Still, the frequency of identification varied. Some displays provided more identifiable exemplars than others did.

Thus, the pattern of results we found in observers' identification of the point-light figures resembles the pattern we found in our detection studies. To examine this parallel more directly, we compared the detection and identification measures of each of the configurations presented. As Fig. 6 suggests, configurations which gave naive observers the impression of human movement were the same configurations which trained observers detected reliably in visual noise. Indeed, performances in the two measures are significantly correlated ($r = 0.986$, $p < .001$). Since task demands should not have increased the likelihood that naive observers would identify the target figure as human, their responses should reflect the degree to which each configuration exhibits characteristics indicative of human form. Our findings therefore are consistent with the proposal that the visual system exploits configural information specifically indicative of human form in the perception of biological motion displays.

8. General discussion

The principal aim of the current studies was to determine whether and how visual motion analyses might be specifically attuned to properties of human locomotion. Experiment 1 showed that limb extremities are not themselves indicative of human form, *per se*, as proposed by Mather et al. (1992). Instead, while the movements of the feet and hands may play a fundamental role in the perception of a walker's heading, such movements are not sufficient for the exquisite sensitivity with which observers identify the presence or absence of human figures. On the other hand, the results of Experiment 2 indicate that the visual system does exploit some structural information that is specifically characteristic of human form during the analysis of human movement. Subjects more accurately determined the presence or absence of human movement when that movement was consistent with the structure of the human body than when that same movement was rearranged such that the limbs were positionally inconsistent with the human form. Experiment 3 examined organizing features which might potentially signal the presence of a human form.

These characteristics, dynamic limb symmetry and the organization of limbs about the principal axis or torso, might play a fundamental or even defining role in the visual analysis of human movement. The visual system treated these figures as equivalent, though the structural properties evident in the stimuli differed. This finding suggests that different cues are sufficient to trigger processing by the mechanism thought to be responsible for the perception of human movement. From an ecological point of view, this result would be expected for observers to be able to identify human locomotion rapidly and accurately across differing view points and conditions of partial occlusion. For example, in the everyday world, observers are clearly able to identify human movement while a person is seated or walking behind a desk. The results of Experiment 4 suggest that the results of Experiments 2 and 3 are robust and generalize beyond presence/absence judgments.

Perhaps the most suitable formulation of our findings would be to propose, as Mather et al. do, that the visual system exploits characteristic features or subconfigurations of human locomotion in the extraction of configural information from biological motion displays. Indeed, taken together, the results of Experiments 2, 3 and 4 indicate that the configurations best detected by observers relying on global visual analyses (that is, masked point light walker displays) were also most likely to be described as depicting a human figure by naive observers. Characteristic subconfigurations may provide a means by which the visual system maps the structure of a percept onto real, often highly familiar, entities in the environment. As such, they are not simply arbitrary patterns of a visual array that the visual system detects and organizes. Instead these subconfigurations are interpreted as they relate to a meaningful framework, the known environment (Gibson, 1979/1986). This may provide the processing flexibility necessary for the robust detection of human movement under varying conditions.

The subconfigurations we presented do not appear to be perceptually decomposed into constituent properties, such as the structure of the principal axis or dynamic symmetry. Rather, these features may be integral dimensions of the body and movement. The integration of dimensions may reflect wholistic processing of those dimensions (Garner, 1974; Kemler Nelson, 1993) in the context of biological motion.

This reading of our findings complements much previous research on the perception of biological motion displays. A substantial body of research has suggested that the perception of biological motion displays engages spatially global visual analyses (e.g., Bertenthal & Pinto, 1994; Cutting et al., 1988; Shiffrar et al., 1997). Though "global", and its counterpart "local", are ill-defined, they hold some heuristic value in describing alternative processing accounts. Local mechanisms are thought to operate early in visual analysis, process elementary image features typical of a wide variety of objects, and integrate information over small spatial extends. The results of these local analyses are then passed on to higher level units that integrate information across larger spatial extents (Van Essen & DeYoe, 1995; Zeki, 1993). In a system reliant on local spatial analyses, the specific structure of a human figure would be incrementally constructed through progressive recombination and integration of local features into a specific composite whole. A mechanism operating on a more encompassing, global level of analysis, on the other hand, would not need

to begin with very general, small features, but rather would be specifically attuned to a global form or some characteristic subconfiguration of a class of objects or events. Such a system would be able to detect a biological form in motion across a variety of local perturbations in the image or motion (e.g., partial occlusion, interference), but might be particularly influenced by transformations that would alter the global form of a moving figure (e.g., inversion). The perception of biological motion exhibits many of the strengths and limitations attributed to global visual analyses. In such a context, we might regard subconfigurations like those we presented to observers as elemental features extending over a relatively large area of the visual field. The results of our studies suggest that the visual detection of limbs, when they are organized consistent with a human form, does not differ significantly from the visual detection of an entire walker. This result suggests that limb pairs may serve as the fundamental building block upon which our perceptions of human locomotion are constructed.

Since human movement is a ubiquitous part of our environment and carries social and survival significance, highly attuned or efficient processing of biological motion might be expected to arise in the course of individual or phylogenetic development. Such developments can be understood in terms of category-specific processes (e.g., Farah, 1991; Farah, McMullen & Meyer, 1991; Rumiati, Humphreys, Riddoch & Bateman, 1994; Warrington & Shallice, 1984). Following cortical injury, some patients show perceptual deficits that are distributed across semantic categories unequally. Warrington and Shallice (1984), for example, described a patient (J.B.R.) who demonstrates impaired object identification even though he was able to identify colors, shapes, and letters. Of particular interest is the fact that his ability to identify pictures of animals was more greatly impaired than was his ability to identify pictures of objects. While our understanding of category specific processing is far from clear (Damasio, Damasio & Tranel, 1990; Damasio, Tranel & Damasio, 1993; Farah & McClelland, 1991; Farah et al., 1991; Gaffan & Heywood, 1993; Warrington & Shallice, 1984), research in this area does suggest that the mature visual system may acquire memory structures and processes specialized for entities or events within a category or domain.

Category-specific effects might involve either feed forward processes especially attuned to the processing of dynamic form (Zeki, 1993) or action (Vaina et al., 1990), or feedback processes through which stored representations influence the sensitivity or operation of lower-level processes. We should bear in mind in considering these possibilities that our findings also provide evidence that the visual system may exploit general configural information in these displays as well. Detection of the inverted control figure, adopted here as a measure of the visual system's ability to detect an unfamiliar figure of identical structure and complexity, surpassed chance. Of course, it is not just possible, but probable that the visual system employs both general and category-specific processes in the perception of familiar objects and events. Certainly, specialized processes can emerge from general processes (Elman et al., 1996; Karmiloff-Smith, 1992). Studies suggest that young infants do not appreciate the organization of limbs about the body until 20 to 28 weeks of age (Pinto, 1996). Interestingly, orientation-specificity also arises in their responsiveness to

biological motion displays during this time. Such orientation-specificity resembles that evident in adult perceptual performance, and is typically interpreted as a reflection of the familiarity of the upright figure. Such findings clearly indicate that the perception of a whole human form in biological motion displays is acquired in the lifetime (albeit the early lifetime) of an individual. They do not preclude, however, the possibility that the perception of parts or subconfigurations may precede the perception of the whole. The results of the research reported here, understood in a developmental perspective, will guide our future investigations of the interaction of general and category-specific processes.

Acknowledgements

References

Bernstein, N. (1967). *The coordination and regulation of movements*, Oxford, England: Pergamon Press.

Bertenthal, B.I., & Davis, P. (1988). Dynamical pattern analysis predicts recognition and discrimination in biomechanical motions. *Proceedings of the Annual Meeting of the Psychonomic Society.* Psychonomic Society Publications, Austin, Texas.

Bertenthal, B.I., & Pinto, J. (1993). Complementary processes in the perception and production of human movements. In: Thelen, E., & Smith, L., *Dynamical approaches to development: Vol. 2. Approaches*, pp. 209–239. Bradford Books, Cambridge, MA.

Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*, *5*, 221–225.

Cutting, J. E. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behaviour Research Methods & Instrumentation*, *10*, 91–94.

Cutting, J. E. (1981). Coding theory adapted to gait perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 71–87.

Cutting, J., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception & Psychophysics*, *44*, 339–347.

Damasio, A.R., Damasio, H. & Tranel, D. (1990). Impairments of visual recognition as clues to the processes of memory. In: G.M. Edelman, W.E. Gall, W.M. Cowan, *Signal and sense: Local and global order in perceptual maps*, pp. 451–473. Wiley, New York, NY.

Damasio, A. R., Tranel, D., & Damasio, H. (1993). Similarity of structure and the profile of visual recognition defects: A comment on Gaffan and Heywood. *Journal of Cognitive Neuroscience*, *5*(3), 371–372.

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.

Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, *22*, 15–22.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*, Cambridge, MA: MIT Press.

Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object recognition. *Cognitive Neuropsychology*, *8*, 1–19.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*(4), 339–357.

Farah, M. J., McMullen, P. A., & Meyer, M. M. (1991). Can recognition of living things be selectively impaired?. *Neuropsychologia*, *29*(2), 185–193.

Gaffan, D., & Heywood, C. A. (1993). A spurious category-specific visual agnosia for living things in normal human and nonhuman primates. *Journal of Cognitive Neuroscience*, *5*(1), 118–128.

Garner, W.R. (1974). *The processing of information and structure*. Hillsdale, Erlbaum, NJ.

Gibson, J. J. (1979/1986). *The ecological approach to visual perception*. Hillsdale, Erlbaum, NJ.

Hoffman, D. D., & Flinchbaugh, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, *42*, 195–204.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Review*, *38*, 379–393.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on Cognitive Science*, Cambridge, MA: MIT Press.

Kemler Nelson, D. G. (1993). Processing integral dimensions: The whole view. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(5), 1105–1113.

Ling, X., & Sanocki, T. (1995). Major axes as a moderately abstract model for object recognition. *Psychological Science*, *6*(6), 370–375.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*, New York: Cambridge University Press.

Marr, D. (1982). *Vision.* Freeman, New York.

Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society of London*, *249*, 149–155.

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*, *4*, 363–391.

Oram, M. & Perrett, D. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience*, *6*, 99–116.

Perrett, D., Harries, M., Mistlin, A.J. & Chitty, A.J. (1990). Three stages in the classification of body movements by visual neurons. In H. B. Barlow, C. Blakemore & M. Weston-Smith, *Images and understanding*, pp. 94–107. Cambridge University Press, Cambridge, England.

Pinto, J. (1996). *Developmental changes in infants' perceptions of point-light displays of human gait.* Unpublished doctoral dissertation, University of Virginia, Charlottesville.

Pinto, J. & Bertenthal, B.I. (1992). Effects of phase relations on the perception of biomechanical motions. *Investigative Ophthalmology and Visual Science, 33 suppl.,* 1144.

Proffitt, D.R. & Bertenthal, B.I. (1988). Recovering connectivity from moving point-light displays. In: W.N. Martin & J.K. Aggarwal (Eds.), *Motion understanding: Robot and human vision*, pp. 297–328. Kluwer, Boston, MA.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Rumiati, R. I., Humphreys, G. W., Riddoch, J. M., & Bateman, A. (1994). Visual object agnosia without prosopagnosia or alexia: Evidence for hierarchical theories of visual recognition. *Visual Cognition*, *1*(2–3), 181–226.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, *1*, 257–264.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, *4*, 379–384.

Shiffrar, M., Lichtey, L., & Heptulla Chatterjee, S. (1997). The perception of biological motion across apertures. *Perception & Psychophysics*, *59*(1), 51–59.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*, Cambridge MA: Harvard University Press.

Stevens, J. A., Fonlupt, P., Shiffrar, M., & Decety, J. (1999). Selective recruitment of motor and parietal cortex during visual perception of apparent human movement. (in submission).

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, *13*, 283–286.

Thornton, I., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, *15*, 535–552.

Tse, P. U. (1999). Complete mergeability and amodal completion. *Acta Psychologica*, *102*, 165–201.

Ullman, S. (1984). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, *13*, 255–274.

Van Essen, D.C. & DeYoe, E.A. (1995). Concurrent processing in primate visual cortex. In M. Gazzaniga (Ed.), *The cognitive neurosciences*, pp. 383–400. MIT Press, Cambridge, MA.

Van Lier, R. (1999). Investigating global effects in visual occlusion: From a partly occluded square to the back of a tree-trunk. *Acta Psychologica*, *102*, 203–220.

Vaina, L., Lemay, M., Bienfang, D., Choi, A., & 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.

Verfaillie, K. (1993). Orientation-dependent priming effects in the perception of biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(5), 992–1013.

Viviani, P., Baud-Bovy, G., & Redolfi, M. (1997). Perceiving and tracking kinesthetic stimuli: Further evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1232–1252.

Viviani, P., & Stucchi, N. (1992). Biological movements look constant: Evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 603–623.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, *107*, 829–854.

Webb, J. A., & Aggarwal, J. K. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence*, *19*, 107–130.

Zeki, S. (1993). *A vision of the brain*, Cambridge, MA: Blackwell.