The visual perception of human and animal motion in point-light displays

Jeannine Pinto
Lafayette College, Easton, PA, USA

Maggie Shiffrar
Rutgers University, Newark, PA, USA

Mounting neurophysiological evidence indicates that the visual analysis of human movement differs from the visual analysis of other categories of complex movement. If different patterns of neural activity underlie visual percepts of human and nonhuman movement, then psychophysical measures should elucidate different patterns of visual sensitivity to human movement and similarly complex, but nonhuman movement. To test this prediction, two psychophysical studies compared visual sensitivity to human and animal motions. Using a simultaneous masking paradigm, observers performed a coherent motion detection task with point-light displays of human and horse gait, presented upright and inverted. While task performance indicated the use of configural processing during the detection of both human and horse motion, observers demonstrated greater visual sensitivity to coherent human motion than coherent horse motion. Recent experience influenced orientation dependence for both types of motion. Together with previous neurophysiological findings, these psychophysical results suggest that the visual perception of human movement is both distinct from and shares commonalities with the visual perception of similarly complex, nonhuman movement.

Keywords: Vision; Biological motion perception; Autism.

INTRODUCTION

As inherently social beings, humans must rapidly perceive and interpret the actions of other people. Consistent with this necessity, humans demonstrate impressive visual sensitivity to the movements of other people (for review see Blake & Shiffrar, 2007). For example, Johansson (e.g., 1973, 1975) famously demonstrated that observers readily detect human action in highly degraded movies known as point-light displays (illustrated in Figure 1a). The perception of human movement in point-light displays is rapid (Johansson, 1976) and tolerates substantial noise across space (Pinto & Shiffrar, 1999), time (Thornton, Pinto, & Shiffrar, 1998), and luminance contrast (Ahlström, Blake, & Ahlström, 1997). In addition, naïve observers can detect a point-light-defined person’s actions (e.g., Dittrich, 1993), emotional states (Chouchourelou, Matsuka, Harber, & Shiffrar, 2006; Clarke, Bradshaw, Field, Hampson, & Rose, 2005; Pollick, Paterson, Bruderlin, & Sanford, 2001), gender (e.g., Pollick, Kay, Heim, & Stringer, 2005), identity (Jokisch, Daum, & Troje, 2006; Loula, Prasad, Harber, & Shiffrar, 2005), intentions (Runeson & Frykholm, 1983; Sebanz & Shiffrar, 2009), vulnerability (Gunn, Johnston, & Hudson, 2002) and potential reproductive fitness (Brown et al., 2005).
In the past few years, substantial progress has been made in identifying the neural circuitry underlying the visual perception of point-light displays of human movement (e.g., Grezes et al., 2001; Puce & Perrett, 2003; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001). Brain imaging has demonstrated that the posterior region of the superior temporal sulcus (STSp) reacts strongly to point-light depictions of coherent, but not scrambled, human movement (Bonda, Petrides, Ostry, & Evans, 1996; Grossman et al., 2000). Neural activity in the STSp is much greater during the perception of canonically oriented than inverted point-light displays of human movement (Grossman & Blake, 2001) and appears to be hemisphere-dependent as the right STSp frequently responds more strongly than the left (Beauchamp, Lee, Haxby, & Martin, 2003; Peuskens, Vanrie, Verfaille, & Orban, 2005). Finally, research with transcranial magnetic stimulation (Grossman, Battelli, & Pascual-Leone, 2005) and lesion patients (Saygin, 2007) suggests that functionality within area STSp is required for the accurate perception of point-light displays of human movement.

Visual perception of point-light displays of human movement also selectively triggers activity in the human premotor cortex (Saygin, Wilson, Hagler, Bates, & Sereno, 2004). Lesion evidence indicates that a functional premotor cortex, like STSp, is a prerequisite for the differentiation of coherent and scrambled point-light actors and the identification of point-light actions (Saygin, 2007). Electroencephalography (EEG) data further support the hypothesis that premotor cortex, part of the mirror system (e.g., Decety et al., 1997), plays a critical role in the perception of coherent, but not scrambled, point-light displays of human motion (Ulloa & Pineda, 2007). Furthermore, significantly greater motor system activity is found when observers view actions that they can perform than actions that they cannot perform (Stevens, Fonlupt, Shiffrar, & Decety, 2000). Psychophysical evidence with typical observers indicates that visual sensitivity to human actions depends on the observer’s past motor experience with the actions being observed (Jacobs, Pinto, & Shiffrar, 2004; Casile & Giese, 2006). Research with hemiplegic patients indicates that lesions to the motor system selectively degrade visual sensitivity to point-light actions corresponding to the observers’ compromised limbs (Serino et al., 2009).
Evidence from psychophysical and brain-imaging studies indicates that the visual perception of human movement differs from the visual perception of object movement. The mirror system responds to the perception of actions but not objects (Rizzolatti, Fogassi, & Gallese, 2001). Differences in STSp activity, as measured by functional magnetic resonance imaging (fMRI), indicate that this region is sensitive to the distinction between human and object movement (Pelphrey et al., 2003). Indeed, STSp responds more strongly to point-light displays of human movement than to point-light displays of moving objects (Beauchamp, Lee, Haxby, & Martin, 2003). MEG activity indicates that analyses of point-light displays of human and object movement diverge approximately 200 ms after stimulus onset when processes in the right temporal lobe, encompassing the STSp, are triggered during the perception of human movement (Virji-Babul, Cheung, Weeks, Kerns, & Shiffrar, 2007). Psychophysical studies indicate that typical observers demonstrate qualitatively different patterns of visual sensitivity to human motion and object motion (Shiffrar & Freyd, 1990, 1993; Shiffrar, Lichtey, & Heptulla-Chatterjee, 1997).

Recent studies have investigated the relationships between the visual perception of human and animal motions. Delayed focus on this issue may reflect a confusing nomenclature. The term “biological motion” is traditionally used to describe point-light displays of human motion (Johansson, 1973). Yet this term implies that the visual system analyzes all biological motions, whether human or animal, similarly. Naïve observers are capable of identifying and classifying animals depicted in dynamic point-light displays (e.g., Bellefeuille & Faubert, 1998; Mather & West, 1993; Pavlova & Sokolov, 2001; Pinto, 1994, 2006), increasing evidence indicates some differentiation in the perception of human and animal motion. For example, when infants view point-light displays of human and animal motion, their ability to differentiate phase-perturbed from canonically timed displays changes over the course of their development (Pinto, 2006). At the age of three months, infants are sensitive to phase differences in point-light human and animal motion. Just two months later, infants respond only to phase differences in upright human motion. This pattern of results suggests that the infant visual system becomes specialized or tuned for the detection of canonical human motion (Pinto, 2006). Specialized perceptual tuning is supported by fMRI data indicating that STSp activity becomes increasingly tuned to human motion as typical children age (Carter & Pelphrey, 2006). Assuming that such perceptual and neural tuning is maintained through adulthood, visual analyses of human and animal motions should diverge in adult observers.

Though PET data suggest significant overlap in the neural areas responsive to point-light human and animal motions (Pitito, Faubert, Gjedde, & Kupers, 2003), EEG measures indicate that the visual perception of human movement engages components of the mirror system while the perception of animal motion does not (Martineau & Cochin, 2003). Furthermore, fMRI data indicate that STSp activity is greater during the perception of human motion than during the perception of animal-like creature motion (Pyles, Garcia, Hoffman, & Grossman, 2007). Patterns of STSp activity are positively correlated with psychophysical measures of visual sensitivity to human and animal-like motions (Pyles et al., 2007). Taken together, these results suggest that increased neural activity in the STSp and mirror system may differentiate visual sensitivity to human and animal motions.

If patterns of neural activity differ during the visual perception of human motion and animal motion, as suggested by the neurophysiological evidence summarized above, then psychophysical profiles of the visual sensitivity to animal and human motions should also differ. The two psychophysical studies reported here tested this prediction.

Past research on the visual perception of point-light displays of human movement has focused on two core characteristics. First, the visual perception of human motion depends on a spatially global mechanism responsive to the overall configuration of the human form in motion (e.g., Ahlström et al., 1997; Bertenthal & Pinto, 1994; Cutting, Moore, & Morrison, 1988; Thornton et al., 1998). Second, the visual perception of human motion is orientation-specific (e.g., Bertenthal & Pinto, 1994; Pavlova & Sokolov, 2000; Shiffrar et al., 1997; Sumi, 1984).

**GLOBAL CONFIGURAL ANALYSIS**

Much of the evidence for spatiotemporally global analyses of human motion comes from tasks involving masked point-light displays. In this paradigm, observers view displays containing a point-light-defined person that is masked by the
superimposition of additional moving point lights. Point-light masks are typically constructed by duplicating the point-light figure and then randomizing the starting locations of each of those duplicate points. Since the size, luminance, and velocities of the duplicate points are unchanged, the motion of each point in the mask is identical to the motion of one of the points defining the person, as illustrated in Figure 2. As a result, only the global configuration of the points distinguishes the person from the mask. The fact that observers are able to detect the presence as well as the direction of an upright point-light person “hidden” within such a mask implies that the mechanism underlying the perception of human movement responds to configural information (Bertenthal & Pinto, 1994; Cutting et al., 1988). This conclusion converges nicely with the finding that STSp responsiveness reflects the configural organization of point-light displays of human motion (Bonda et al., 1996; Grossman et al., 2000).

**ORIENTATION DEPENDENCE**

The perception of human motion is orientation-specific. When a point-light human walker is turned upside-down, detection drops significantly (Bertenthal & Pinto, 1994; Pavlova & Solokov, 2000; Pinto & Shiffrar, 1999). Identification of coherent walkers (Sumi, 1984) and complex actions (Dittrich, 1993) also drops when point-light displays are inverted. By the age of five months, observers differentiate phase-typical from phase-perturbed walkers in upright, but not inverted, point-light displays (Pinto, 2006). STSp responsiveness to point-light walkers is also orientation-dependent (Grossman & Blake, 2001).

These two processing characteristics—global configural analysis and orientation dependence—have been interpreted as setting the perception of some stimuli—namely, faces and bodies—apart from the perception of other categories of stimuli (e.g., Bruce & Humphreys, 1994; Diamond &

---

**Figure 2.** Construction of a display in which a target figure was simultaneously masked with visual noise. (A) The human point-light display is composed of 11 visible elements, numbered here for identification. (B) Visual noise is created from two copies of the target figure by scrambling the locations of the points. Each point, as suggested by their numeric labels, maintains the individual motion of the original target points. As a result, the final display includes three copies of every element (e.g., three right shoulders). The target figure is distinguished from the noise only by the configuration of its points. The outline of a person represents the perceived organization of the moving elements. It does not appear in the actual displays.
Carey, 1986; Pinto & Shiffrar, 1999; Tanaka & Farah, 1993). This raises the question of whether visual sensitivity to animal movement is global and orientation-dependent to the same extent as visual sensitivity to human movement.

EXPERIMENT 1
Comparing the perception of human and horse gaits

Is the visual perception of animal movement subject to the same configural, orientation-specific processes that are triggered during the visual perception of human movement? To address this question, we compared the perception of human gait to the perception of horse gait. Human and horse gaits can be defined as nested hierarchies of pendula that are comparable in complexity. Both consist of four jointed limbs organized about a single principal axis. Previous studies have shown that adults can reliably recognize a point-light horse (Mather & West, 1993; Bellefeuille & Faubert, 1998) and that infants can categorize them as a quadruped (Arterberry & Bornstein, 2002).

The stimuli consisted of point-light depictions of a walking person and a walking horse (Figure 1a and 1b). Both target figures were masked with additional point lights so that the target figures, whether human or horse, could only be detected by global, configural analyses. To the extent that the visual perceptions of horse gait and human gait rely on similar neural mechanisms, one would expect to find similar levels of detection sensitivity and orientation-dependence in the horse and human conditions of this experiment.

Method

Participants

All 103 observers were students at the Newark campus of Rutgers University. Data from 13 additional participants were excluded because they needed, but failed to wear, corrective lenses. Participants received credit in a psychology course. None had prior experience with point-light displays or knowledge of the hypothesis under investigation.

Equipment and software

Stimuli were displayed on a Macintosh 21-inch (40 × 30 cm) RGB monitor set to 16 color planes, a 1152 × 870 pixel resolution and a 75 Hz refresh rate. Stimulus generation and presentation were controlled by custom software on a Macintosh Quadra 950 with a processor speed of 33 MHz. This equipment and software were used in both of the experiments reported here.

Stimuli and procedure

Human (Figure 1a). Following Cutting (1978), human gait was simulated as a hierarchy of nested pendula depicted with 11 points. The points mimicked the motions of the head, shoulder, elbows, wrists, hip, knees, and ankles of a walking person viewed sagittally. A single gait cycle was presented in 40 frames at a rate of 40 ms/frame. At its fullest extent, the figure measured 7.9 cm in height and 3.3 cm in width. The point-light person appeared to walk in place as if on a treadmill.

Horse (Figure 1b). The horse figure was depicted with 20 points. A single gait cycle was presented in 75 frames at a rate of 40 ms/frame. The horse figure was created from a digitized segment of a film. The film, a promotional videotape produced by the Rocky Mountain Horse Society, provided footage of animals demonstrating a typical walk. Rocky Mountain horses are spotted. Thus, the point-light horse display was created from the natural markings on a horse’s head, shoulders, hips, knees, and hooves. Using Adobe Premiere and Adobe PhotoShop, we isolated individual frames of the video sequence and identified the 2D coordinates of each point of articulation on the horse’s body. The figure measured 4.5 cm from hind hoof to hind hip and 6.3 cm from front hoof to the head. At its fullest extent, the hind and front hooves extended 6.6 cm horizontally. The horse appeared to walk in place.

Inversion. Inverted versions of the human and horse figures were created by rotating the coordinates 180° about a horizontal axis in the picture plane.

Masking. We employed a simultaneous masking paradigm to test for spatially global configural processing of the target figures. Observers were
presented with the dynamic target figures superimposed with visual noise, as illustrated in Figure 2. Each noise mask contained twice as many points as the corresponding target. Visual noise was created by copying the local motion vectors of the points defining the target, duplicating them, and then relocating them randomly within a 12.5 × 13.5 cm area in the center of the computer monitor. Thus, for each of the four target stimuli (horse or human, inverted or upright), the points defining the target and mask were identical in size, color, motion, and shape (Cutting et al., 1988). All points were white against a black background and subtended 0.2° of visual angle from the observer’s viewpoint.

Two types of trial were constructed from the masked displays, as follows.

Target-present trials. On each trial, the point-light target (human or horse, upright or inverted) was superimposed with the point-light mask. The location of the target figure was randomized within the mask. No part of the target was unmasked. On half of the target-present trials, the points defining the target depicted rightward facing and gait direction. On the other half of the trials, leftward facing and gait direction were depicted. Direction was randomized across trials. Target-present trials comprised 50% of the experimental trials.

Target-absent trials. In the target-absent trials, the points defining the target were replaced with an equal number of noise points. Each of these noise or positionally-scrambled points underwent the same motion as one of the points defining the target. Thus, only the global, configural organization of the points indicated whether a target was present or absent. The scrambled target points were located randomly within an area the approximate size of the target figure so that the target-present and target-absent displays could not be distinguished by systematic variations in the distribution of elements. Target-present trials comprised the other 50% of the experimental trials.

In both target-present and target-absent trials, trial duration equaled the amount of time needed for the target to complete two full gait cycles (3200 ms for the human and 6000 ms for the horse). Because horses and humans differ structurally, they are impossible to equate perfectly. Therefore, we “stacked the deck” against the human motion stimuli by depicting them with less information (fewer points) over less time (shorter stimulus duration) than the horse stimuli.

Each participant completed two blocks of 26 experimental trials. In a between-subjects design, participants were randomly assigned to one of the four stimulus conditions.

Procedure. Participants were tested individually in a quiet, dark room. At the start of each session, participants completed a brief questionnaire about their vision and well-being in order to identify any extraneous factors, such as extreme sleep deprivation or failure to wear one’s corrective lens, that might interfere with performance.

To accustom observers to the displays and the task, participants first completed a set of training trials with a point-light car. The masked point-light car stimuli were constructed in the same way as the masked point-light horse stimuli. At the start of training, observers were shown the point-light car unmasked and were asked to identify it. On the rare occasions when participants failed to identify the car, the experimenter identified it correctly for them. The experimenter then described the masked displays as looking like a “snowstorm” and described the observer’s task as judging whether or not the car was present in the snowstorm. Observers were told explicitly that the direction of the dynamic car’s heading would vary across trials but that the car in the mask would be identical to the unmasked car. Each observer then completed a 26-trial training set with the masked point-light car stimulus. Training was repeated if the observer’s performance did not exceed 75% correct, suggesting that the performance did not differ from chance by a binomial test at an alpha smaller than .05.

Subsequent to training, participants were familiarized with the unmasked target stimulus (human or horse) that they were to detect in the experimental trials. Participants viewed the target figure, presented upright, for two gait cycles and then identified it. If needed, the experimenter provided the correct label. The point-light horse or human gait was then presented a second time for the same duration as the first presentation. On request, observers were permitted up to two additional presentations of the stimulus. Immediately following this, each observer completed two blocks of experimental trials. They were instructed to decide whether their assigned target figure (upright human, inverted human, upright horse, or inverted horse) was present within the “snowstorm” on each trial. No feedback was given during the experimental trials.

Participants were seated comfortably at arm’s length, approximately 43 cm, from the display...
monitor. They used the computer keyboard to initiate each trial and to indicate whether or not they detected their target figure on each trial. Together, training and testing required 25–35 min. In this between-participants design, 22 observers completed the upright human condition, 24 the inverted human condition, 27 the upright horse condition, and 30 the inverted horse condition.

Results

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). Statistical analyses were computed with the mean performance across the two test blocks.

Were observers sensitive to the presence of the walking human and horse targets? The data are summarized in Figure 3a. Single-sample $t$-tests, conducted for each stimulus condition, showed that $d'$ in all conditions was significantly above chance (all $t$ values > 4.8, all $p$ values < .001). Above chance performance in all four conditions suggests that observers are capable of performing spatially global visual analyses of both human and horse motions when presented in either upright or inverted orientations.

To examine the effects of orientation on detection sensitivity, we computed an analysis of variance on the mean $d'$, using target and orientation as between-subjects factors. This analysis revealed main effects for both target and orientation and an interaction approaching statistical significance, $F(1, 99) = 3.74, p = .06$; partial $\eta^2 = .04$. The human figure was detected more accurately than the horse, $F(1, 99) = 13.8, p < .01$; partial $\eta^2 = .12$, and the upright figures more accurately than the inverted figures, $F(1, 99) = 25.14, p < .01$; partial $\eta^2 = .20$. Subsequent analyses for each target figure yield consistent patterns and confirm that orientation and target do not interact. Observers were reliably more sensitive to the upright human figure than to the inverted human figure, $F(1, 44) = 14.96, p < .01$; partial $\eta^2 = .25$. Similarly, observers detected the upright horse more accurately than the inverted horse, $F(1, 55) = 8.37, p < .01$; partial $\eta^2 = .23$. The orientation-specific patterns of responses to the human and horse figures are comparable, as suggested by the absence of a statistical interaction between figure and orientation.

Though observers received no feedback during the experimental trials, we examined changes in performance over time. We computed a repeated measures ANOVA to compare task performance during the first and second blocks of trials as a function of target and orientation. The omnibus analysis revealed a three-way interaction between blocks, target and orientation, $F(1, 99) = 6.67, p < .02$, a two-way interaction between block and target, $F(1, 99) = 6.24, p < .02$, and main effects of target, $F(1, 99) = 13.66, p < .01$, orientation, $F(1, 99) = 25.2, p < .01$, and block, $F(1, 99) = 6.99, p < .02$. Subsequent comparisons showed that
accuracy held steady in both horse motion conditions and in the inverted human condition (all $F$ values < 1), but increased from block 1 to block 2 in the upright human motion condition, $F(1, 21) = 18.65, p < .01$.

**Discussion**

Several conclusions can be drawn from the results of this experiment. Above chance levels of performance in all four conditions suggest that observers performed global configural motion analyses of both the human and horse displays in both upright and inverted orientations. Pavlova and Sokolov (2000) similarly reported above-chance detection performance with inverted point-light walkers in a mask. Nonetheless, the detection of the human figure was superior to detection of the horse figure. Thus, while the visual system employs global configural processing in the perception of both figures, something facilitates the perception of human motion.

Enhanced performance with human motion relative to horse motion occurred even though the horse stimuli were defined by more points (20 for the horse versus 11 for the human) and were displayed for longer trial durations (6000 ms for the horse versus 3200 ms for the human). If observers’ performance was determined by the amount of stimulus information available, detection of horse motion should have been superior to detection of human motion. Yet the reverse pattern of results was found. Given that we “stacked the deck” against performance in the human motion condition, the current results are likely underestimates of the perceptual superiority of human motion over horse motion.

Upright displays were detected more accurately than inverted displays for both the human and horse figures. Numerous studies of human motion perception have reported that visual sensitivity to canonical displays is superior to that of inverted displays (Bertenthal & Pinto, 1994; Shiffrar et al., 1997; Sumi, 1984). Indeed, most things are more difficult to perceive upside-down. Faces and bodies appear to be especially difficult, however. The size of the inversion effect appears to be, in part, a function of visual experience. Dog trainers show a larger inversion effect than non-experts in the perception of dogs’ bodies (Diamond & Carey, 1986). Visual experience also influences the size of the inversion effect with human faces (e.g., Hancock & Rhodes, 2008). Static images of human body postures also show a substantial inversion effect (Reed, Stone, Bozova, & Tanaka, 2003). One such study compared the size of the inversion effect for static postures of a human body and a dog body (Reed, Stone, & McGoldrick, 2006). In both cases, the postures were either frequently occurring (a bowing man and a begging dog) or rare (a begging man and a bowing dog). Like the present study, inversion effects were found for both human and dog body postures. Visual experience modulated the size of the inversion effect but was insufficient to account for all of the results. Thus, inversion effects can be understood, in part, as reflecting perceptual experience with a particular orientation. Experiment 2 described below examines the impact of immediate experience on animal and human movement perception.

A functional STSp and premotor cortex are required for the accurate perception of point-light displays of human gait (Grossman et al., 2005; Saygin, 2007). Two aspects of the current data set are consistent with patterns of neural activity in area STSp. First, STSp is more responsive during the perception of coherent human motion than scrambled human motion (Grossman et al., 2000) and during the perception of upright than inverted human gaits (Grossman & Blake, 2001). Second, STSp activity is greater during the perception of human motion than during the perception of creature motion (Pyles et al., 2007). Consistent with STSp activity, participants in the current study demonstrated global configural processing, greater visual sensitivity to upright than inverted displays, and greater visual sensitivity to human motion than to horse motion. This correlation between patterns of neural activity and psychophysical sensitivity measures is compatible with the hypothesis that STSp activity contributes to visual sensitivity to point-light displays of human movement (Pyles et al., 2007). Importantly, the perception of socially relevant stimuli depends on a vast network of neural areas that extends well beyond STSp (e.g., Pelphrey & Carter, 2008). As such, the current results cannot be understood as reflecting neural activity in any single area. For example, neural mechanisms underlying body schemata contribute to visual percepts of human movement (Stevens et al., 2000). And, as summarized below, motor areas also enhance visual sensitivity to human movement.

The current results suggest that the visual perception of human motion may not represent a qualitatively distinct phenomenon, but rather one
that differs from animal motion perception in a
graded fashion. This interpretation echoes pre-
vious findings. Despite substantial differences in
visual experience with moving dogs and seals, both
dog trainers and seal trainers demonstrate the
greatest visual sensitivity to point-light human
motion, significantly less sensitivity to point-light
dog motion, and the least sensitivity to point-light
seal motion (Cohen, 2002). Cohen (2002) conjec-
tured that visual sensitivity might reflect the degree
of structural similarity between an observed ani-
mal’s body and the observer’s own body such
that the greater the similarity, the greater is the
observer’s visual sensitivity to that animal’s ac-
tions. Consistent with this, visual percepts of
human body motion are also related to the
structural similarity between the observer’s body
and an observed person’s body (Funk, Shiffrar, &
Brugger, 2005). Theories of the mirror system
make a similar prediction. The mirror system is
thought to facilitate action perception by taking
advantage of the similarity between perceived
actions and the observer’s own motor repertoire
of possible actions (e.g., Calvo-Merino, Glaser,
Grezes, Passingham, & Haggard, 2005). The tigh-
ter the match between perceived and performable
actions, the greater is the visual sensitivity to the
perceived action (e.g., Prinz, 1997; Rizzolatti et al.,
2001; Shiffrar & Pinto, 2002; Viviani & Stucchi,
1992). Obviously, human observers can more
accurately perform actions performed by another
person than actions performed by a horse. If motor
processes facilitate the visual perception of motor
acts as the above theories predict, then one should
find greater motor system activity, and a corre-
sponding increase in visual sensitivity, when hu-
mans view actions performed by another human
than actions performed by a horse. Consistent with
these predictions, greater motor system activation
is found during the perception of human locomotion
than during the perception of horse locomotion
(Martineau & Cochin, 2003) and observers in
the current study demonstrated greater visual
sensitivity to human locomotion than to horse
locomotion.

Action performability is gravity-dependent.
When the movie of a point-light walker is inverted,
the depicted gait is rendered physically impossible.
Since the motor system cannot simulate actions
that it cannot perform, numerous theories predict
decrements in visual sensitivity to impossible
actions (e.g., Prinz, 1997; Rizzolatti et al., 2001;
Shiffrar & Pinto, 2002; Viviani & Stucchi, 1992).
Furthermore, motor system activity during the
perception of physically impossible human actions
drops to baseline (e.g., Stevens et al., 2000). STSp
activity also decreases when point-light displays
depict inverted, and thus impossible, actions
(Grossman & Blake, 2001). Consistent with these
patterns of neural activity, observers in the current
experiment showed significant decrements in their
visual sensitivity to inverted gaits. Thus, several
aspects of the current psychophysical results mir-
ror previously described patterns of neural activity
in social brain areas.

**EXPERIMENT 2**

**Priming the detection of point-light
gaits**

As described above, visual sensitivity to human
and horse locomotion is orientation-dependent.
Visual experience contributes to orientation-
dependent visual sensitivity to static human and
animal bodies (Reed et al., 2006). Visual sensi-
tivity to human motion is dependent on both
long-term visual experience (Jacobs et al., 2004)
and short-term priming (Verfaillie, 1993). Neural
activity in area STSp during the perception of
point-light displays of human movement is also
experience-dependent (Grossman, Blake, & Kim,
2004) in that STSp activity increases as observers
gain more visual experience with point-light dis-
plays of human motion. Age-dependent tuning of
STSp activity may also reflect the impact of long-
term visual experience (Carter & Pelphrey, 2006).

Previous research has shown that priming an
observer with an unmasked point-light human
walker in a particular orientation alters detection
at that orientation (Pavlova & Sokolov, 2000;
Verfaillie, 1993). Does recent visual experience
also influence visual sensitivity to animal motion?
An understanding of the impact of recent experi-
ence on visual sensitivity to point-light displays of
human and animal movement is needed to address
some apparently conflicting findings. A recent
psychophysical study suggests that visual sensitiv-
ity to human movement does not differ from visual
sensitivity to other types of structured movement
(Hiris, 2007). In a Herculean effort, the four
observers in Hiris’s (2007) experiment each com-
pleted 12,600 trials of a motion coherence discrimi-
nation task with masked displays. Across blocks
of trials, the target stimulus to be detected was a
dynamic point-light walker, a fixed view of a point-
light walker that rigidly translated, or a cloud of
points that rotated or translated. The four expert observers achieved amazing levels of performance (with mean $d'$ scores frequently approaching 5.0) and under some conditions, showed equivalent patterns of detection sensitivity to human and nonhuman motions. This result suggests that the visual analysis of human movement does not always differ from the visual analysis of nonhuman movement. How can this conclusion be integrated with differential patterns of neural activity during the perception of human motion and nonhuman motion (e.g., Beauchamp et al., 2003; Pelphrey et al., 2003; Pyles et al., 2007; Virji-Babul et al., 2007)?

Given the importance of visual experience, we wondered whether recent visual experience might be sufficient to manipulate visual sensitivities to human and nonhuman movements. To test this hypothesis, Experiment 1 was repeated with one modification. In this experiment, participants were primed with their unmasked target immediately before each trial. The orientations of the prime and target were always identical. To the extent that recent experience modifies psycho-physical measures of visual sensitivity to point-light displays, the addition of the prime should produce a pattern of performance that differs from that found in Experiment 1. The prime might enhance observers’ target detection in a manner roughly similar to the ways in which experts’ elaborated representations of targets and task constraints enhance their perceptual abilities (e.g., Ericsson & Lehmann, 1996).

Method

Participants

Seventy-six observers were students of Rutgers University in Newark. Five additional participants were eliminated because they failed to wear needed corrective lenses. Participants were recruited and selected by the same means and criteria employed in Experiment 1. None had participated in Experiment 1.

Stimuli and experimental design

As in Experiment 1, participants in this experiment performed a coherent motion detection task with point-light displays of horse gait or human gait depicted upright or inverted. As before, the coherent target was present on half of the trials and absent on half of the trials. However, unlike in Experiment 1, at the start of each trial the target figure appeared unmasked for 20 frames (800 ms). Immediately after this prime, the trial began with the target figure randomly relocated within the masked portion of the display. As before, there were two types of trial: target-present and target-absent. In the target-present trials, the coherent point-light figure presented in the prime was also present in the subsequent mask. The gait cycle of the walking figure continued to evolve smoothly across the mask onset transition. In the target-absent trials, the coherent figure in the prime was scrambled at the moment of mask onset. All other aspects of the stimulus displays and presentation were identical to those described in Experiment 1.

This experiment employed the same test procedure used in Experiment 1, including the training procedure with the point-light car. Stimulus conditions were assigned randomly between subjects. There were 21 participants in the upright human condition, 17 participants in the inverted human condition, 19 participants in the upright horse condition, and 19 participants in the inverted horse condition.

Results

When primed, could observers detect the presence of coherent horse and/or human motion within the point-light masks? The data are summarized in Figure 3b. Single-sample $t$-tests, conducted for each target figure and orientation, showed that $d'$ in all stimulus conditions was significantly above chance (all $t$ values > 6.3, all $p$ values < .001). Reliable detection of the targets suggests that, when primed, the visual system is sensitive to the global configuration of both human and horse movement in both canonical and inverted orientations.

To examine the effects of orientation on detection sensitivity, we computed an analysis of variance on the mean $d'$, using target figure and orientation as between-subjects factors. Detection performance did not vary as a function of orientation, $F(1, 72) = 0.55, ns$, or as a function of an interaction between orientation and target figure, $F(1, 72) = 0.03, ns$. To confirm this result, we separated the data by target figure and submitted the data from the human figure condition to an ANOVA with orientation as the between-participants factor. This analysis confirmed that orientation exerted no reliable influence on performance,
In the inverted conditions, recent experience (inverted person or horse) conflicted with past experience (upright people and horses). This suggests that short-term experience can facilitate target detection when stimuli diverge significantly from an observer’s long-term representations of the corresponding events. Such priming-dependent facilitation likely reflects the deployment of task-dependent strategies and stimulus representations that are not commonly employed outside the task itself. If so, this might explain why participants in the motion coherence task of Hiris (2007) did not show enhanced visual sensitivity to human motion.

How can enhanced detection of the inverted, and thus physically impossible, human motion be understood in relation to previous suggestions that impossible human movements cannot be readily primed (Kourtzi & Shiffrar, 1999; Pavlova & Sokolov, 2000)? In the current masking study, each participant viewed only one stimulus in only one orientation. This was done to eliminate the possibility of priming across orientations and figures. As a consequence, participants may have been able to detect targets by matching a single template of the figure. Extensive experience can also facilitate the use of rapid template-matching strategies in perceptual discriminations of biological and non-biological bodies (e.g., Biederman & Shiffrar, 1987). In previous priming studies, participants judged a wide variety of forms (Kourtzi & Shiffrar, 1999) and orientations (Pavlova & Sokolov, 2000) across trials or blocks of trials. The use of variable stimuli likely promotes the use of detection strategies that differ from those employed in the current experiment.

In sum, the results of this experiment indicate that visual sensitivity to human motion may not reflect obligatory processing constraints.

Discussion

The results of this experiment converge with the results of Experiment 1 in indicating that visual sensitivity to human motion is greater than visual sensitivity to animal, or at least horse, motion. Thus, enhanced visual sensitivity to canonical human movement appears to be a robust phenomenon.

Previous studies have demonstrated that point-light displays of human gait are subject to reliable short-term priming effects (Verfaillie, 1993). The results of this experiment indicate that the detection of animal motion is also prime-dependent and that orientation-dependent performance with point-light displays can be eliminated by recent experience. Thus, orientation dependence cannot be used as a universal “gold standard” characteristic of visual analyses of human motion.

Interestingly, the addition of the prime facilitated detection of the inverted targets but not the upright targets. In the upright conditions, recent experience provided by the prime (upright person or upright horse) was consistent with long-term, real-world experience (upright horses and people).
information (e.g., Brothers, 1997; Pelphrey & Carter, 2008; Thompson & Hardee, 2008). The goal of the psychophysical work described here was to test whether such differentiated neural analyses of social and nonsocial information have perceptual consequences. Specifically, are human observers better able to detect the presence of human motion, a socially relevant event, than the presence of an equally complex but nonhuman motion?

To address that question, observers in Experiment 1 viewed masked displays of human gait or horse gait, presented upright or inverted. In all cases, observers performed the same perceptual task; namely, detection of coherent motion. Mean detection performance was always above chance. Because task performance required global analyses, above-chance performance indicates that observers were able to globally process both human motion and horse motion. In both cases, detection performance was better with upright than inverted targets. Importantly, detection performance with human motion was superior to detection performance with horse motion. This performance difference is consistent with the prediction of social brain theories since observers demonstrated greater visual sensitivity to a socially relevant event. However, above-chance performance and orientation-dependence in the detection of both human and horse motions suggest that at least some of the same global motion processes contributed to both.

Experiment 2 investigated the role of recent experience in detection performance. When an unmasked point-light prime appeared at the onset of each trial, detection performance was orientation-independent for both human and horse motion. In the absence of a prime, observers must rely on long-term representations of a target in order to detect it. Under those conditions, detection of upright targets was significantly better than detection of inverted targets. Priming facilitated detection of inverted displays that were inconsistent with observers’ long-term representations of the target (e.g., horses walking on a ceiling). Thus, long-term and short-term experience differentially impact visual sensitivity to human and animal motions. Yet even when short-term priming could be used to facilitate target detection, something still enhanced visual sensitivity to human motion over horse motion.

Neurophysiological studies have identified a network of areas involved in the visual perception of human movement (e.g., see Blake & Shiffrar, 2007; Puce & Perrett, 2003 for review). The detection of point-light-defined human actions requires functional processing within area STSp and premotor cortex (Saygin, 2007). STSp activity during the perception of point-light human motion reflects several characteristics, such as configural processing (Thompson, Clarke, Stewart, & Puce, 2005) and orientation dependence (Grossman & Blake, 2001), of psychophysical measures of visual sensitivity to point-light human motion. Indeed, STSp activation increases as observers become more sensitive to point-light human actions (Grossman et al., 2004). Furthermore, the STS region is more responsive during visual perception of human motion than during the perception of either meaningful object movement (Pelphrey et al., 2003) or animal movement (Pyles et al., 2007). Similarly, activity in the mirror system is greater during the perception of human movement than horse movement (Martineau & Cochin, 2003). Of course, the current psychophysical results cannot be used to identify the precise contributions of STSp, the mirror system, or any other neural area to motion coherence detection. Nonetheless, perceptual indicators of enhanced visual sensitivity to human motion are consistent with the hypothesis that social brain areas, including the STS region and the mirror system, respond differentially during the perception of human motion and animal motion.

Movements of the human body convey extensive social information and the typical human visual system is well equipped for the detection of those social cues (Shiffrar, Kaiser & Chouchouroulou, 2009). Observers with autism spectrum disorder (ASD) are compromised in their social abilities including the comprehension of bodily gestures (American Psychiatric Association, 2006). These individuals also exhibit selective deficits in their visual sensitivity to point-light displays of human, but not object, movement (Kaiser, Delmofino & Shiffrar, 2009; Kaiser & Shiffrar, 2009). Interestingly, the structure and function of STSp are compromised in observers with ASD (e.g., Boddart et al., 2004; Freitag et al., 2008; Zilbovicius et al., 2006), as is the mirror system (e.g., Hadjikhani, Joseph, Snyder, & Tager-Flusberg, 2006), although that point is controversial (e.g., Hamilton, Brindley & Frith, 2007). In any case, there appears to be an intriguing co-occurrence of abnormalities in the social brain areas related to action perception and selective deficits in visual sensitivity to other people’s actions. To the extent that areas of the social brain are selectively
tuned for the detection of human motion (e.g., Carter & Pelphrey, 2006; Pyles et al., 2007), observers with ASD should not exhibit greater visual sensitivity to human motion than to animal motion. We are currently testing this hypothesis.

References


