

Learning to discriminate complex movements: Biological versus artificial trajectories

Jan Jastorff

Laboratory for Action Representation and Learning,
Department of Cognitive Neurology,
Hertie Institute for Clinical Brain Research,
University Clinic, Tübingen, Germany



Zoe Kourtzi

Max Planck Institute for Biological Cybernetics,
Tübingen, Germany, &
Birmingham University, School of Psychology,
Edgbaston, Birmingham B15 2TT, United Kingdom



Martin A. Giese

Laboratory for Action Representation and Learning,
Department of Cognitive Neurology,
Hertie Institute for Clinical Brain Research,
University Clinic, Tübingen, Germany



The recognition of complex body movements and actions is a fundamental visual capacity very important for social communication. It seems possible that movement recognition is based on a general capability of the visual system to learn complex visual motion patterns. Alternatively, this visual function might exploit specialized mechanisms for the analysis of biologically relevant movements, for example, of humans or animals. To investigate this question, we trained human observers to discriminate novel motion patterns that were generated, exploiting a new technique for stimulus generation by motion morphing. We tested the learning of different classes of novel movement stimuli. One group of stimuli was fully consistent with human movements. A second class of stimuli was based on artificial skeleton models that were inconsistent with human and animal bodies. A third group of stimuli specified the same local motion information as human movements but was inconsistent with an underlying articulated shape. Participants learned both classes of articulated movements very fast in an orientation-dependent manner. Learning speed and accuracy were strikingly similar and independent of the similarity of the stimuli with biologically relevant body shapes. For the class of stimuli without underlying articulated shape, however, we did not observe significant improvements of the discrimination performance after training. Our results indicate the existence of a fast visual learning process for complex articulated movement patterns, which likely is relevant for biological motion perception. This process seems to operate independently of the consistency of the patterns with biologically relevant body shapes but seems to require the compatibility of the learned movements with a global underlying shape.

Keywords: biological motion, perceptual learning, motion morphing, movement recognition, visual learning, discrimination

Introduction

The ability to recognize complex movements and actions is critical for the survival of many species. Consequently, the human visual system is very skilled in the extraction of information from movements, even for strongly impoverished stimuli like point-light displays (Johansson, 1973). Neonates are able to imitate facial and manual gestures (Meltzoff & Moore, 1977), suggesting that the recognition of complex movements might, at least partially, depend on innate mechanisms for the processing of biologically important human movements (Fox & McDaniel, 1982).

However, motion recognition might also critically depend on visual learning. A dependence on learning has been demonstrated for a broad range of visual tasks ranging from lower level vision tasks like orientation discrimina-

tion, hyperacuity, or direction discrimination (e.g., Ball & Sekuler, 1987; Mayer, 1983; Poggio, Fahle, & Edelman, 1992) to higher level tasks such as face or object recognition (see Fine & Jacobs, 2002; Goldstone, 1998; Tarr & Bülthoff, 1998, for a review). Studies on object recognition indicate that observers are able to learn novel complex shapes, exploiting representations that show view dependence (Edelman & Bülthoff, 1992). Moreover, neurophysiological experiments in monkeys support the existence of neurons in the inferotemporal cortex, which learn to respond selectively to novel complex three-dimensional shapes (Logothetis, Pauls, & Poggio, 1995). Many of these neurons show view-dependent tuning.

The central role of learning in the visual recognition of complex shapes motivates the hypothesis that the recognition of complex movement patterns might also be based on learning. Evidence supporting this hypothesis was

provided by studies showing that human observers learn to recognize individuals based on their facial or full-body movements (e.g., Hill & Pollick, 2000; Kozlowski & Cutting, 1977; O'Toole, Roark, & Abdi, 2002; Troje, Westhoff, & Lavrov, 2005). Moreover, the detection of point-light walkers in dynamic noise can be improved through visual learning (Grossman, Blake, & Kim, 2004; Hiris, Krebeck, Edmonds, & Stout, 2005). Furthermore, the recognition of biological motion is dependent on stimulus orientation, like the recognition of stationary objects (Bertenthal, Proffitt, & Kramer, 1987; Pavlova & Sokolov, 2000; Sumi, 1984). Consistent with these psychophysical findings, biological-motion-sensitive neurons in the superior temporal sulcus (STS) of monkeys show view-dependent modulation of their firing rate (Perrett et al., 1985), and imaging studies indicate reduced fMRI activity in human STS for the presentation of inverted point-light walkers (Grossman & Blake, 2001). This suggests that complex movements and static shapes might be encoded by similar orientation-dependent and, potentially, view-dependent mechanisms (Verfaillie, De Troy, & Van Rensbergen, 1994). That such learning mechanisms provide a computationally powerful explanation of biological motion recognition is suggested by theoretical models that account for a variety of experimental results (Giese, 2000; Giese & Poggio, 2003; Lee & Wong, 2004).

This study investigates the learning of complex motion patterns that were either biologically relevant human-like actions or artificial movements. Two classes of artificial patterns were tested. One class of stimuli was articulated movements based on skeleton models without biological relevance. The other class of artificial stimuli had the same local motion information as human motion but was inconsistent with an underlying skeleton. Opposed to the human-like stimuli, the artificial stimuli were not interpreted consistently as biological actions that are executed by humans or animals.

As paradigm for the study of learning, we chose the discrimination learning of motion patterns because this allowed us a relatively accurate matching of the complexity and the low-level properties of the different stimulus classes. A control of the low-level properties was important because we wanted to study the learning of global motion patterns, rather than discrimination based on specific salient local cues or features (like average or maximum speed of individual dots). To generate stimuli with highly controlled spatiotemporal similarity properties, we exploited motion morphing and a new method for the approximation of motion morphs by real human movements. Our experiments tried in particular to address which properties of complex motion stimuli are critical for fast learning and for the generalization to similar stimuli.

We conducted four main experiments and two control experiments. [Experiment 1](#) shows that humans can learn novel articulated movements very quickly. In addition, it demonstrates that human-like and artificial articulated

movements are learned equally fast and accurately. The similarity of the learning of these two stimulus classes was confirmed by [Experiment 2](#), which demonstrates that the learned representation for both stimulus types is orientation dependent, like normal biological motion perception. [Experiment 3](#) rules out the possibility that the observed similarity between human-like and artificial articulated movement patterns is due to the fact that both stimulus classes were generated by motion morphing. Consistent with [Experiment 1](#), we obtained striking similarities between the learning of (nonmorphed) real human movements and artificial patterns. [Experiment 4](#) finally demonstrates that the presence of an underlying skeleton seems crucial for fast learning and robust generalization.

Methods

Participants

A total of 34 individuals (15 male, 19 female; mean age, 27.6 years) participated in the experiments (11 in [Experiment 1](#), 9 in [Experiment 2](#), 7 in [Experiment 3](#), and 7 in [Experiment 4](#)). All participants had normal or corrected-to-normal vision. Many of them had participated in psychophysical experiments before, but no one had ever been exposed to the same or similar morphing stimuli. Participants were tested individually, gave written informed consent to participate in the study, and were paid for participation.

Stimuli

Visual stimulus presentation

The stimuli were presented as point-light walkers consisting of 10 dots. The dot trajectories were generated by motion morphing (see below) between three prototype trajectories (natural human or artificial movements). The stimuli were displayed using an Apple Macintosh G4 computer and a Sony color monitor (75 Hz frame rate; $1,024 \times 768$ pixel resolution) that was viewed binocularly from a distance of 40 cm. Stimulus presentation and recording of the participants' responses were accomplished using the MATLAB Psychophysics Toolbox (Brainard, 1997). The stimuli were shown as black dots on a gray background, and each dot had a diameter of 0.5 deg visual angle ([Figure 1](#)).

To prevent participants from using low-level strategies for accomplishing the task, the stimulus dots were not presented on the major joints. Instead, for every frame, the dot positions were chosen randomly and uniformly distributed along the bone segments that were immediately adjacent to the relevant joint (cf. Beintema & Lappe, 2002, for a similar stimulus manipulation). The maximum

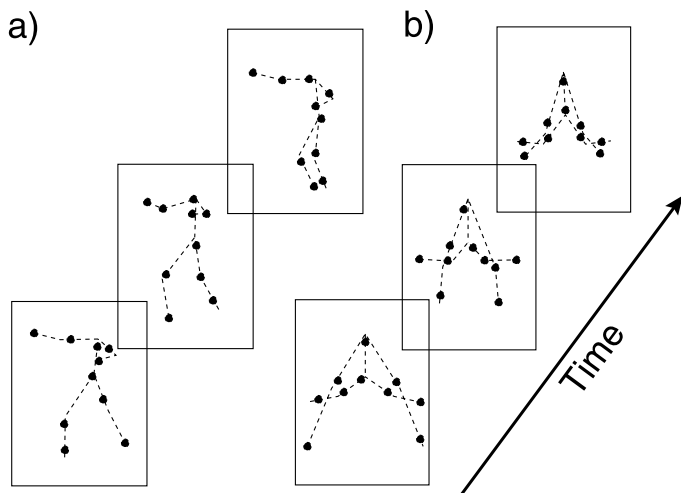


Figure 1. Human-like and artificial articulated stimuli. Snapshots from the sequence of a human-like (a) and an artificial articulated (b) point-light stimulus. The points were not displayed directly on the joint but randomly shifted along the bones of the skeleton in each frame of the animation (dashed lines not shown during the experiment). Examples are also linked as auxiliary [Movie 1](#) and auxiliary [Movie 2](#).

displacements were 30% of the bone length away from the joints. This manipulation does not disrupt the perception of biological motion. However, we have shown in previous experiments that it efficiently prohibits the use of local strategies, like comparing the relative positions of individual dot pairs (Jastorff, Kourtzi, & Giese, 2003). Also, it makes it very difficult to memorize dot configurations of individual frames. The size of the stimuli was 5×10 deg, and their position was randomized within an area of ± 2 deg horizontally and vertically.

In each trial, we presented first a movement stimulus for four gait cycles, followed by a 500-ms blank interval and the second movement stimulus, which was also presented for four cycles. Each gait cycle lasted for 1,300 ms, resulting in a total duration of 10.9 s per trial.

Tracking of human movements

Prototypical human movement trajectories were obtained by tracking the two-dimensional joint positions in videos showing an actor performing different movements. The actor moved along a line that was orthogonal to the view direction of the camera (side view). All movements were executed periodically for multiple cycles, but only a single cycle of each movement was used for motion morphing. First, the translation of the body center was subtracted by fitting the translation of the hip with a linear function of time. The resulting movement looks like a person performing the movement on a treadmill. The body points that were tracked manually were the head, shoulders, elbows, wrists, hip, knees, and ankles. For the generation of the point-light stimuli, the positions of the shoulder and the

head markers were averaged, resulting in a stimulus with 10 dots. The tracked trajectories were time-normalized and smoothed by fitting them with a second-order Fourier series. The resulting periodic trajectories served as prototypes for the generation of human-like novel movements. The prototypes of the human movements are listed in [Table 1](#). Some of the movements were right–left symmetric, while others were asymmetric.

Generation of artificial articulated movements

Prototypical trajectories for the artificial stimuli were generated by animation of multiple artificial skeleton models with nine bone segments that were linked in the same way as a human skeleton ([Figure 1b](#)). The shapes of these skeletons were chosen to be highly dissimilar from biologically relevant body structures. The trajectories of the joint angles $\alpha_n(t)$ of the eight joints of the skeletons were given by sinusoidal functions of time:

$$\alpha_n(t) = a_n + b_n \sin(\omega t + \phi_n).$$

The frequency, ω , and amplitudes, b_n , were matched with typical values of joint trajectories of human actors during natural movements. In this way, low-level properties and the complexity of the human and artificial prototype movements were coarsely balanced.

Motion morphing

Our stimuli were generated by motion morphing between triples of prototypical trajectories. These triples were chosen either from the human or from the artificial articulated movements. Motion morphing algorithms interpolate between trajectories, resulting in new trajectories that blend between the style properties of the prototypes (e.g., Bruderlin & Williams, 1995; Wiley & Hahn, 1997). Recently, such methods have been exploited to generate stimuli for psychophysical experiments (Giese & Lappe, 2002; Troje, 2002). For our experiment, we applied a method that creates new trajectories by *linear combination* of three prototypical trajectories in space–time (Giese & Poggio, 2000). It has been shown that this method results in natural-looking morphs for interpolations between different natural gaits and also interpolates between the

Group	Prototype		
1	Out-to-in crescent kick	Front kick	Rock n' roll kick
2	Forward jumping jack	Skating	Sirtaki step
3	Front punch	Knife hand strike	Hooking punch
4	Marching	Limping	Running

Table 1. Description of the movements that were linearly combined for the generation of the human-like stimulus groups.

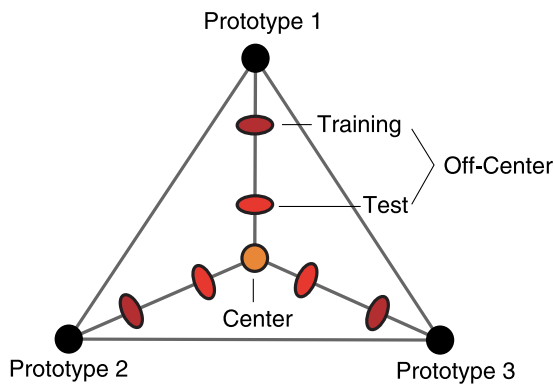


Figure 2. Pattern space defined by motion morphing. Morphs were generated by linear combination of the joint trajectories of three prototypical patterns (Prototypes 1 to 3). Three groups of stimuli were generated by choosing different combinations of linear weights: Center stimuli with equal weights of all three prototypes, Test Off-Center stimuli with one prototype weighted slightly higher than the others, and Training Off-Center stimuli for which the weight of one prototype significantly exceeded the equal weights of the other two.

perceived properties of such patterns (see Giese & Lappe, 2002, for details). Formally, the morphs were given by the equation

$$\begin{aligned} \text{New motion pattern} &= c_1(\text{Prototype 1}) \\ &+ c_2(\text{Prototype 2}) \\ &+ c_3(\text{Prototype 3}) \end{aligned}$$

where the weights c_i determine how much the individual prototypes contribute to the morph. When the weight of a prototype is high, the linear combination strongly resembles this prototype. (Weight combinations always fulfilled $c_1 + c_2 + c_3 = 1$). The weight vectors (c_1 , c_2 , and c_3) define a metric space, and the distances between these vectors provide a measure for the perceived spatiotemporal similarity of the corresponding trajectories. This had been verified in previous studies by analysis of perceptual ratings for motion morphs and by applying multidimensional scaling for the reconstruction of a perceptual metrics from similarity judgments for pairs of such motion morphs (Giese & Lappe, 2002; Giese, Thornton, & Edelman, 2003). By varying the distance between the weight vectors, we were able to control the difficulty of the stimulus discrimination precisely.

Using this algorithm, we generated three different stimulus groups for each type of movement (human-like and artificial patterns). The first group, called *Center* stimuli in the following, is characterized by equal weights of all prototypes in the linear combination ($c_1 = c_2 = c_3$). The other two groups were characterized by nonequal weights of the prototypes in the motion morph. For the group of *Test Off-Center* stimuli, the weights for one prototype only slightly exceeded the weights of the other

two ($c_i > c_j = c_k$, with i , j , and k defining the three prototypes). For the third group, called *Training Off-Center* stimuli in the following, one prototype was weighted significantly higher than the others ($c_i \gg c_j = c_k$; Figure 2). The weights for the Test Off-Center stimuli were adjusted, based on a preceding pilot experiment ensuring that average performance was about 60% correct discriminations before training. The largest weight of the Training Off-Center stimulus was then set to 1.25 times the largest weight of the corresponding Test Off-Center stimulus (maintaining $c_1 + c_2 + c_3 = 1$).

We confirmed that the physical distances between the trajectories of the Center and Test Off-Center stimuli for human-like and artificial stimuli were comparable by computing the mean Euclidean distance between the dot trajectories (0.073 for human-like and scrambled human-like stimuli and 0.085 for the artificial patterns). This makes it unlikely that our results can be explained by simple, low-level motion or spatial differences between the dot trajectories.

Selection of the prototypes

The prototypes for each group of the human-like movements were carefully chosen to guarantee that the resulting morphs always looked in such a way that a human actor could execute them. In a pilot experiment, we tested 21 human movements, presented as side views, to determine triples of prototypes that resulted in smooth-looking morphs. We collected naturalness ratings (scale: 1, *unnatural*; 5, *natural*) for each of the morphed stimuli, and only the stimuli with the highest naturalness ratings (Levels 4 or 5) were selected for the generation of the human-like motion morphs. This procedure led to the selection of four triplets of human movements: one contains locomotion patterns and the rest include boxing movements, aerobics movements, and kicks (see Table 1). We verified that the resulting morphs were consistently interpreted as human movements by naive observers.

Twelve prototypical artificial movements were generated and selected based on the requirement that the resulting morphs did not contain any salient visual cues (e.g., large accelerations and bizarre asymmetries). Similar to the prototypes for natural human movements, these prototypes were divided into four groups containing three prototypes. The prototypes within the same group differed gradually with respect to the length of the bones of the skeleton and the joint angle trajectories. Prototypes in different groups were more dissimilar with respect to these parameters. Two of the groups contained mirror-symmetric prototypes, and the other two groups contained asymmetric patterns.

In a pilot experiment, we verified that naive participants were not able to provide consistent interpretations for any of these artificial movements. Most common interpretations were the following: peculiar animal movements (38% of observers), dynamic mechanical devices (21% of observers), and articulated shapes without any particular global

interpretation (35% of observers). Examples of the different stimuli are provided in [Figure 1](#) and as Auxiliary [Movies 1](#) (human like) and [2](#) (artificial articulated).

Procedure

For each observer, two human-like stimulus groups and two artificial stimulus groups were chosen randomly out of the four groups of each type. The experiment started with a brief practice session of four trials, showing four example stimuli without feedback (two human-like and two artificial stimuli).

In all trials, participants had to compare two movement stimuli in a pair comparison paradigm. Each trial started with the presentation of a Center stimulus followed either by the same Center stimulus or by an Off-Center stimulus from the same group that was randomly selected. Both stimuli were presented for four gait cycles. The random selection of the Off-Center stimulus reduced the presentation of the individual Off-Center stimuli, decreasing the amount of spurious learning. In a two-alternative forced-choice paradigm, participants had to report whether the second stimulus matched the first one. They were instructed to respond immediately after making their decision. However, no explicit time constraint was imposed.

The experiment consisted of three test blocks that were interleaved by two training blocks. In the test blocks, Center stimuli had to be discriminated from Test Off-Center stimuli (difficult discrimination, see [Figure 2](#)). Each stimulus group was presented three times in random order, resulting in 12 trials overall. During test trials, no feedback about correct discrimination was provided. The training blocks consisted of 32 trials (8 repetitions per stimulus group). During training, participants had to discriminate between Center and Training Off-Center stimuli (simple discrimination, see [Figure 2](#)) and received feedback about their performance. By testing the transfer between Training Off-Center and Test Off-Center stimuli, we were able to verify whether the learned representation generalized to similar stimuli.

The results were analyzed using repeated measures analyses of variance (ANOVAs), applying a Greenhouse–Geisser correction for nonsphericity.

Results

Experiment 1: Learning of human-like versus artificial articulated movements

The first experiment compared discrimination learning between two types of articulated motion stimuli: movements that closely resembled human movements ([Movie 1](#)) and movements based on artificial skeleton models that were quite dissimilar from biologically relevant move-

ments of humans or animals and were, thus, not consistently perceived as such movements ([Movie 2](#)). With this experiment, we tried to clarify two questions: (1) Are humans able to learn the discrimination between artificial articulated motion patterns without immediate biological interpretation, and how fast is this learning? (2) Is there a difference between the learning of biologically relevant and artificial articulated movements?

Participants were trained with two types of point-light stimuli: human-like stimuli, generated by morphing between real human movements, and artificial articulated movements, generated by morphing between trajectories that were generated from artificial skeleton models (see [Methods](#) section). If the human visual system contains special mechanisms for the learning of biologically important movements, one would expect that learning of human-like patterns should be faster and potentially more accurate than the learning of completely artificial patterns. If, however, the visual system is disposing of a general mechanism for the learning of articulated movements that is independent of their biological importance, no difference would be expected.

Results and discussion

Participants perceived the human-like stimuli as human movements, whereas the artificial articulated patterns typically resulted in very inconsistent interpretations between participants (e.g., “mechanical device” or “weird spider,” see [Methods](#) section). [Figure 3a](#) shows the discrimination performance (percentage correct) for the human-like stimuli (dark gray) and the artificial articulated patterns (light gray) for the three test blocks. Starting close to chance level, participants show very similar improvements of discrimination performance for both stimulus types. Two training blocks with only 16 repetitions of each Center stimulus were sufficient to improve performance to a level above 80% correct responses for both stimulus types. A two-way repeated measures ANOVA reveals a significant main effect of the number of the test block, $F(2,20) = 9.2$, $p < .01$. Neither the main effect of stimulus type, $F(1,10) < 1$, $p = .91$, nor the interaction was significant, $F(2,20) < 1$, $p = .76$. An additional ANOVA tested the influence of the stimulus group (triple of prototypes) within the human-like stimuli and within the artificial stimuli. This analysis verified that the learning effects did not depend significantly on the specific stimulus group (nonsignificant interaction between stimulus group and number of the test block). This result verifies that the matching of task difficulty before training between the stimulus groups was efficient.

[Figure 3b](#) shows the response times for the two stimulus types. Consistent with the increase in performance, the response times decrease after training for both stimulus types in a very similar way. This observation was confirmed by an ANOVA showing a significant main effect of the number of test blocks, $F(2,20) = 5.5$, $p < .05$,

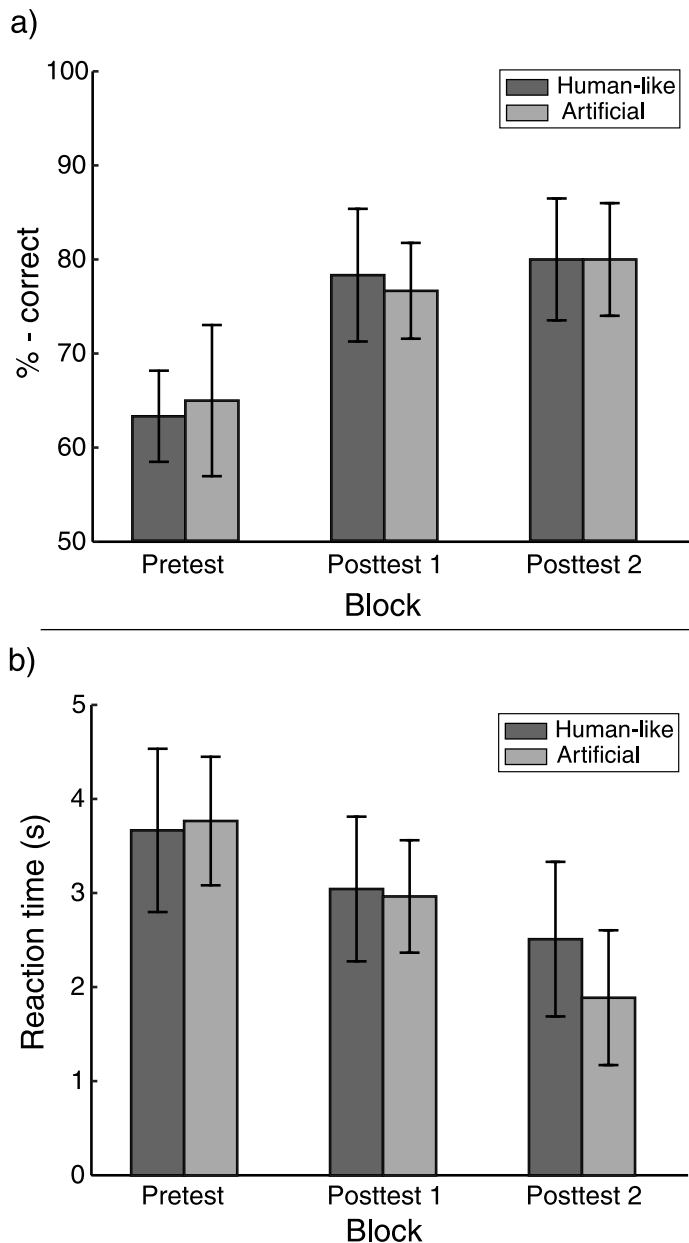


Figure 3. Discrimination learning for human-like versus artificial articulated movements. Panel a shows the mean percentages of correct responses for Experiment 1 (\pm SEM) in the three test blocks for the human-like stimuli (dark gray) and for artificial articulated movements (light gray). Panel b shows the mean reaction times (\pm SEM) in the three test blocks measured after the presentation of the first stimulus separately for human-like and the artificial articulated movements ($n = 11$).

but no significant influence of stimulus type, $F(1,10) < 1$, $p = .63$, and no significant interaction, $F(2,20) = 2.4$, $p = .12$.

These results indicate that the human visual system disposes of a learning mechanism for articulated movement patterns, which works equally well for biologically important and artificial articulated patterns. Furthermore, less

than 20 stimulus repetitions during training were sufficient for participants to increase their performance significantly. This demonstrates relatively fast visual learning of such articulated movement patterns. A similar number of stimulus repetitions seems to be required for the learning of the detection of novel point-light stimuli in noise (Hiris et al., 2005).

Because the discrimination during the test blocks of our experiments (discrimination between Center and Test Off-Center stimuli) was more difficult than the discrimination during the training blocks (discrimination between Center and Training Off-Center stimuli, see Methods section), successful learning of a simpler discrimination facilitated the more difficult discrimination during the test blocks. The same phenomenon has been reported in other perceptual learning experiments (e.g., Ahissar & Hochstein, 1997; Liu & Weinshall, 2000; Mackintosh, 1974). In addition, this result demonstrates that observers were able to generalize from the learned representation to similar novel movement stimuli.

To further investigate the properties of this generalization, we conducted two additional control experiments. The first experiment tested the transfer of the learning to novel untrained stimuli generated from different prototypical movements. For every participant, the four stimulus groups of the human-like and the artificial patterns were randomly divided into two sets. The first set was based on two triples of prototypes of the human-like and the artificial articulated movements. The second set contained the remaining two triples of prototypes of each stimulus class. The stimuli of the first set were presented throughout the experiment (e.g., in the test blocks and the training blocks). The second set was presented only during the test blocks and served for the testing of generalization. If learning is unspecific for the trained movements, one would expect transfer to the novel stimuli of the second set, although they were derived from different prototypes and thus represent a different set of actions. If, however, learning is specific for the trained movements and shows only generalization to similar actions, transfer would be expected only within the trained stimulus groups.

Figure 4 shows the results of this first control experiment. We only observed improvements in performance for the trained stimuli but not for the novel ones. This was confirmed by a three-way ANOVA that revealed a significant main effect of the number of the test block, $F(2,16) = 4.2$, $p < .05$, and a marginally significant interaction, $F(2,16) = 3.5$, $p = .06$, between the number of test block and the familiarity (Trained vs. Novel stimuli). An additional contrast analysis based on the marginally significant interaction revealed significant improvements for the trained groups, human-like: $F(1,8) = 4.7$, $p < .05$; artificial articulated: $F(1,8) = 21.4$, $p < .01$, but not for the untrained groups, human-like: $F(1,8) < 1$, $p = .47$; artificial articulated: $F(1,8) < 1$, $p = .98$. This result implies that the learning was specific for the trained motion patterns and did not transfer to untrained patterns that were generated

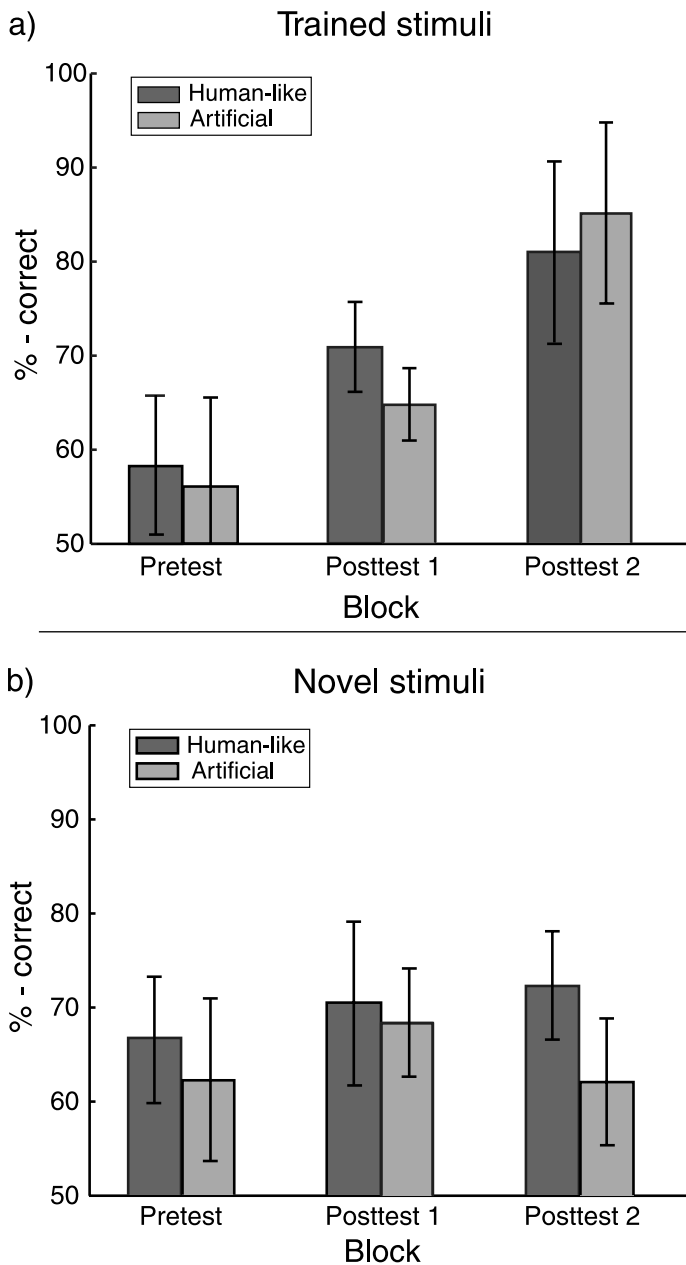


Figure 4. Control experiment testing generalization to novel stimuli generated from different prototypes. The figure shows the mean percentages of correct responses (\pm SEM) for the trained stimuli (Panel a) and the novel stimuli (Panel b) in the three test blocks. The human-like stimuli are presented in dark gray, and the artificial articulated movements are presented in light gray.

from different prototypical movements. This shows that the selectivity of the learned representation is sufficient for distinguishing different action categories, for example, boxing and kicking. Moreover, this result rules out that our results can be explained by general factors, like increasing familiarity with the task or increasing efficiency of the processing of biological motion.

The second control experiment tested the necessity of the training blocks with feedback for the learning process. The

three test blocks were presented without intermediate training. In this case, discrimination performance did not increase significantly (no main effect of the number of the test block), $F(2,18) < 1$, $p = .91$. Similar to [Experiment 1](#), we did not observe significant differences between human-like and artificial articulated stimuli in this control experiment: no main effect of stimulus type, $F(1,9) < 1$, $p = .93$, and no significant interaction, $F(2,18) < 1$, $p = .46$. This implies that training together with feedback seems to be essential for the fast improvement during learning. In addition, this experiment provides additional evidence against the explanation of the observed changes by unspecific familiarity or practice effects.

Experiment 2: Orientation dependence of the learned representation

A characteristic property of the recognition of biological motion is its strong orientation and view dependence. Rotation of point-light walkers in the image plane against the familiar upright orientation substantially degrades recognition performance (Bertenthal et al., 1987; Pavlova & Sokolov, 2000; Sumi, 1984). Likewise, it has been shown that recognition of point-light walkers, as well as the response of biological motion-selective neurons in monkey cortex, is modulated by rotation of the stimuli in depth (Bülthoff, Bülthoff, & Sinha, 1998; Oram & Perrett, 1996).

[Experiment 2](#) tested whether the learned representations of novel human-like and artificial articulated patterns show orientation dependence, like the recognition of natural biological motion patterns. To test orientation dependence, we modified [Experiment 1](#) by training the participants with stimuli that were rotated by 90 deg in the image plane against the test stimuli, which were presented upright.

Results and discussion

An analysis of the performance during the training blocks shows that participants were able to learn the rotated stimuli in the same way as the stimuli presented upright in [Experiment 1](#). Performance increased significantly from the first to the second training block: main effect of training block, $F(1,8) = 9.8$, $p < .05$, but there was no significant difference between the stimulus classes, $F(1,8) < 1$, $p = .69$, and no significant interaction, $F(1,8) < 1$, $p = .46$.

In accordance with the orientation dependence of biological motion recognition, we did not obtain any significant improvement of discrimination performance for both stimulus types in the test blocks, when training stimuli were rotated against the test stimuli ([Figure 5](#)). This was confirmed by a repeated measures ANOVA showing no significant influence of the number of the test block, $F(2,16) < 1$, $p = .44$. In accordance with the results of

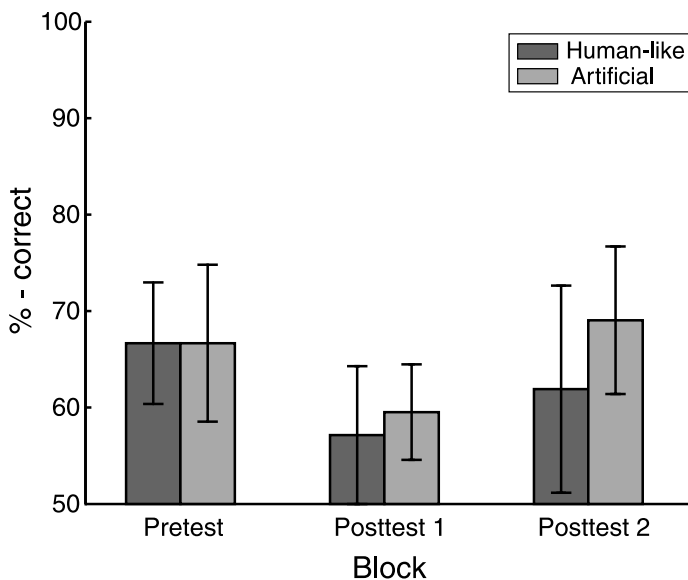


Figure 5. Orientation dependence of the learned representation. Mean percentages of correct responses (\pm SEM) for three test blocks. Training stimuli were rotated against the test stimuli in the image plane (by 90 deg). Recognition rates are shown for the human-like movements (dark gray) and for the artificial articulated movements (light gray; $n = 9$).

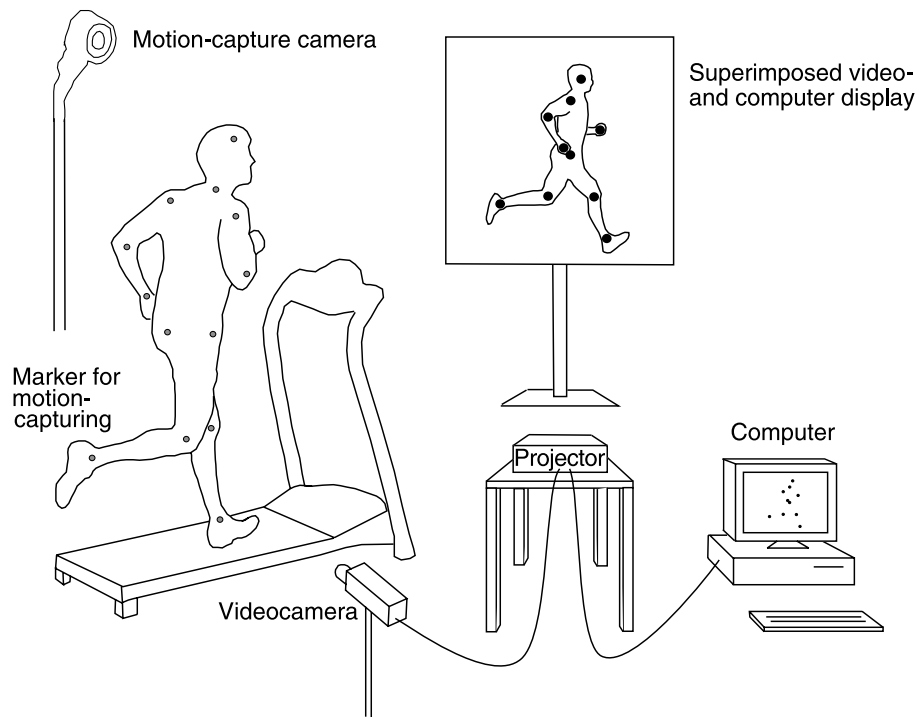


Figure 6. Setup for approximating human-like motion morphs by real human movements. A point-light walker, animated with the trajectories generated by motion morphing, is superposed to an online video of the actor walking on a treadmill. The movements of the actor, who tries to imitate the movements of the point-light walker, are motion-captured (using reflecting markers on shoulders, elbows, wrists, knees and ankles, head and hip). The recorded trajectories closely approximate the trajectories of the human-like motion morph but fulfill exactly the laws of motion of real human movements. Based on the processed trajectories, point-light stimuli with 10 dots were generated by averaging the positions of the head and the shoulder markers as well as the hip markers. A movie showing a real human sample stimulus can be obtained as Auxiliary [Movie 3](#).

Experiment 1, there was no significant difference between the two stimulus classes, $F(1,8) < 1$, $p = .45$, and no significant interaction, $F(2,16) < 1$, $p = .79$.

This result implies that, similar to normal biological motion patterns, the representations for the novel learned patterns are strongly orientation dependent. The fact that observers' performance substantially improved during the training blocks confirms that they were able to learn rotated novel patterns but showed a lack of transfer to the normal orientation. This result seems consistent with the hypothesis that biological motion recognition is based on the learning of orientation-dependent or, potentially, view-dependent templates (Giese & Poggio, 2003; Verfaillie, 2000). This implies that biological motion recognition might be based on similar principles as the recognition of complex static shapes (see Tarr & Bülthoff, 1998, for review).

Experiment 3: Comparison with real human movements

All stimuli in [Experiments 1](#) and [2](#) were generated by motion morphing, that is, by interpolation between prototypical trajectories. Theoretically, even when

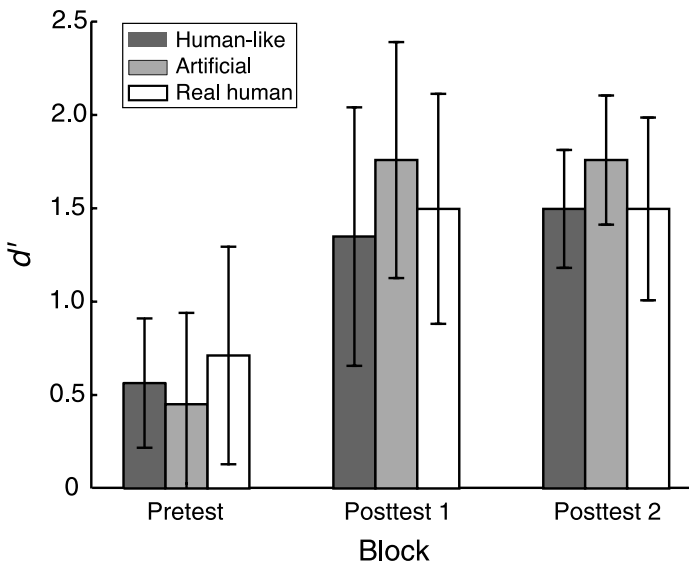


Figure 7. Real human movements compared with artificial movements. Shown are the d' values for the human-like stimuli (dark gray), the artificial articulated stimuli (light gray), and real human movements approximating human-like morphs (white) for the three test blocks ($\pm SEM$; $n = 7$).

generated from human trajectories, such interpolations might violate specific properties of human movements that may be critical for their perception. This fact might account for the lack of a difference between human-like and artificial motion morphs. To rule out this possible confound, we conducted an additional control experiment comparing artificial morphs with real human movements, which exactly fulfill the human “laws of motion.”

Our control experiment exploits the fact that the human-like motion morphs are very similar to human movements, in that they can be executed by a human actor. To create a stimulus set that was very similar to Experiment 1, we developed a method that approximates human-like motion morphs by the movements of a real human actor, exploiting the setup illustrated in Figure 6.

A human actor on a treadmill tried to imitate human-like motion morphs. The actor viewed a superposition of an online video of his own movement and a point-light stimulus, whose movements were defined by the motion morph. The actor tried to align the joints of his body, as accurately as possible, with the positions of the dots of the point-light walker, monitoring himself on the video screen. After several minutes of training, the actor was able to accomplish a relatively accurate reproduction of the movements of the point-light walker (reproducing 88% of the variance of the two-dimensional joint trajectories).

The movements of the actor were recorded using a VICON 612 motion capture system with six cameras. The three-dimensional positions of 22 reflecting markers were recorded with a sampling frequency of 120 Hz and a spatial accuracy below 1 mm. The resulting trajectories were processed using commercial software by VICON. From

each movement, several movement cycles were recorded, and the cycle that was most similar (after projection on the two-dimensional plane) to the imitated point-light stimulus was selected for stimulus generation. Based on the processed marker trajectories, point-light stimuli with 10 dots were generated and presented with the same frame rate as the original stimuli.

Results and discussion

The design of Experiment 3 was identical to that of Experiment 1. However, we compared the learning of three stimulus classes: artificial articulated movements (Movie 2), human-like morphs (Movie 1), and real human movements (Movie 3).

The results of this experiment are summarized in Figure 7. Based on a different set of participants, this experiment replicates the results of Experiment 1. For the analysis of differences between the different stimulus types, we applied a signal detection analysis and computed d' values. In accordance with the improvement in performance in Experiment 1, we found a marginally significant increase of d' values with the number of test blocks, $F(2,12) = 3.8$, $p = .07$, whereas there was no significant influence of the stimulus type, $F(2,12) < 1$, $p = .47$, and no significant interaction, $F(4,24) < 1$, $p = .67$. To verify the marginally significant effect of the training, we conducted a separate analysis also for the percentage of correct responses, similar to what we did in Experiments 1 and 2. For this analysis, we found a significant main effect of the number of the test block, $F(2,12) = 7.1$, $p < .05$, supporting the improvement due to training, but there was no significant influence of the stimulus type, $F(2,12) < 1$, $p = .65$, and no significant interaction, $F(4,24) < 1$, $p = .79$.

These results rule out the possibility that the similarities between the learning of human-like and artificial articulated patterns obtained in Experiment 1 are only caused by the fact that all stimuli were generated by motion morphing. The same similarity is found between real human movements and artificial articulated patterns.

Experiment 4: Learning of stimuli with and without global underlying shape

Experiments 1, 2, and 3 have demonstrated strong similarities for the learning of human-like and artificial articulated movement patterns. This raises the question if all movement patterns of comparable complexity can be learned, even if they do not support the perception of a global underlying shape or skeleton. To test this question, we compared the learning of human-like movements (Movie 1) with the learning of movement patterns without an underlying skeleton but with identical local motion information (Movie 4). In this way, we tested whether an underlying skeleton facilitates discrimination learning and generalization. Such facilitation would be specifically

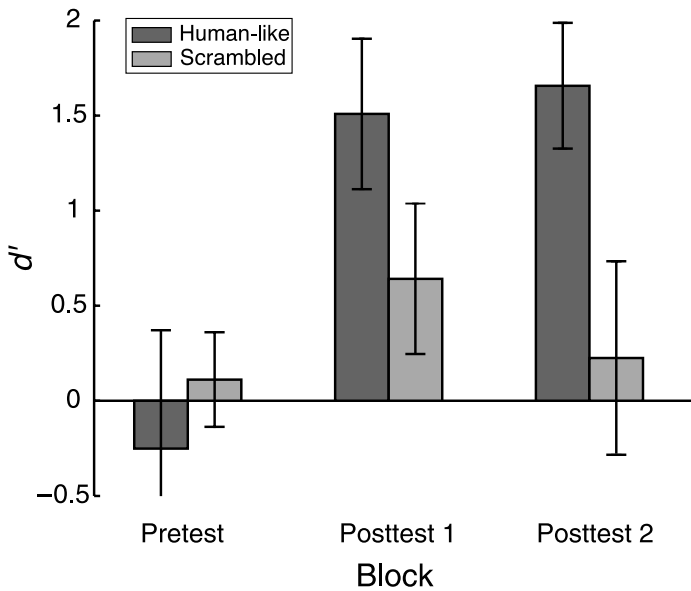


Figure 8. Human-like movement stimuli compared with scrambled human-like stimuli. Shown are the d' values for the intact human-like stimuli (dark gray) and the scrambled stimuli (light gray) in the individual test blocks (\pm SEM; $n = 7$).

expected if the learning of motion patterns is dependent on shape models or on a hierarchical grouping of local movements (e.g., Johansson, 1986; Restle, 1979).

To generate stimuli that are not compatible with an underlying skeleton, we spatially scrambled the human-like movement patterns; that is, we added temporally constant random position offsets to the individual dot positions. This operation destroys the consistency of the movements with an underlying articulated shape. However, because the offsets are temporally constant, it does not affect the local motion information of the stimuli. The offsets were constrained to ensure that the scrambled stimuli covered the same spatial area as the original human-like stimuli. None of the tested participants were able to recognize an articulated shape in any of these scrambled stimuli.

The experimental design of this experiment was identical to that of Experiment 1. For each participant, two human-like morphs were randomly chosen and presented intact, while the remaining two human-like morphs were presented as scrambled stimuli.

Results and discussion

The results of Experiment 4 from the three test blocks are presented in Figure 8. In this case, the learning process seems to be different for the two stimulus classes. Although the initial d' values for both stimulus types are close to 0, only the performance for the natural-looking morphs is increasing significantly (like in the previous experiments). The performance for the scrambled stimuli

in the test blocks did not increase after the training. This observation was confirmed by a repeated measures ANOVA with the factors stimulus type and number of the test block. We obtained a significant main effect of stimulus type, $F(1,6) = 11.1, p < .05$, and of the number of the test block, $F(2,12) = 8.3, p < .01$. The interaction was marginally significant, $F(2,12) = 3.8, p = .08$. A separate analysis of the correct responses in percentage correct supported these results. We found a significant main effect of the number of the test block, $F(2,12) = 8.5, p < .01$, and a significant interaction, $F(2,12) = 7.5, p < .05$, whereas the main effect of stimulus type was not significant, $F(1,6) < 1, p = .89$.

Interestingly, a separate analysis of the training blocks suggests no differences in performance during training between the intact and the scrambled stimuli. A repeated measures ANOVA reveals a significant improvement from the first training block to the second training block, $F(1,6) = 10.9, p < .05$, for both stimulus types, but there was no significant effect of stimulus type, $F(1,6) < 1, p = .41$, and no significant interaction, $F(1,6) < 1, p = .78$. This result suggests that participants learned to discriminate between the scrambled training stimuli. However, they were not able to generalize to test stimuli with smaller distances between Center and Off-Center stimuli in morphing space.

An additional analysis of the response times shows that the observed difference between intact and scrambled movements in the test blocks cannot be attributed to a speed–accuracy trade-off. Comparing the response times between the two stimulus types (in the test blocks), we found no significant differences, $F(1,6) < 1, p = .72$. This clearly indicates that participants were trying to succeed for both stimulus types.

Experiments 1, 2, and 3 demonstrated relatively fast discrimination learning for articulated movements, that is, movements with an underlying skeleton. Experiment 4 suggests that compatibility with a skeleton is critical for this fast learning or at least for the generalization to more difficult discriminations of similar stimuli. Yet, we cannot exclude that stimuli without underlying skeletons are learned with longer training times, for example, exploiting specific local feature combinations.

Discussion

Our study investigated the visual learning of complex movement patterns that shared specific properties with natural human movements. We studied discrimination learning for stimulus pairs that were generated by motion morphing, allowing for a precise control of the spatiotemporal similarities between the stimuli and their low-level properties. Our experiments show that humans can learn to discriminate between novel articulated movement patterns

very quickly, after less than 20 stimulus repetitions. This repetition number is consistent with the learning speed in a recent study on learning of the detection of point-light patterns in noise (Hiris et al., 2005). In addition, the learned representation generalized to similar motion patterns, enabling participants to accomplish more difficult discriminations after the training. Similar to normal biological motion, the learned visual representation seems to be strongly orientation dependent; that is, discrimination performance after training increased only when the test patterns were presented with the same orientation as the training patterns.

Interestingly, learning speed and accuracy for human movements and completely artificial articulated patterns were quite similar. Familiarity or biological relevance of the underlying kinematics or skeleton seems thus not to be critical for the visual learning process. Contrasting with this result, motion stimuli without underlying skeleton could not be learned equally fast, even when their local motion properties were identical to human-like movements. This suggests that the “binding” or grouping of the individual stimulus elements into a global percept might strongly facilitate the learning. The similarity of our stimuli with normal biological motion stimuli and the fact that the learning was orientation dependent make it possible that the investigated learning process is also relevant for normal biological motion recognition.

The observed strong similarity between learning of human-like and artificial articulated movements was highly reproducible. This similarity was observed in Experiments 1, 2, and 3 and in the two additional control experiments (46 participants and 8 different movement stimuli in total). In addition, we found reproducible improvements of performance with training and highly selective differences between trained and untrained stimuli. These observations, and the fact that we obtained a difference between the learning of normal and scrambled human stimuli, rule out the possibility that the observed similarity between the two classes of articulated stimuli just reflects a lack of sensitivity of our paradigm or behavioral measures.

Our study provides some insights that contribute to the question of what is “special” about biological motion perception or at least for the visual learning of such stimuli. Biological movements are characterized by several properties that might be critical for their perception: (1) general smoothness properties of the trajectories, (2) the consistency with an underlying global shape or skeleton, and (3) the familiarity or biological relevance of these shapes and the associated motion patterns. Our experiments show that the last property is not critical for fast learning of complex movements. However, the second property seems to be critical.

By construction, all our stimuli shared the first property, similar smoothness of the trajectories. Smoothness is closely related to the consistency of the movements with laws of motion that are typical for biological motor

behavior (Viviani & Flash, 1995). An example is the “two-thirds power law” (Lacquaniti, Terzuolo, & Viviani, 1983), signifying that curvature and speed of planar human movements are linked by a power law. Psychophysical experiments have shown that simple motion stimuli fulfilling this law appear smoother (Viviani & Stucchi, 1992). The joint trajectories used for the generation of our stimuli were in agreement with this motor invariant. (The exponents of the power law were determined by linear regression, applied to the logarithms of velocities and curvatures, yielding exponents ranging between .31 and .36 for human-like and artificial articulated movements). It remains to be clarified in future experiments whether a violation of general smoothness properties impairs fast visual learning. Well-controlled experiments of this type might be very difficult to realize because it would have to be excluded that differences between stimuli with different smoothness are not just reflecting differences in low-level motion processing, induced by the different motion energy distributions of such stimuli.

There are several possible explanations why learning of human-like and artificial articulated patterns is similar, whereas learning of scrambled patterns is much more difficult. First, there might be a mechanism that recognizes complex movements by matching the underlying articulated shape (e.g., Beintema & Lappe, 2002; Giese, 2000; Giese & Poggio, 2003; Marr & Vaina, 1982; Vaina & Bennour, 1985; Webb & Aggarwal, 1982). This mechanism might operate independently of the biological relevance of such shapes. Second, the recognition of motion patterns might be based on features of intermediate complexity, which only arise for motion that is derived from smoothly deforming or articulated shapes. Third, there could be top-down influences of shape recognition that facilitate the learning of motion patterns. The existence of top-down influences is suggested by a number of psychophysical and imaging studies showing that local motion perception and the activity in motion-related brain areas are modulated by the recognition of shapes, in particular when they are typically associated with body movements (e.g., Chatterjee, Freyd, & Shiffrar, 1996; Kourtzi & Kanwisher, 2000; Peuskens, Vanrie, Verfaillie, & Orban, 2005; Senior et al. 2000).

The importance of top-down mechanisms seems also consistent with our observation that more difficult discriminations between articulated movements were facilitated by previously learned simpler discriminations of similar patterns. The same phenomenon has been observed in other visual learning experiments (e.g., Liu & Weinshall, 2000; Mackintosh, 1974) and has also been termed perceptual “Eureka” (Ahissar & Hochstein, 1997). It has been explained by the learning of a more effective allocation of attention to features or stimulus dimensions that are relevant for the discrimination. Several psychophysical studies suggest the existence of top-down influences in biological motion perception (Cavanagh, Labianca, &

Thornton, 2001; Thornton, Rensink, & Shiffrar, 2002). The results of our experiments suggest that an underlying global shape and, potentially, form recognition might facilitate such top–down processes.

Another implication of our study is that consistency with a preexisting internal (dynamic) body model seems not to be required for fast visual learning of complex movements. Such internal models likely contribute to the perception of imitable body movements (e.g., Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Grafton, Fadiga, Arbib, & Rizzolatti, 1997; Prinz, 1997; Wilson & Knoblich, 2005; Wolpert, Doya, & Kawato, 2003). Recent imaging experiments suggest that neural structures that are involved in the representation of such internal models are also activated by point-light walkers (Saygin, Wilson, Hagler, Bates, & Sereno, 2004). However, we think that it is unlikely that the learning of the artificial articulated patterns was based on such internal body models because their kinematics differed strongly from human bodies and specified nonimitable movements. The contribution of internal models to the recognition of artificial stimuli could potentially be clarified in brain imaging studies that compare activity distributions for the two articulated stimulus types.

Several studies have suggested that infants have an innate preference for the processing of biological motion (Fox & McDaniel, 1982; Grezes et al., 2001; Johansson, von Hofsten, & Jansson, 1980; Meltzoff & Moore, 1977; Pavlova, Staudt, Sokolov, Birbaumer, & Krageloh-Mann, 2003). A preference for biological motion stimuli has also been observed in animals (e.g., Blake, 1993). For example, inexperienced newly hatched chicks demonstrate an innate predisposition to approach motion stimuli that share specific low-level properties with biological movements (Vallortigara, Regolin, & Marconato, 2005). However, this preference seems not to be selective for the global stimulus structure because the animals were equally attracted by scrambled and intact point-light displays of hens. Also, this innate preference seemed not to be selective for movements of different species, for example, own species versus predators. Yet, an innate predisposition to attend to stimuli with low-level properties that are typical for biological movements might be very helpful to support the learning of more subtle biologically important differences between complex motion stimuli. Our results complement these studies about potential unspecific innate factors by providing a detailed investigation of stimulus properties that seem to be critical for the learning of detailed distinctions between complex movement patterns.

Summarizing, we have demonstrated the existence of a fast visual learning process for the holistic structure of complex motion patterns. This process shares important properties with normal biological motion perception and seems not to differentiate between biologically relevant and nonbiological articulated movements, as long as they convey the percept of a global form. Learning might thus

play an important role for understanding the perception of biological motion and actions.

Acknowledgments

We thank I. Bühlhoff, Q. Vuong, P. Thier, and L. M. Vaina for providing helpful comments on the manuscript and interesting discussions. Also, we thank M. Pavlova, J. E. Cutting, and D. R. Proffitt for helpful suggestions. This work was supported by the DFG, SFB550-C10, HFSP, the Volkswagenstiftung, and the Max Planck Society. M. Giese is a visiting research fellow at the Department of Biomedical Engineering, Boston University.

Commercial relationships: none.

Corresponding author: Martin A. Giese.

Email: martin.giese@uni-tuebingen.de.

Address: Laboratory for Action Representation and Learning, Schaffhausenstr. 113, Hertie Institute for Clinical Brain Research, D-72072 Tübingen, Germany.

Bibliography

- Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–406. [[PubMed](#)]
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*, 953–965. [[PubMed](#)]
- Beintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 5661–5663. [[PubMed](#)] [[Article](#)]
- Bertenthal, B. I., Proffitt, D. R., & Kramer, S. J. (1987). Perception of biomechanical motions by infants: Implementation of various processing constraints. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 577–585. [[PubMed](#)]
- Blake, R. (1993). Cats perceive biological motion. *Psychological Science*, *4*, 54–57.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. [[PubMed](#)]
- Bruderlin, A., & Williams, L. (1995). Motion signal processing. *Computer Graphics*, *29*, 97–104.
- Bühlhoff, I., Bühlhoff, H., & Sinha, P. (1998). Top-down influences on stereoscopic depth-perception. *Nature Neuroscience*, *1*, 254–257. [[PubMed](#)] [[Article](#)]
- Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition*, *80*, 47–60. [[PubMed](#)]

- Chatterjee, S. H., Freyd, J. J., & Shiffrar, M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 916–929. [PubMed]
- Edelman, S., & Bühlhoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, *32*, 2385–2400. [PubMed]
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611. [PubMed]
- Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning tasks: A review. *Journal of Vision*, *2*(2), 190–203, <http://journalofvision.org/2/2/5/>, doi:10.1167/2.2.5. [PubMed] [Article]
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, *218*, 486–487. [PubMed]
- Giese, M. A. (2000). Neural field model for the recognition of biological motion. In H. Bothe & R. Rojas (Eds.), *Second International ICSC Symposium on Neural Computation (NC 2000)* (pp. 828–835). ICSC Academic Press.
- Giese, M. A., & Lappe, M. (2002). Measurement of generalization fields for the recognition of biological motion. *Vision Research*, *42*, 1847–1858. [PubMed]
- Giese, M. A., & Poggio, T. (2000). Morphable models for the analysis and synthesis of complex motion pattern. *International Journal of Computer Vision*, *38*, 59–73. [Article]
- Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews. Neuroscience*, *4*, 179–192. [PubMed]
- Giese, M. A., Thornton, I. M., & Edelman, S. (2003). Metric category spaces of biological motion [Abstract]. *Journal of Vision*, *3*(9), 83a, <http://journalofvision.org/3/9/83/>, doi:10.1167/3.9.83.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612. [PubMed]
- Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *NeuroImage*, *6*, 231–236. [PubMed]
- Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? *NeuroImage*, *13*, 775–785. [PubMed]
- Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research*, *41*, 1475–1482. [PubMed]
- Grossman, E. D., Blake, R., & Kim, C. Y. (2004). Learning to see biological motion: Brain activity parallels behavior. *Journal of Cognitive Neuroscience*, *16*, 1669–1679. [PubMed]
- Hill, H., & Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science*, *11*, 223–228. [PubMed]
- Hiris, E., Krebeck, A., Edmonds, J., & Stout, A. (2005). What learning to see arbitrary motion tells us about biological motion perception. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1096–1106. [PubMed]
- Jastorff, J., Kourtzi, Z., & Giese, M. A. (2003). Role of learning in biological motion recognition [Abstract]. *Journal of Vision*, *3*(9), 84a, <http://journalofvision.org/3/9/84/>, doi:10.1167/3.9.84.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211.
- Johansson, G. (1986). Relational invariance and visual space perception: On perceptual vector analysis of the optic flow. *Acta Psychologica*, *63*, 89–101. [PubMed]
- Johansson, G., von Hofsten, C., & Jansson, G. (1980). Event perception. *Annual Review of Psychology*, *31*, 27–63. [PubMed]
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*, 48–55. [PubMed]
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*, 575–580.
- Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica*, *54*, 115–130. [PubMed]
- Lee, J., & Wong, W. (2004). A stochastic model for the detection of coherent motion. *Biological Cybernetics*, *91*, 306–314. [PubMed]
- Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, *40*, 97–109. [PubMed]
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552–563. [PubMed] [Article]
- Mackintosh, N. J. (1974). *The psychology of animal learning*. London: Academic.
- Marr, D., & Vaina, L. (1982). Representation and recognition of the movements of shapes. *Proceedings of the Royal Society of London: Series B, Biological Sciences*, *214*, 501–524. [PubMed]

- Mayer, M. J. (1983). Practice improves adults' sensitivity to diagonals. *Vision Research*, *23*, 547–550. [PubMed]
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 74–78. [PubMed]
- O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, *6*, 261–266. [PubMed]
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, *76*, 109–129. [PubMed]
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*, 889–899. [PubMed]
- Pavlova, M., Staudt, M., Sokolov, A., Birbaumer, N., & Krageloh-Mann, I. (2003). Perception and production of biological movement in patients with early periventricular brain lesions. *Brain*, *126*, 692–701. [PubMed] [Article]
- Perrett, D. I., Smith, P. A., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., et al. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: A preliminary report. *Behavioral Brain Research*, *16*, 153–170. [PubMed]
- Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *The European Journal of Neuroscience*, *21*, 2864–2875. [PubMed]
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science*, *256*, 1018–1021. [PubMed]
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154.
- Restle, F. (1979). Coding theory of the perception of motion configurations. *Psychological Review*, *86*, 1–24. [PubMed]
- Saygin, A. P., Wilson, S. M., Hagler, D. J., Jr., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *The Journal of Neuroscience*, *24*, 6181–6188. [PubMed] [Article]
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., et al. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Current Biology*, *10*, 16–22. [PubMed] [Article]
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, *13*, 283–286. [PubMed]
- Tarr, M. J., & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, *67*, 1–20. [PubMed]
- Thornton, I. M., Rensink, R. A., & Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception*, *31*, 837–853. [PubMed]
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*(5), 371–387, <http://journalofvision.org/2/5/2/>, doi:10.1167/2.5.2. [PubMed] [Article]
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics*, *67*, 667–675. [PubMed]
- Vaina, L., & Bennour, Y. (1985). A computational approach to visual recognition of arm movements. *Perceptual and Motor Skills*, *60*, 203–228. [PubMed]
- Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biology*, *3*, e208. [PubMed] [Article]
- Verfaillie, K. (2000). Perceiving human locomotion: Priming effects in direction discrimination. *Brain and Cognition*, *44*, 192–213. [PubMed]
- Verfaillie, K., De Troy, A., & Van Rensbergen, J. (1994). Transsaccadic integration of biological motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 649–670. [PubMed]
- Viviani, P., & Flash, T. (1995). Minimum-jerk, two-thirds power law, and isochrony: Converging approaches to movement planning. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 32–53. [PubMed]
- Viviani, P., & Stucchi, N. (1992). Biological movements look uniform: Evidence of motor–perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 603–623. [PubMed]
- Webb, J. A., & Aggarwal, S. J. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence*, *19*, 107–130.
- Wiley, D., & Hahn, J. (1997). Interpolation synthesis of articulated figure motion. *IEEE Computer Graphics and Applications*, *17*, 39–45.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*, 460–473. [PubMed]
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, *358*, 593–602. [PubMed] [Article]