Biological motion as a cue for the perception of size

Daniel Jokisch　　　　Institute of Cognitive Neuroscience,
Ruhr-University, Bochum, Germany

Nikolaus F. Troje　　　　Institute of Cognitive Neuroscience,
Ruhr-University, Bochum, Germany

Animals as well as humans adjust their gait patterns in order to minimize energy required for their locomotion. A particularly important factor is the constant force of earth's gravity. In many dynamic systems, gravity defines a relation between temporal and spatial parameters. The stride frequency of an animal that moves efficiently in terms of energy consumption depends on its size. In two psychophysical experiments, we investigated whether human observers can employ this relation in order to retrieve size information from point-light displays of dogs moving with varying stride frequencies across the screen. In Experiment 1, observers had to adjust the apparent size of a walking point-light dog by placing it at different depths in a three-dimensional depiction of a complex landscape. In Experiment 2, the size of the dog could be adjusted directly. Results show that displays with high stride frequencies are perceived to be smaller than displays with low stride frequencies and that this correlation perfectly reflects the predicted inverse quadratic relation between stride frequency and size. We conclude that biological motion can serve as a cue to retrieve the size of an animal and, therefore, to scale the visual environment.

Keywords: biological motion, gravity, size perception, gait

Introduction

The perception of motion is a fundamental property of the visual system. One of the most complex but also most familiar types of motion are the nonrigid movement patterns of living organisms. For animals as well as for humans, animate motion patterns contain a wide variety of information. Correct interpretation of this information is an important ability. In the animal kingdom, accurate and fast movement recognition of a prey or predator animal increases an animal's fitness and, therefore, its chance of survival. For humans, the ability to identify, interpret, and predict the actions of others is of particular relevance in the context of successful social interaction that plays a major adaptive role.

Visualizing the position of the main joints of a walking person by bright dots is enough to convey a vivid impression of a human figure in motion. The percept collapses into a meaningless array of unconnected dots when the walker stands still, demonstrating that the interpretation is carried solely by the dynamics of the display (Johansson, 1973). Observers require only 100–200 ms to organize such point-light displays into a coherent percept (Johansson, 1976). The rudimentary information contained in point-light displays of biological motion is sufficient even to solve sophisticated recognition tasks. Observers are able to recognize the gender of a walking person (Barclay, Cutting, & Kozlowski, 1978; Cutting, 1978; Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Troje, 2002), can identify friends by their gait (Cutting &

Kozlowski, 1977), and can even recognize themselves from a recorded point-light display of their own movements (Beardsworth & Buckner, 1981). Mather and West (1993) extended the point-light display paradigm to animations of four-legged animals and showed that human observers can identify different animals by their movement pattern. Inversion effects of biological motion displays of animal movements were investigated by Pinto and Shiffrar (1999). The ability to perceive biological motion is not restricted to humans. It has been shown that cats are able to identify point-light displays of conspecifics (Blake, 1993), that pigeons are capable of discriminating between categories of conspecifics' walking and pecking when presented as point-light displays (Dittrich, Lea, Barrett & Gurr, 1998), and that chicks and quails also have the ability to perceive point-light displays of biological motion of conspecifics (Yamaguchi & Fujita, 1999). The ability of nonhuman primates to perceive biological motion was indicated by the finding of single cells responding selectively to biological motion displays (Oram & Perrett, 1994).

Animals as well as humans adjust their gait patterns in order to minimize the energy required for their locomotion. The energy costs are determined by the properties of the physical world. A particularly important factor in this context is the constant force of earth's gravity. For many dynamic events occurring under constant gravity conditions, a fixed relation between temporal and spatial parameters is maintained. This relation is particularly valid for inanimate motion systems, such as pendulum motion or ballistic motion.

However, it also seems to hold for many animate motion patterns. Therefore, from a theoretical point of view, time can be used as an information source about spatial scale
in visually recognizable events under the influence of gravity. Several studies have investigated the perception of scale properties in inanimate dynamic events.

Pittenger (1985, 1990) examined the perception of the scale properties in pendulum motion. The length of a freely swinging pendulum is proportional to the square of its period. Pittenger (1985) found that observers could estimate the length of a pendulum when given information about its period. The estimated lengths were found to be a linear function of actual lengths, though with wide differences in slopes among individual observers. When viewing normal pendulums with physically correct periods and perturbed pendulums with either shorter or longer periods, observers could rate the naturalness of motion with a high degree of acuity (Pittenger, 1990).

The same idea has also been applied to the perception of the distance of objects in free fall (Saxberg, 1987a, 1987b; Watson, Banks, von Hofsten, & Royden, 1992). The law of free fall motion relates the height of a fall to the duration of the event. Analogous to pendulum motion, the height of fall is proportional to the square of its duration. In a simulated catching task, in which observers should predict the position where a ball approaching along a parabolic trajectory would fall, Saxberg (1987b) tested whether observers make use of this information. When the display contained information both from image expansion and vertical component of free fall, observers performed this task well, but when information of image expansion was eliminated, they failed. The authors concluded that the latter finding demonstrated a lack of using the information mediated by the relation between height of fall and its duration. However, Watson et al. (1992) argued that this failing was based on conflicting sources of information and not purely on the inability to retrieve the relation between height and duration of the event.

Stappers and Waller (1993) tested people's ability to use the time of free fall of objects as a reference to spatial scale and showed that observers reliably matched gravitational acceleration to apparent depth in a computer simulation.

Hecht, Kaiser, and Banks (1996) examined whether observers could utilize size and distance information provided by gravitational acceleration by presenting observers with displays of the motion of rising and falling objects. Observers were able to use the information to some extent but were more sensitive to average velocity than to gravitational acceleration.

Another study that investigated the perception of spatio-temporal patterns of object motion (Warren, Kim, & Husney, 1987) demonstrated observers' ability to make accurate perceptual judgments of elasticity of bouncing objects by detecting single period duration visually or auditorily in absence of height information.

McConnell, Muchisky, and Bingham (1998) tested observers' ability to judge object size in event displays that eliminated all information other than time and trajectory forms. Initially, judgment variability was substantial, but after feedback on one event, observers performed better and generalized training to other events. Observers were sensitive to the general form of the spatio-temporal scaling relation, but required feedback to attune event-specific constants.

The general form of the relation between a spatial scale s and a temporal scale T in events governed by gravity is given by

$$s = kT^2 \text{ ,} \tag{1}$$

where k is a constant factor specific to the event being considered.

The above findings document that the human visual system seems to be able to use this quadratic relation in order to achieve size information from temporal cues. The absolute quantitative relation expressed in the constant k, however, is not as easily obtainable.

Psychophysical studies considering the relationship between temporal and spatial parameters as visual cues for event perception have not been restricted to inanimate dynamic systems. In the domain of animate motion, such visual cues are proposed to play a role in action perception. Runeson and Frykholm (1981, 1983) have shown that the weight of an object can be readily estimated by observing another person lifting and carrying it when the person is represented as point-light display. They concluded that the crucial information is embedded in the kinematics of the action pattern, in which an object's weight is specified by the magnitude of postural adjustments relative to the acceleration of the object. Bingham (1987, 1993a) provided further empirical evidence for the content of information about an object's weight in the kinematic pattern.

The studies by Runeson and Frykholm (1981, 1983) and Bingham (1987, 1993a) investigate the ability to derive additional information from visual point-light displays of human actions employing knowledge about the effects of gravity on objects in the physical world. Therefore, these studies are related to our question. However, they do not directly address the question whether temporal parameters from biological motion can be used as a cue about size information of animate beings.

From a physical point of view, the relation between temporal and spatial parameters described above is also evident in animate locomotion patterns. A simple model for a walking biped is an inverted pendulum that idealizes the total body mass to a point mass on a rigid mass-less leg (Alexander, 1977). More complex models consider humans and animals as a set of coupled,

articulated pendulum segments. No mechanical energy is needed to maintain the movements of an ideal undamped pendulum because kinetic and gravitational potential energy fluctuations are equal in amplitude and exactly 180° out of phase. In humans, the pendulum-like mechanism conserves about 65% of the mechanical energy from step to step at the preferred walking speed (Cavagna, Thys, & Zamboni, 1976). Pendulum-like energy exchange diminishes at faster walking speeds because of a mismatch in the magnitudes and phases of the fluctuations of the two forms of mechanical energy. Thus, at non-optimal speeds, the muscles must provide additional mechanical power. The relation between the length l and the period T of an ideal pendulum is

$$l = gT^2 / 4\pi^2 , \tag{2}$$

with g being gravitational acceleration. In order to obey this relation, smaller animals have to move with a higher stride frequency $f = 1/T$ than larger animals.

The major force that determines the pendulum-like movements during walking is gravity, which must be at least equal to the centripetal force needed to keep the center of mass traveling along a circular arc. The centripetal force needed is equal to mv^2/L, where m is body mass, L is leg length, and v is forward speed (Kram, Domingo, & Ferris, 1997). The ratio between the centripetal force and the gravitational force $(mv^2/L)/mg = v^2/gL$ is the dimensionless Froude number (Alexander, 1989). Therefore, if animals travel with equal Froude number, their speeds v are proportional to the square root of the leg length L. If they move in dynamically similar fashion (Alexander & Jayes, 1983), the stride length l is proportional to the leg length and hence the stride frequency $f = v/l$ is inversely proportional to the square root of the leg length. Pennycuick (1975) measured the stride frequencies of African mammals moving spontaneously in their natural habitat and found that they are in fact inversely proportional to the square root of the stride length to a very good approximation. Thus, the findings show that the relation between spatial and temporal scales expressed in Equation 1 is also reflected in the locomotion patterns of animals.

In this study, we examined whether the human visual system is able to use this relation to derive the size of an animal in the absence of other cues. To achieve this, we presented observers with point-light displays of a dog. Varying the playback speed, we asked observers to estimate the size of the dog. We predicted that animals are perceived to be larger in animations presented with low stride frequency and smaller in animations with high stride frequency. More specifically, we assumed that the relationship between the stride frequency f of an animal and its estimated size s_{dyn} is

$$s_{dyn} = \frac{c_1}{f^2} , \tag{3}$$

where c_1 is a constant factor quantifying the spatio-temporal scaling relation. The absolute value of c_1 depends on gravitational acceleration and on the gait pattern (e.g., trotting, cantering, etc.).

However, the kinematics of the animation may not be the only source of information about the dog's size. Additional size cues might be contained in an animal's posture or proportions of body segments. For example, Pittenger and Todd (1983) have shown that changes of static body proportions of line drawings of a human body have an effect on perception of growth, and, therefore, also have an indirect effect on the perception of size. Studies using other biological objects have also shown that the perception of size can be influenced by form information. Bingham (1993b, 1993c) showed that properties of tree form could be used to estimate the height of trees.

The size information embedded in body proportions is independent of the temporal scaling factor and can be described as a second constant:

$$s_{stat} = c_2 . \tag{4}$$

s_{dyn} and s_{stat} exist simultaneously and both may contribute to a size estimate. Here, we assume linear integration, and we introduce a factor λ accounting for the relative weight of the two terms:

$$s_{all} = \lambda c_1 \frac{1}{f^2} + (1 - \lambda)c_2 . \tag{5}$$

In order to test this hypothesized model, we conducted two experiments presenting observers with point-light displays depicting a dog moving across the screen. We chose a dog as a model because dogs cover a wide range of different sizes ensuring that size estimations made by observers are not restricted too much by the range of possible sizes. Our point-light dog was shown as walking through a three-dimensional scene depicting a desert landscape.

When observing the image of such a scene, the perceived size of different objects within the scene depends, on the one hand, on the visual angle covered by the objects and, on the other hand, on the perceived position in depth within the scenery. As a consequence of this size-distance ambiguity, there exist two methods to change the size of an object within the scene: (1) varying its position in depth while maintaining a fixed visual angle or (2) showing the object at a fixed distance and varying the size of the object's visual angle. For both methods, the size of other objects embedded within the scene provides an absolute reference.

In Experiment 1, observers were asked to adjust the apparent size of the dog by changing its position in depth while maintaining its projected size on the screen, and,

therefore, its subtended visual angle. In Experiment 2, observers were allowed to change the size of the dog directly. In Experiment 2, we also added a second task: In addition to estimating the size of the dynamic point-light displays, observers were required to estimate the size of a static stick-figure display.

Experiment 1

The observers' task was to estimate the size of the dog animations. The point-light displays were presented in a desert landscape with varying stride frequencies. Perspective and texture gradient created a three-dimensional percept. Reference objects (cactuses and posts) were scattered across the scene to provide size references at different depths. With the visual angle subtended by the dog remaining constant, observers could place the animation at different locations in depth in order to indicate the perceived size.

Method

Participants

Sixteen students (11 females and 5 males) between the ages of 20 and 39 years from the psychology and biology departments at the Ruhr-University participated in this experiment. They received course credit for their participation. All participants had normal or corrected-to-normal vision. They were naive as to the objectives of this experiment.

Stimuli

Synthetic motion data of a dog ("Animania Dog" by Credo Interactive Inc.) were presented in saggital view as point-light displays. The display consisted of 20 dots altogether. Three dots represented the position of each leg's main joints (forelegs: elbow, carpal, and phalange; hind legs: knee, tarsal, and phalange). The positions of the pelvis and the scapula were both represented by two dots each. Two dots represented the position of the head and two represented the position of the thoracic and coccygeal vertebrae. Each dot had a size of 4 mm^2 and was displayed in a bright green coloring. An additional set of 20 black dots represented the shadows of the dots depicting the dog's body. Adding a shadow ensures that observers perceive the animal's legs to have contact to the ground. The point-light display had a size of 4 cm on the screen corresponding to 4 deg of visual angle at the viewing distance of 58 cm. This distance was fixed by using a wooden chinrest. The image sizes of the point-light displays were held constant across all trials.

In order to determine exactly the gait pattern of our animated dog, we examined the phase relations between the feet. The difference between various gait patterns is described by the phase relations between the movements of the four legs. For instance, the trot is a symmetrical

gait in which diagonal pairs of legs move together. In cantering animals, this symmetry is broken. Whereas one diagonal pair of legs moves in synchrony, the other pair is out of phase, with the respective foreleg being ahead of the contralateral hind leg. According to Alexander (1984), the phase difference of this asynchronous pair is 140 deg. In our data, the phases of the legs with respect to the left foreleg were 155, 205, and 0 deg for the right foreleg, the left hind leg, and the right hind leg, respectively. This pattern clearly shows the asynchronous characteristic of the canter, but the phase difference between foreleg and hind leg of the asynchronous leg pair is smaller than described by Alexander (1984). We still term the gait pattern of our animated dog in the following experiments as "canter," accepting some mismatch between the phase relation in our data and data reported in the literature.
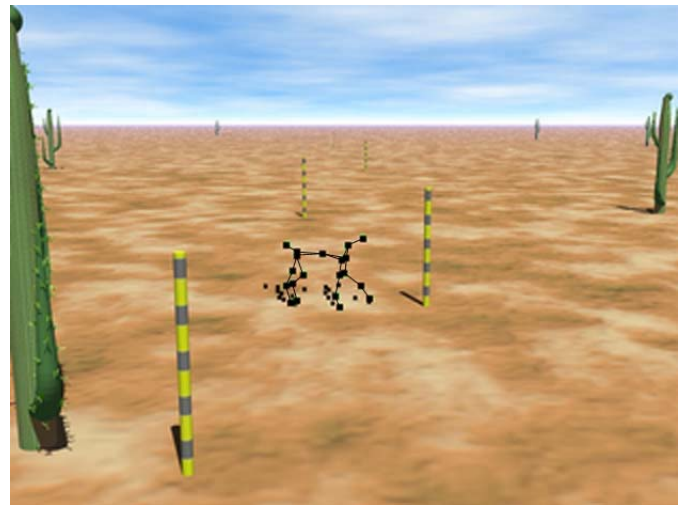

Figure 1. Display of a dog on the perspective background. The lines connecting the dots were shown only in the stick-figure depictions of the second subtask of Experiment 2. They were omitted in Experiment 1 and in the first subtask of Experiment 2. Clicking into the image will evoke an interactive animation similar to the ones shown in the experiment.

The point-light displays were presented on a background depicting a perspective landscape (Figure 1). The landscape was designed with the software Bryce 4 by Meta Creations. It portrayed a desert scene in which were embedded some objects (cactuses and posts) serving as reference objects. All objects belonging to the same class had the same size within the perspective scene (posts 1 m; cactuses 2 m), resulting in varying image sizes on the screen according to their positions in spatial depth. Posts were positioned in regular distances on two parallel lines. Cactuses were arranged in random order. The lens of the camera recording this scenery was positioned 1.5 m above the ground having a tilt angle of 8°. The scenery subtended a visual angle of 35.5 * 24.5 deg.

Procedure

Animated dogs moved across the scene from the left-hand side to the right-hand side. The playback speed was varied systematically, resulting in five different stride frequencies (2.54, 3.02, 3.59, 4.27, and 5.08 cycles/s). These frequencies corresponded to 71, 84, 100, 119, and 141% of the original stride frequency.

By pressing the arrow buttons on the keyboard, participants could change the vertical position of the point-light display on the screen and hence the perceived position in depth in 21 steps. The physical size of the point-light display remained constant. Due to the perspective background, each vertical screen position corresponded to one position in spatial depth, resulting in a changed size impression. Apparent size changed from one position to the adjacent one by factor 1.09. According to the 21 different positions, apparent size changed altogether by a factor of 5.66 within the whole range.

The experiment took place in a separate experimental room. Animations were presented on a 19-inch monitor (90 Hz) at a frame rate of 45 Hz. Observers were told that they would be shown with dogs of different sizes animated as point-light displays. They were instructed to adjust the apparent size of the dogs so that the display on the screen looked as natural as possible.

In each trial, observers were allowed to try different positions as often as they wanted. Each time the observers hit a key to change the size, the dog started at the initial position on the left side of the screen (click Figure 1 to evoke an animation illustrating the stimulus). A trial was completed when the observers had selected one position and confirmed their choice by pressing the space bar. Time for solving the task was unlimited. No feedback was given following the size judgments. Before starting the experimental trials, observers were shown six demonstration trials in order to familiarize them with the displays and the setup. During those demonstration trials, the experimenter pointed out the perspective properties of the scene and drew attention to the various sizes of the objects (posts and cactuses) serving as reference scale.

The experiment was conducted using a one factorial repeated measures within-subjects design. The independent variable encoded the five different stride frequencies of the dog animation. In each condition, 11 repeated trials were presented. Each trial started with different initial sizes covering the whole range of possible sizes. The order of the 55 trials was randomized individually for each participant.

Results and Discussion

The effect of stride frequency on perceived size was significant as tested by an analysis of variance (ANOVA) $(F(4,60) = 11.85, p < .001)$. On average across all participants, animated dogs moving with high stride

frequency were perceived to be smaller than dogs moving with low stride frequency (Figure 2). This outcome confirms the hypothesis that observers retrieve size information from the stride frequency that animals use for locomotion. Recall that the instructions did not explicitly draw observers' attention to the stride frequency of the animated animals. According to the instructions, observers were requested to adjust the position so that the scene looked as natural as possible. Therefore, observers seem to use implicit knowledge to make their size judgments.

Based on the assumptions formulated in Equation 5, the function

$$s_{all} = k_1 \frac{1}{f^2} + k_2 \tag{6}$$

was fitted to the data. Using averages across observers, the best fitting values are $k_1 = 141$ and $k_2 = 35$. With these values, the Equation 6 correlates with $r^2 = 0.96$ to the means of estimated sizes across all observers. Only 4% of the variance of the data remains unexplained. A linear fit, on the other hand, correlates to the empirical data with $r^2 = 0.88$, therefore leaving 12% of the variance unexplained.
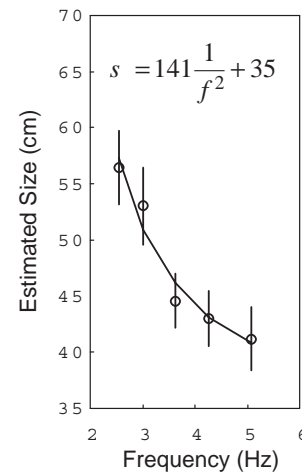

Figure 2. Means across all 16 observers in Experiment 1. The estimated size is plotted for each stride frequency. Error bars indicate SEM. The graph corresponds to the fit of the theoretical model. The coefficient of determination between the function and the means across all observers is $r^2 = .96$.

When focusing on the patterns of results obtained from each observer, clear interindividual differences in consideration of the spatio-temporal scaling relation become obvious, as is indicated by the variability of k_1 (Figure 3).

Substantial individual differences are also evident in the correlation between the empirical data and the model fit (Table 1). Out of 16 observers estimating the size of the animated dogs, 9 showed a response pattern correlating significantly to the model fit. The response

pattern of the others failed to reach a level of significant correlation.

One of the observers (T.B.) reaching a significant level of correlation between his response pattern and the model fit interpreted the temporal scaling factor in

opposition to the expected direction. This observer associated high stride frequencies with large sizes and low stride frequencies with small sizes, resulting in a negative value for k_1.
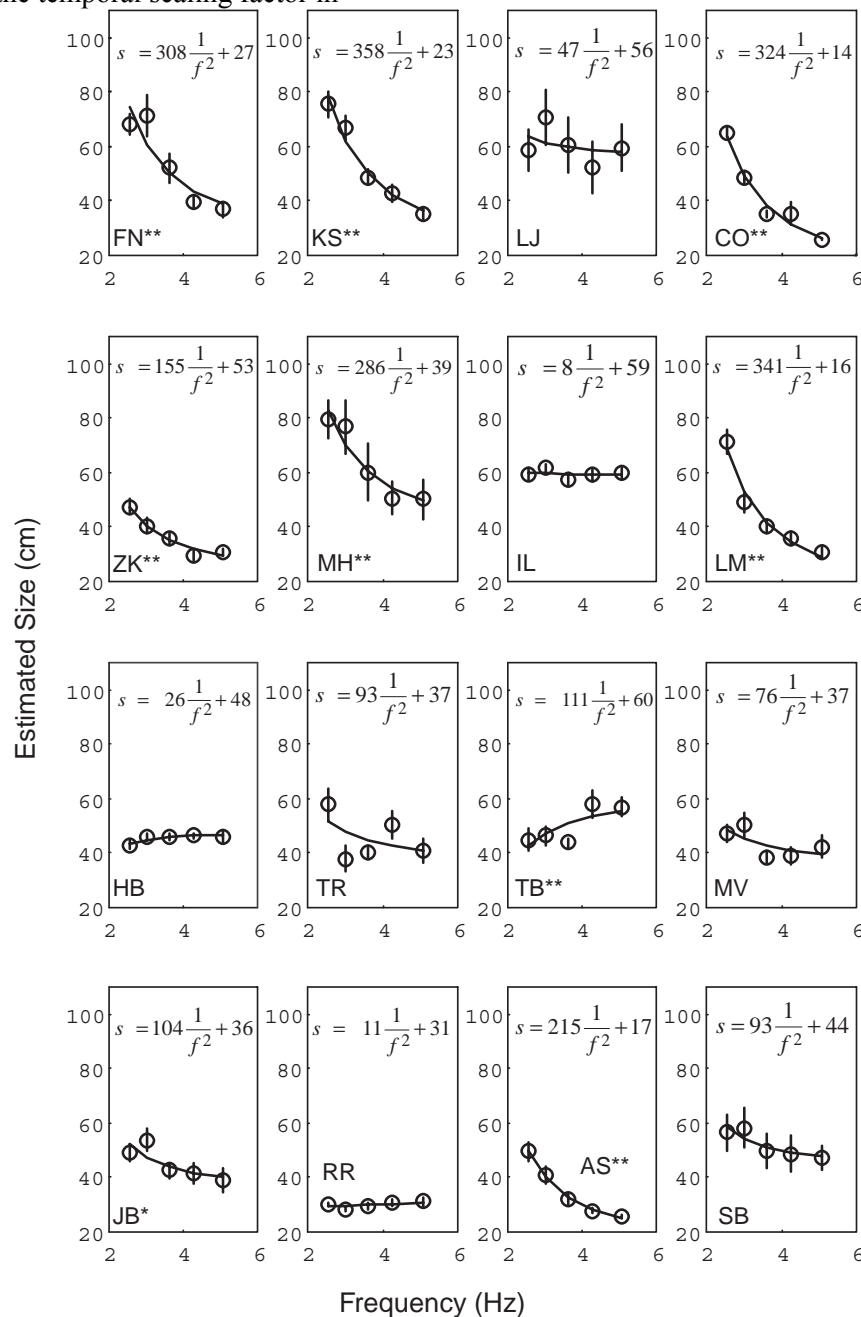

Figure 3. Mean estimated size for each observer for each stride frequency in Experiment 1 (n = 11). Error bars indicate SEM. The graph corresponds to the fitted model to each observer individually. *p < .05; **p < .01 indicates the level of significance of the correlation between the model and individual size estimations.

Consistent size information could be retrieved by 50% of the observers in the setting realized in Experiment 1. This outcome indicates substantial interindividual differences in the ability to retrieve information from the spatio-temporal scaling relation. Such an outcome might have at least two possible

sources. One explanation is that some observers neglect the spatio-temporal scaling relation in their estimations and refer only on other size cues. Alternatively, it may be possible that some observers did not understand the relationship between changes in vertical position and spatial depth, and, therefore, had major problems to

indicate their size impression adequately within this experimental setup.

We inferred perceived size by requiring observers to adjust the position of the dog animation in the landscape. One objection to this task could be that the phenomenon of visual depth compression might cause perceptual distortions of the otherwise well-defined relation between the distance of an object and its projected size. However, as Sedgewick (1993) points out, this would not affect frontal plane dimensions of a projected object.

Table 1. Model Parameters From Experiment 1

Participant	k_1	k_2	r^2
F.N.	308	27	.38**
K.S.	358	23	.62**
L.J.	47	56	.00
C.O.	324	14	.75**
Z.K.	155	23	.45**
M.H.	286	39	.17**
I.L.	8	59	.00
L.M.	341	16	.66**
H.B.	-26	48	.05
T.R.	93	37	.05
T.B.	-111	60	.12**
M.V.	76	37	.08
J.B.	104	36	.11*
R.R.	-11	31	.01
A.S.	215	17	.61**
S.B.	93	44	.04

Parameters of the theoretical model (Equation 6) fitted to the data of individual participants. r^2 = coefficient of determination. *p < .05; **p < .01.

In addition, the scene provides reference objects at different depths. The observers therefore did not have to rely on distance provided by depth cues alone. The size of the dog could be indicated simply in relation to the size of the cactuses and posts scattered around the scene.

Moreover, from the setting realized in Experiment 1, neither the weight factor λ providing information about the individual weights of both sources of information (static vs. dynamic) nor the constants c_1 and c_2 can be calculated directly, because λ is confounded with the constant scaling factors c_1 and c_2 (Equation 5). The constant k_1 combining λ and c_1 only weakly reflects the tendency to what extent the temporal scaling relation is considered.

As a consequence of the above discussed issues, we designed a second experiment. In this experiment, observers were allowed to directly change the size of the dog. While this may facilitate indication of perceived size for the observers, it also rules out any remaining concerns about depth-compression effects. Furthermore, a second subtask was added to deal with the problem of confoundation of the weight factor with the scaling factors.

Experiment 2

In this experiment, we changed the mechanism for indicating perceived size. Observers could change perceived size of the dog directly by changing its projected size while its position in spatial depth remained constant.

In the supplementary task, with the goal to get a direct size estimate based on cues independent of the stride frequency, observers were requested to estimate apparent size of a static stick-figure depiction to derive a direct measure of c_2 in Equation 5. In combination with measurements k_1 and k_1 obtained from the first part of Experiment 2, this was used to derive values for λ and c_1. By this procedure, we are able to separate size information from static and dynamic sources and to calculate how the sources of information are integrated.

Method

Participants

Sixteen students (8 females and 8 males) between the ages of 19 and 32 years from the psychology department of the Ruhr-University participated in this experiment. None of these participants had participated in Experiment 1. Participants received course credit for their participation. All participants had normal or corrected-to-normal vision. They were naive as to the objectives of this experiment.

Stimuli

Stimuli were identical with the ones used in Experiment 1 with the exception that rather than displaying the dog with constant projected size at 21 different positions in depth, this time we generated 21 differently sized dogs and displayed all of them at the same position. The range of apparent sizes covered by this mechanism was the same as in the previous experiment. The visual angle of the dog animation varied from 2.2 deg for the smallest animation to 12.4 deg for the largest animation. The pixel size of the dots describing the positions of the main joints and their shadows on the ground were adjusted accordingly. As in Experiment 1, five different stride frequencies were used: 2.54, 3.02, 3.59, 4.27, and 5.08 cycles/s.

For the second part of the experiment, we generated a static stick-figure depiction of the point-light display on the perspective background used before. The stick figure was positioned in the middle of the screen. Dots belonging to adjacent joints were connected, illustrating the articulation of the joints (Figure 1).

Procedure

The procedure in the first subtask in Experiment 2 was performed similarly to the one used in Experiment 1. The only difference was the mechanism for indicating size. Observers' instructions were similar to the ones in

the former experiment, but were adapted to the new procedure. Six demonstration trials preceded the 55 experimental trials, in which observers gave their size estimates by choosing the dog with the size that looked most natural. Observers were given no feedback following their size judgments. The experiment was conducted using a one factorial repeated measures within-subjects design. In each of the five different frequency conditions, 11 repeated trials were presented. Each trial started with different initial sizes covering the whole range of possible sizes. The order of the 55 trials was randomized individually for each participant.

Having completed the first part of the experiment, participants were instructed about the second subtask, in which they were presented with 11 trials showing static stick-figure displays of a dog. Observers were explicitly told that all stick-figure displays were based on the same animal, varying only on its initial display size and the state (i.e., the phase) of the stride cycle. Using the arrow keys on the computer keyboard, their task was to indicate the size of the stick-figure dogs by the same mechanism as in the first subtask.

Results and Discussion

The results of the first part of this experiment were analyzed as in Experiment 1. Similar to the previous experiment, on average across all observers, dogs moving with high stride frequency were estimated to be smaller than dogs moving with low stride frequency (Figure 4). This effect was significant as tested by an ANOVA ($F(4,60) = 20.67$, $p < .001$).
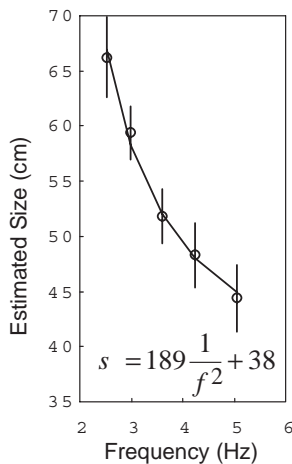

Figure 4. Means across all 16 observers in Experiment 2. The estimated size is plotted for each stride frequency. Error bars indicate SEM. The graph corresponds to the fit of the theoretical model. The coefficient of determination between the function and the means across all observers is $r^2 = 0.98$.

This finding again supports the spatio-temporal scale hypothesis. The following function provides the best fit between the theoretical model and the empirical data:

$$s = 189\frac{1}{f^2} + 38 \ .$$

The coefficient of determination between this function and the means of estimated sizes across all observers was $r^2 = 0.98$. A linear fit correlates to the model with $r^2 = 0.94$. Comparing the proposed model fit with a linear fit, the proposed model leaves only 2% of the variance unexplained, whereas the linear fit leaves 6% of the variance unexplained.

The median of the static figure size estimations of each observer in the second subtask was taken as value for c_2, representing size information independent of any temporal scaling cue. On average across all observers, c_2 assumes a value of 61.47 cm. The standard deviation of 13.90 cm is relatively small, indicating a generally uniform behavior in this subtask. Individual measures for c_2 were used to determine the weight factor $\lambda = 1 - k_2 / c_2$ and the spatio-temporal scaling factor $c_1 = k_1 * c_2 /(c_2 - k_2)$ for each observer, according to Equation 5 (Table 2).

Table 2. Model Parameters from Experiment 2

Participant	k_1	k_2	c_1	c_2	λ	r^2
A.C.	491	20	732.84	60.00	.67	.71**
J.A.	186	30	413.33	55.00	.45	.25**
H.O.	219	38	521.43	65.43	.42	.29**
U.A.	204	34	340.02	84.82	.60	.24**
A.A.	-13	69	86.67	60.00	-.15	.00
S.I.	68	53	566.67	60.00	.12	.02
J.N.	143	49	572.01	65.43	.25	.08
N.K.	427	12	514.46	71.34	.83	.81**
C.N.	184	33	408.89	60.00	.45	.52**
D.M.	-5	70	-20.83	92.50	.24	.00
P.P.	97	33	440.91	42.41	.22	.18**
M.H.	-27	43	128.57	35.67	-.21	.06
A.G.	205	34	427.03	65.43	.48	.26**
C.K.	180	29	382.98	55.00	.47	.36**
M.K.	283	21	435.38	60.00	.65	.45**
J.C.	373	33	1065.71	50.43	.35	.18**

Characteristics of the theoretical model (Equation 5) fitted to the data of individual participants. Note: $k_1 = \lambda\, c_1$; $k_2 = (1-\lambda)\, c_2$. c_2 was derived from the median of the size estimations per observer given in the static stick-figure trials. r^2 = coefficient of determination. *$p < .05$; **$p < .01$.

The individual response patterns again showed considerable inter-individual differences in the use of the spatio-temporal scaling factor (Figure 5). In this experiment, a very clear division into two groups became apparent.

Whereas for 11 out of 16 observers the correlation with the proposed model was highly significant ($p < .01$), there was no correlation at all for the remaining 5 observers ($p > .05$). Showing very flat curves, these observers did not seem to pay any attention to the different stride frequencies. Their response patterns seemed to be completely ignorant with respect to the
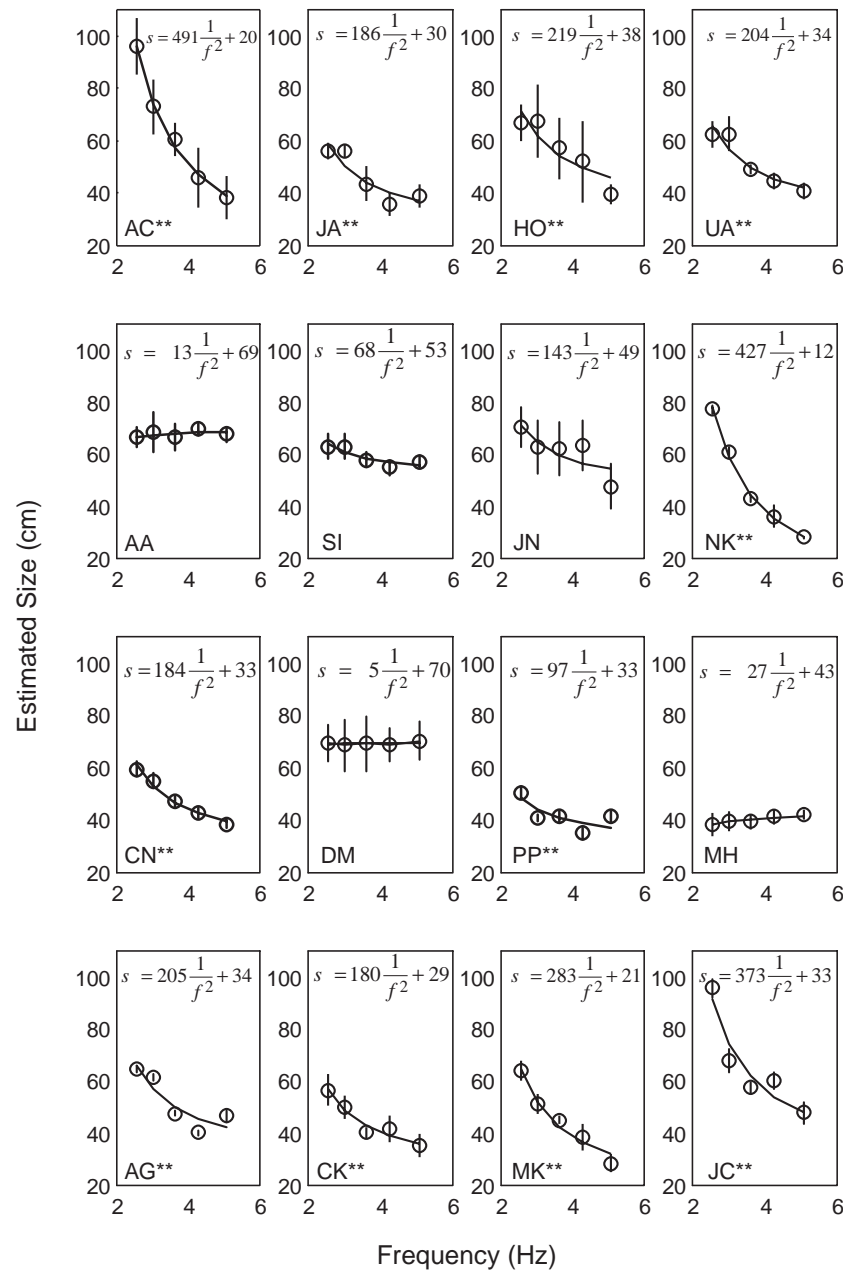
Figure 5. Mean estimated size for each observer for each simulated stride frequency in Experiment 2 (n = 11). Error bars indicate SEM. The graph corresponds to the fitted model to each observer individually. *p < .05; **p < .01 indicates the level of significance of the correlation between the model and individual size estimations.

independent variable (i.e., the stride frequency). Two observers (J.N. and D.M.) also showed very large variances across similar stimulus repetitions, which indicates that they responded in a disoriented manner. Observers from this group also gave the largest and smallest values for the size of the statically displayed dog. Consequently, for some of them, very low (and in two cases even negative) values for λ are obtained.

Disregarding the five participants that did not show any meaningful behavior, the results show that the inverse quadratic relation between characteristic size and stride frequency is employed by the visual system when

estimating the size of an animal in the absence of other cues.

General Discussion

As summarized above, previous experimental work has shown that observers are able to judge object size in inanimate dynamic systems governed by gravity. The experiments reported here provide the first empirical evidence that those findings can be extended to the domain of animate motion as well. The human visual system uses the physically determined relation between

spatial and temporal scales to obtain the size of a moving animal in the absence of other cues.

In both experiments conducted to test the spatio-temporal scale hypothesis, we found the predicted effect of stride frequency on perceived size.

Nevertheless, when investigating the individual size estimations in terms of the parameters of the proposed model, substantial interindividual differences became evident. These differences were more pronounced in Experiment 1 than in Experiment 2. The results obtained in the modified setting show that observers retrieved the motion-mediated size information more efficiently. The data show less intersubject variability and larger values for k_1 when compared to Experiment 1, in which we had attempted to provide a method for transforming observers' size impression into a corresponding response while maintaining a constant retinal size of the stimulus.

In the two experiments reported here, we presented to the observers a single scaling relation between time and space with the requirement to yield judgment of spatial scale based on temporal variations. One might argue that observers simply assign numbers to the temporal variations without really detecting these variations as information about scale. However, if this were the case, one would expect observers to assign the direction of the mapping between time and space arbitrarily. Only one of 32 observers showed a reversed correlation between perceived size and stride frequency. Moreover, we found a quadratic relation rather than a simple linear one, which reflects the physical properties of the temporal spatial relation. Simply assigning numbers to temporal variations would probably lead to a linear relation instead of a quadratic one.

Altogether, seven observers in Experiment 1 and five observers in the optimized setting in Experiment 2 neglected the temporal-spatial scaling relation by showing a random pattern in their results. A reason for this pattern of results might be the methodological approach. We used a method similar to Pittenger (1985), in which participants were given only timing as information about spatial scale in pendulum motion. Pittenger's results were similar to the current results in that they were noisy with strong individual differences. In a related study concerning pendulum motion (Pittenger, 1990), the observers were given precise information about spatial scale, but the timing of the event was manipulated to be either consistent or inconsistent with the pendulum law. Rather than having to readjust the correct timing, observers had to judge only its correctness. Observers performed with high accuracy on this task. According to Pittenger's results, observers seem to be more sensitive to violation of the temporal-spatial scaling relation than to transforming temporal information about spatial parameters into size judgments. A similar effect may have also played a role in our setup.

Given constant stride length, a higher stride frequency goes along with a higher locomotion speed. One might be concerned about this confoundation of stride frequency and locomotion speed, arguing that the current results could depend on simple translational speed rather than on the details of the gait itself. In a previous study (Jokisch, Midfort & Troje, 2001), we used point-light displays of biological motion of dog animations, having subtracted the translational motion component. Consequently, the position of the point-light animal remained constant in the center of the screen. Varying the stride frequency, we found a significant effect on perceived size. Therefore, we are confident that the crucial source conveying size information in the experiments we are reporting here is the stride frequency itself.

Nevertheless, we cannot entirely exclude that translational speed may contribute to the size judgment. In a natural display stride frequency, locomotion speed and stride length cannot be unconfounded. However, we did not want to make any issue about the details of the perceptual cues used to derive size from biological motion. Instead, we wanted to test whether the human visual system is able to employ the relation between temporal and spatial scales, which is physically defined through gravitational acceleration.

Human observers seem to be able to employ the general inverse quadratic relation between size and stride frequency to derive information about size from temporal parameters. In addition to this qualitative result, the measurements taken in Experiment 2 can also be used to make quantitative comparisons between the absolute size indicated by the observers and the size of real animals that walk with the respective stride frequencies. The relation between size and stride frequency of walking animals is expressed by the factor c_1 in Equation 3. Summarizing the results of Experiment 2, we compute c_1 as the median of the 11 observers that did respond in a consistent manner. The resulting value amounts to 435 cm s^{-2}.

Unfortunately, the only set of data that we are aware of which can be used to derive the spatio-temporal relation factor from natural locomotion patterns is the one reported by Pennycuick (1975), who compared stride frequencies and shoulder heights of 14 African quadruped mammal species for different gait patterns. The smallest animal in this study (Thomson's gazelle) had a shoulder height of 60 cm; the largest one (elephant) had a shoulder height of 310 cm. From Pennycuick's Figure 13, we calculated c_1 to amount to 410 cm s^{-2} for cantering animals. This value is very close to the one obtained from our data.

The close matching between the empirical data for cantering animals (Pennicuick, 1975) and the data obtained in our experiments seems to imply that the human visual system not only takes into consideration the general inverse quadratic relation between stride

frequency and size but also takes advantage from implicit knowledge about the particular observed gait pattern. We want to note, however, that the good quantitative fit between Pennicuick's and our data may well be accidental. There are a number of factors that introduce uncertainty into the absolute value of the spatio-temporal scaling factor c_1 as derived from our experiments. For instance, the perceived height of the reference objects in the scenery may deviate from their "real" height. The posts were intended to have a height of 1 m and the cactuses a height of 2 m. Those numbers were given to the observers in their introduction to the experiment. However, the reference objects may still have been perceived to be larger or smaller, changing the reference frame used to indicate the dog's size. Another critical point is the determination of the constant c_2 in Equation 5. In the second subtask of Experiment 2, we tried to measure the perceived size as given by cues that are independent from stride frequency. We did that by asking the observers to estimate the size of a static stick-figure display. However, this procedure may not be sufficient to accurately derive the desired information. It is still possible that a moving dog does provide cues about its size, which are not available in the static display but which are still not depending on the stride frequency. A last factor that adds uncertainty is the fact that living animals, even if they try to minimize energy consumption during locomotion, are still different from inanimate dynamic systems. In a swinging pendulum or a bouncing ball, the relation between temporal and spatial parameters is exactly defined by gravity, because no other forces affect these motions. In contrast, in dynamic animate systems, muscular forces controlled by intentional behavior play an important role. They are not used only to simply compensate for damping effects in the articulated pendulum system of the body; they can also be used to significantly alter the motion pattern to cover a wider range of stride frequencies within a given gait pattern.

Conclusions

In summary, we can state that human observers are able to employ implicit knowledge about the general inverse quadratic relation between size and stride frequency to derive information about the size of an animal from temporal parameters. The exact scaling of this relation is dependent on a number of parameters that are beyond the control of our current experiments. We are therefore critical with respect to the perfect accordance of our data with quantitative predictions involving knowledge about the biomechanics of particular quadruped gaits. It would be interesting, however, to measure whether the perceived size of animals traveling with a given stride frequency changes in a predictable way as a function of the gait pattern.

Acknowledgments

References

Alexander, R. M. (1977). Mechanics and scaling of terrestrial locomotion. In T. J. Pedley (Ed.), *Scale Effects in Animal Locomotion* (pp. 93-110). New York: Academic Press.

Alexander, R. M. (1984). The gaits of bipedal and quadrupedal animals. *The International Journal of Robotics Research, 3*, 49-59.

Alexander, R. M. (1989). Optimization and gaits in the locomotion of vertebrates. *Physiological Reviews, 69*, 1199-1227. [PubMed]

Alexander, R. M., & Jayes, A. S. (1983). A dynamic similarity hypothesis for the gaits of quadrupedal mammals. *Journal of Zoological Society of London, 201*, 135-152.

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics, 23*, 145-152. [PubMed]

Beardsworth, T., & Buckner, T. (1981). The ability to recognize oneself from a video recording of one's movements without seeing one's body. *Bulletin of the Psychonomic Society, 18*, 19-22.

Bingham, G. P. (1987) Kinematic form and scaling: Further investigations on the visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 155-177. [PubMed]

Bingham, G. P. (1993a) Scaling judgments of lifted weight: Lifter size and the role of the standard. *Ecological Psychology, 5,* 31-64.

Bingham, G. P. (1993b) Perceiving size of trees: Form as information about scale. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1139-1161.

Bingham, G. P. (1993c) Perceiving size of trees: Biological form and the horizont ratio. *Perception and Psychophysics, 54,* 485-495. [PubMed]

Blake, R. (1993). Cats perceive biological motion. *Psychological Science, 4*, 54-57.

Cavagna, G. A., Thys, H., & Zamboni, A. (1976). The sources of external work in level walking and running. *Journal of Physiology, 262*, 639-657. [PubMed]

Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception, 7*, 393-405. [PubMed]

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society, 9*, 353-356.

Dittrich, W. H., Lea, S. E. G., Barrett, J.,& Gurr, T. R. (1998). Categorization of natural movements by pigeons: Visual concept discrimination and biological motion. *Journal of the Experimental Analysis of Behavior*, *70*, 281-299.

Hecht, H., Kaiser, M. K., & Banks, M. S. (1996). Gravitational acceleration as a cue for absolute size and distance? *Perception & Psychophysics, 58*, 1066-1075. [PubMed]

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

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research, 38*, 379-393. [PubMed]

Jokisch, D., Midford, P.E., & Troje, N. F. (2001). Biological motion as a cue for the perception of absolute size [Abstract]. *Journal of Vision, 1*(3), 357a. http://journalofvision.org/1/3/357, DOI 10.1167/1.3.357. [Abstract]

Kram, R., Domingo, A., & Ferris, D. P. (1997). Effect of reduced gravity on the preferred walk-run transition speed. *The Journal of Experimental Biology, 200*, 821-826. [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.

Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London Series B, 258*, 273-279.

Mather, G., & West, S. (1993). Recognition of animal locomotion from dynamic point-light displays. *Perception, 22*, 759-766. [PubMed]

McConnell, D. S., Muchisky, M. M., & Bingham, G. P. (1998). The use of time and trajectory forms as visual information about spatial scale in events. *Perception & Psychophysics, 60*, 1175-1187. [PubMed]

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

Pennycuick, C. J. (1975). On the running of the gnu (*Connochaetes taurinus*) and other animals. *Journal of Experimental Biology*, 63, 775-799.

Pinto, J., & Shiffrar, M. (1999). Visual analysis of human and animal biological motion displays [Abstract]. *Abstracts of the Psychonomic Society, 4,* 1.

Pittenger, J. B. (1985). Estimation of pendulum length from information in motion. *Perception, 14*, 247-256. [PubMed]

Pittenger, J. B. (1990). Detection of violations of the law of pendulum motion: Observers' sensitivity to the relation between period and length. *Ecological Psychology, 2*, 55-81.

Pittenger, J. B., & Todd, J. T. (1983) Perception of growth from changes in body proportions. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 945-954. [PubMed]

Runeson, S., & Frykholm, G. (1981) Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 733-740. [PubMed]

Runeson, S., & Frykholm, G. (1983) Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition and deceptive intention. *Journal of Experimental Psychology: General, 112,* 585-615.

Saxberg, B. V. (1987a). Projected free fall trajectories. I. Theory and simulation. *Biological Cybernetics, 56*, 159-175. [PubMed]

Saxberg, B. V. (1987b). Projected free fall trajectories. II. Human experiments. *Biological Cybernetics, 56*, 177-184. [PubMed]

Sedgwick, H. A. (1993) The effects of viewpoint on the virtual space of pictures. In S. R. Ellis, M. K. Kaiser & A. Grunwald (Eds.), *Pictorial communication in virtual and real environments*. New York: Taylor & Francis.

Stappers, P. J., & Waller, P. E. (1993). Using the free fall of objects under gravity for visual depth estimation. *Bulletin of the Psychonomic Society, 31*, 125-127.

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]

Warren, W. H., & Kim, E. E., & Husney, R. (1987) The way the ball bounces: Visual and auditory perception of elasticity and control of the bounce pass. *Perception, 16,* 309-336. [PubMed]

Watson, J. S., Banks, M. S., von Hofsten, C., & Royden, C. S. (1992). Gravity as a monocular cue for perception of absolute distance and/or absolute size. *Perception, 21*, 69-76. [PubMed]

Yamaguchi, M.K., & Fujita, K. (1999) Perception of biological motion by newly hatched chicks and quail. *Perception, 28(Suppl.),* 23-24.