Perceptual Effects of Inconsistency in Human Animations

SOPHIE KENNY, Queen's University

NAUREEN MAHMOOD, Max Planck Institute for Intelligent Systems

CLAIRE HONDA, Queen's University

MICHAEL J. BLACK, Max Planck Institute for Intelligent Systems

NIKOLAUS F. TROJE, Queen's University, York University

The individual shape of the human body, including the geometry of its articulated structure and the distribution of weight over that structure, influences the kinematics of a person's movements. How sensitive is the visual system to inconsistencies between shape and motion introduced by retargeting motion from one person onto the shape of another? We used optical motion capture to record five pairs of male performers with large differences in body weight, while they pushed, lifted, and threw objects. From these data, we estimated both the kinematics of the actions as well as the performer's individual body shape. To obtain consistent and inconsistent stimuli, we created animated avatars by combining the shape and motion estimates from either a single performer or from different performers. Using these stimuli we conducted three experiments in an immersive virtual reality environment. First, a group of participants detected which of two stimuli was inconsistent. Performance was very low, and results were only marginally significant. Next, a second group of participants rated perceived attractiveness, eeriness, and humanness of consistent and inconsistent stimuli, but these judgements of animation characteristics were not affected by consistency of the stimuli. Finally, a third group of participants rated properties of the objects rather than of the performers. Here, we found strong influences of shape-motion inconsistency on perceived weight and thrown distance of objects. This suggests that the visual system relies on its knowledge of shape and motion and that these components are assimilated into an altered perception of the action outcome. We propose that the visual system attempts to resist inconsistent interpretations of human animations. Actions involving object manipulations present an opportunity for the visual system to reinterpret the introduced inconsistencies as a change in the dynamics of an object rather than as an unexpected combination of body shape and body motion.

CCS Concepts: • **Computing methodologies** → **Motion capture**; • **Applied computing** → *Psychology*;

Additional Key Words and Phrases: Retargeting, human animation, animated avatars, shape capture, realism, inconsistency, perception, discrimination, action

This work was supported by a NSERC Discovery grant and the Humboldt Research Award to NFT, as well as a NSERC Alexander Graham-Bell Graduate Scholarship to S.K. Additional travel funds were provided to S.K. by the International Research and Training Grant "The Brain in Action" (CREATE/DFG). M.J.B. has received research gift funds from Intel, Nvidia, Adobe, Facebook, and Amazon. While M.J.B. is a part-time employee of Amazon, his research was performed solely at, and funded solely by, MPI. N.M. is a founder and CEO at Meshcapade GmbH, which is commercializing body modeling technology. Her research was performed solely at, and funded solely by, MPI.

Authors' addresses: S. Kenny and C. Honda, Queen's University, Department of Psychology, Kingston, Canada, K7L 3N6; emails: kenny.s, claire.honda@queensu.ca; N. Mahmood and M. J. Black, Max Planck Institute for Intelligent Systems, Tübingen, Germany, 72076; emails: nmahmood@tue.mpg.de, black@tuebingen.mpg.de; N. F. Troje, Queen's University, Department of Psychology, Kingston, Canada, K7L 3N6 and York University, Center for Vision Research, Toronto, Canada, M3J 1P3; email: troje@yorku.ca.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2019 Copyright held by the owner/author(s).

1544-3558/2019/02-ART2

<https://doi.org/10.1145/3301411>

ACM Reference format:

Sophie Kenny, Naureen Mahmood, Claire Honda, Michael J. Black, and Nikolaus F. Troje. 2019. Perceptual Effects of Inconsistency in Human Animations. *ACM Trans. Appl. Percept.* 16, 1, Article 2 (February 2019), 18 pages.
<https://doi.org/10.1145/3301411>

1 INTRODUCTION

When looking at other people and assessing their actions and behaviour, observers are likely to employ a wealth of prior knowledge about statistical dependencies between body build and kinematic aspects of body movements [5]. Size, weight, relative limb length, and the distribution of mass over the body all influence the way a body moves. Observers have been shown to use shape and motion information in a variety of people perception tasks, including sex discrimination [26, 45, 46], person identification [50, 52], and action perception [2, 3, 8, 31, 35, 36, 39, 40]. According to the Kinematic Specification of Dynamics principle [35, 36], observers learn to relate the visually available kinematic patterns of action with the latent forces that would have been exerted to produce the observed motion pattern. Learning occurs through an observer's life-long personal experience with manipulating their natural environment and observing other people in similar situations [10]. Innate predispositions and sensory filter mechanisms may also play a role [20, 48].

Given the importance of both body shape and body motion as sources of information for people perception, it is likely that the visual system is sensitive to internal inconsistencies introduced by perturbing the relationship between shape and motion. In computer animation, internal inconsistencies can arise as a result of animation retargeting. In this procedure, the recorded motion of a performer is used to animate a humanoid computer graphics character. In the general case, the shapes of the performer and of the animated character will differ [11]. The aim of animation retargeting procedures is to preserve the motion of the performer such that the high-level features of an action are preserved in the animated character [9, 30]. However, retargeting the motion of a performer onto a character with a different body shape is likely to impact perception, because the shape of the character implies dynamic properties that generally differ from those of the performer [47].

In a recent review on empirical studies on the notion of the “uncanny valley,” the discomfort sometimes elicited by renderings of people who appear almost real but not quite so, Kättsyri and colleagues [22] compared evidence for a number of different theories that may underly the uncanny valley effect. The authors discuss empirical work [14, 32, 41] in favour of the “mismatch hypothesis.” The theory suggests that the negative responses associated with the uncanny valley stem from inconsistencies between different aspects that contribute to the rendering. The visual system expects a human shape to move like a human and a robot shape to move like a robot. It follows that when a robotic motion is used to animate a human body shape, the resulting animation should be perceived as less attractive and more eerie than when the same robotic motion is used to animate a robotic body shape [38]. It has also been suggested that shape-motion inconsistency might trigger a warning mechanism that implies that an individual differs from the norm or may be unpredictable in some way. As such, shape-motion consistency influences perceptual judgements and biases participants to perceive an animated avatar to suffer from a biological insult [15] to present psychological pathologies [43] or to have an atypical sexual orientation [19].

The perceptual mismatch hypothesis has resonated with animators and researchers alike. It is used both as an explanation for why technologically impressive human animations can fail to successfully appeal to their audiences [4, 42] and as a pre-emptive rationale for animators and engineers who purposefully reduce the human-likeness of androids or avatars to increase their acceptability [25, 27, 51].

Perception researchers using computer animations have also shown concern for the potential influence of shape-motion inconsistency. Some researchers routinely acknowledge that large shape-motion inconsistencies might negatively bias the perception of animations in their studies. They anticipate the issue by designing re-targeted stimuli whose shapes and motions are as consistent as possible. In one example, McDonnell et al. [28] studied the perception of animated crowds of walking humans and noted that they took care to select performers

whose body shapes, ages, and sexes were roughly similar to each of their 20 unique avatars. In a related example, Hoyet and colleagues [17] studied the perception of motion of multiple performers retargeted onto a single avatar and carefully matched the body shapes of the performers to the single avatar to avoid confounds due to inconsistency. However, the extent to which the human visual system is sensitive to uncorrected discrepancies is unknown, since the effect of inconsistency is rarely studied with realistic human avatars.

Instead, research into human action perception has often been conducted using point-light displays [18], which reduce the rendering of a human actor to a relatively small number of dots representing the major joints of the body. The only information about body shape that can be directly obtained from these displays is based on relative limb lengths and the coarse geometry of the articulated array of joints. However, even these reduced displays are sufficient to demonstrate that observers have expectations about the relations between shape and motion when assessing human motion. A study that provided clues about the ability of the visual system to detect shape-motion inconsistency was conducted by Runeson and Frykholm [36]. The authors demonstrated that observers could identify a performer's attempts to move in a way that is inconsistent with their own body shape—for instance, when they pretended to move like a person of the opposite gender. Studies have shown that the way men and women move is partially determined by the dynamics of their body shape [45, 46]. As a result, mimicking the opposite gender disrupts the expected relationship between body shape and motion. For another example, observers seem to be able to use knowledge of the physically determined relationship between stride length and stride frequency of a walking animal to appropriately determine the size of an animal [21]. It was also shown that naturalness ratings of point-light walkers presented at a veridical, sped-up, or slowed-down playback speed betray the visual system's sensitivity to inconsistency: Observers rate the veridical playback speeds of point-light walkers as most natural [49]. In yet another experiment, Klüver et al. [23] demonstrated that exchanging the shape and motion information between pairs of point-light walkers directly resulted in lower attractiveness for hybrid walkers compared to consistent walkers.

Information from movement kinematics is often sufficient to support the perception of action outcomes. For example, throwing distance of invisible objects can be inferred from the arm kinematics of point-light displays [31, 36, 53]. For other actions, body shape is more likely to influence how kinematic information is interpreted by the visual system. For instance, although kinematics support the perception of a lifted weight [2, 8, 39], the weight of the performer may modulate how kinematics are used. According to Runeson and Frykholm [35], light performers who lift increasingly heavy boxes adjust their posture to stabilize themselves as a counterweight to the box, while heavier performers do not require postural adjustments as they have a more stable base of support. These examples highlight the action-specificity in the way shape and motion are integrated in the perception of action outcomes.

The above-mentioned studies all employed point-light displays. Similar, well-controlled work with more realistic animations that provide a richer visualization of individual body shape is rare. The reason for this might be rooted in the difficulty of generating a sufficiently large dataset that contains body shape and body kinematic from the same individuals. Traditionally, body kinematics is sampled with optical motion capture technology. Body shape can be assessed with three-dimensional (3D) whole-body scanners. Few labs have both and none has produced a comprehensive dataset yet.

Recently Loper and colleagues [24] devised a new technique that fits a parametric model of body shape and pose directly to 3D motion capture markers. The body shape is constrained to lie in the space of human body shapes, which is represented by a SCAPE body model [1] learned from approximately 4,000 3D body scans [13]. MoSh, as this technique has been named, at least in its current state and given the marker set that we used, is not able to reconstruct idiosyncratic details of hands and face, but the system recovers individual 3D body shape and its dependencies on pose with a high degree of accuracy [24].

In summary, evidence suggests that animation retargeting may result in a disruption of the normal shape-motion integration process if the body shapes of the performer and of the CG character are dissimilar. In the current article, we analyze the effects of stimulus inconsistency introduced by retargeting motion between

Fig. 1. Example of a height-matched pair of light and heavy performers.

Table 1. Characteristics of the Recorded Performers and the Resulting Pairs

Performer	Pair ID	Age	Weight (kg)	Weight Category	Height (m)	BMI
1	A	29	66.4	Light	1.78	21.0
2	A	20	112	Heavy	1.78	35.5
3	B	26	73.3	Light	1.82	22.3
4	B	21	113	Heavy	1.80	34.9
5	C	20	72.3	Light	1.86	20.9
6	C	22	125	Heavy	1.84	37.0
7	D	24	60.4	Light	1.84	17.9
8	D	19	101	Heavy	1.86	29.3
9	E	25	77.1	Light	1.88	21.8
10	E	25	91.3	Heavy	1.97	23.7

people with different body masses. Both motion and body shape were obtained from motion capture data using MoSh [24]. To uncover the perceptual consequences of retargeting animations, we conducted three experiments that employed different measures of sensitivity to inconsistencies between body shape and kinematics in human animations. In the first experiment, we studied the ability of observers to explicitly identify such inconsistencies. In the second experiment, we examined the effect of inconsistency on judgments of three animation characteristics. Finally, in the third experiment we assessed the effect of inconsistency on perceived action outcomes. We begin by describing the general methods that were constant to each experiment, before describing the three experiments in succession.

2 GENERAL METHODS

2.1 Performers

We recruited 10 male performers who varied widely in body shape. We restricted ourselves to male performers to focus the variance among our stimuli on the weight differences and to avoid confounding these differences with sex-related interactions. We then paired the 10 performers into five groups. Our goal was to maximize the difference in weight between the two individuals of each pair while keeping their body heights as similar as possible (Figure 1, Table 1). The five light performers weighed on average 69.9kg ($SD = 6.55$ kg), and the five heavy performers weighed on average 108kg ($SD = 12.83$ kg). A paired sample t -test indicated that the weights differed significantly ($t(4) = -5.92$, $p = 0.004$). Neither height nor age were significantly different between light and heavy performers.

The 38kg difference between the light and heavy performers is accompanied by large differences in body mass distribution. According to the guidelines of the Canadian Medical Association [6], the Body Mass Index and waist circumference measurements of the light and heavy performers indicate large differences in body fat distributions for the two groups. The light performers had normal Body Mass Indices ($M = 20.8\text{kg/m}^2$, $SD = 1.69\text{kg/m}^2$)

Fig. 2. Realistic human animations of performers recorded while pushing, throwing, and lifting objects.

and a narrow waist circumference ($M = 77.2\text{cm}$, $SD = 2.86\text{cm}$), while the heavy performers had obese class I Body Mass Indices ($M = 32.0\text{kg/m}^2$, $SD = 5.54\text{kg/m}^2$) and wide waist circumferences ($M = 102\text{cm}$, $SD = 10.1\text{cm}$).

Performers were recorded executing the following three actions: pushing a training sled weighing between 23kg to 123kg, in 20kg increments, over a distance of 2m; throwing beanbags weighing 200g toward targets located 2m to 6m away, in 1m increments; and lifting boxes situated two step-lengths away from them, on the floor, that weighed from 1kg to 11kg, in 2kg increments (Figure 2).

2.2 Motion and Shape Estimation

We used an 18-camera optical motion capture system (Vicon iQ, Vicon Motion Systems Ltd.) to record the actions of 10 male volunteers from Queen's University, Canada. The marker set that we used followed the suggestions given in Reference [24]. Data were captured at 120fps.

Next, we used the MoSh algorithm [24] to obtain correlated estimates for body shape and body motion for each performer, each action, and each object weight and throwing distance. MoSh produces a 3D surface composed of a triangulated mesh with 10,777 vertices, using a data-driven model that parameterizes the interindividual variance in body shape, as well as pose-dependent body shape deformations. MoSh achieves reconstruction errors that are smaller than 1cm (mean distance between true and estimated body surface) [24].

2.3 Creation of Consistent and Inconsistent Animations

Consistent stimuli obtained shape and motion data from the same performer, while inconsistent stimuli were generated by hybridizing the shape of one performer with the motion of another. To maximize the effect of hybridization, we always combined shape and motion of performers that were drawn from different weight groups. Therefore, the inconsistent stimuli showed animated characters with the shape of a heavy performer and the motion of a height-matched light performer or the shape of a light performer combined with the motion of a height-matched heavy performer (Figure 3).

The combination of 10 actors, 6 different weights, and both consistent and hybrid versions resulted in a stimulus set that consisted of 120 stimuli each for the box lifting task and for the sled pushing task. Since there were only 5 different distances in the throwing task, this resulted in a stimulus set of 100 different throwing stimuli. See the supplementary video for examples of the stimuli.

2.4 Presentation Apparatus

Stimuli were presented to participants in Experiments 1, 2, and 3 while they sat in front of a keyboard and wore a head-mounted display (Oculus Rift DK2, runtime version 0.8.0.0, experiment designed in Unity3D, version 5.1.4 with the Unity Oculus Plugin).

The headset simulated a stereoscopic 3D depiction of a testing room. Real-time head tracking was used to enable motion parallax and to allow participants to actively explore the scene by moving and turning their head.

Fig. 3. Consistent stimuli were animated characters whose shape and motion were taken from the same performer, and inconsistent stimuli were animated characters whose shape and motion were taken from height-matched performers with dissimilar body weights.

However, active head movements were not required to maintain the stimuli within view. All characters were rotated 20° counterclockwise from facing the participant. All actions were presented such that at the end of the animation, the performer concluded at a distance of 4m in front of the observer. Since the actions of pushing and lifting included locomotion, their animations were initiated farther away from the observers. The maximal initial distance was about 6m away for the pushing actions.

3 EXPERIMENT 1: DETECTION OF INCONSISTENT HUMAN ANIMATIONS

In the first experiment, we tested the detectability of shape-motion inconsistency. We presented pairs of consistent and inconsistent animations and asked observers to identify the inconsistent stimulus. We predicted that when forced to make a direct comparison between the two stimuli, participants would be able to detect the conflicting dynamics. Based on work that demonstrates that variability in visual appearance reduces the ability to respond to subtle variations in motion [29], we also predict that detection would be easier if the two characters were shown with differing motions as compared to differing body shape.

3.1 Methods

3.1.1 Participants. Twenty-four volunteers from Queen’s University were recruited. Participants were 15 females and 9 males, and they were between 18 and 24 years of age ($M = 19.4$ years old, $SD = 1.58$). Participants were tested individually and reported normal or corrected-to-normal vision. None had taken part in the motion capture sessions from which the stimuli were created.

3.1.2 Procedure. The stimuli were the five height-matched pairs of performers described in Section 2.1, but we used only the highest level of object property of the three actions: the 123kg sled pushes, the 6m throws, and the 11kg box lifts. With 10 performers and three action clips for each, we used a total of 30 veridical, consistent stimuli and their respective hybridized versions.

We informed participants that they would be seeing 10 male performers lifting invisible boxes, pushing invisible sleds, and throwing invisible beanbags. Pairs of consistent and inconsistent animations were presented to participants, separated by a 500ms interstimulus interval. We instructed participants to select the “mismatching” animation, i.e., to indicate which of the two animations of a pair was created using the appearance of one performer, but the movement of another performer. Each stimulus pair was presented twice for a total of 60 trials per block. Two blocks were run, for a total of 120 trials. In one block, the same shape was used for both stimuli of a pair, and the motion differed. In the other block, the same motion was used for both stimuli of a pair, and the shape differed. Participants were aware of this manipulation. The order in which the two stimuli of a pair were presented was random. The order of the trials within a block was also random. The order of the two blocks was counterbalanced across participants (Figure 4).

Fig. 4. Participants indicated which of the two stimuli they believed was the mismatching, retargeted animation. Actions were randomly presented within a block. In one block, the body shape remained the same between the two stimuli, and in the other, the motion remained the same.

Fig. 5. Correct response rates across 24 participants for the inconsistent stimulus detection task. Participant performance was marginally better than chance (.50) in the box pushing and lifting animations. Error bars represent ± 1 SEM.

3.2 Results

The results of the inconsistency discrimination task are presented in Figure 5. Performance was in all cases very close to chance level. One-tailed one-sample t -tests comparing the correct response rates to chance level were not statistically significant for object throws ($M = 0.497$, $SEM = 0.015$, $t(23) = -0.216$, $p = 0.584$) and only marginally significant for box lifts ($M = 0.527$, $SEM = 0.016$, $t(23) = 1.72$, $p = 0.050$), and sled pushes ($M = 0.540$, $SEM = 0.023$, $t(23) = 1.76$, $p = 0.046$).

A two-way repeated-measures ANOVA with actions and blocks as factors did not provide evidence that performance varied as a function of the action viewed or the information rated. There was no main effect of action: Participants selected the inconsistent stimuli over the consistent stimuli at an equally low rate across all actions, $F(2, 46) = 1.76$, $p = 0.184$, $\eta_p^2 = 0.071$. There was also no main effect of blocks: Participant performance was the same when either the shape or the motion information had been manipulated within a pair of stimuli, $F(1, 23) = 0.141$, $p = 0.711$, $\eta_p^2 = 0.006$. There was also no interaction between the two factors, $F(2, 46) = 0.892$, $p = 0.417$, $\eta_p^2 = 0.037$.

Participants rated each of the 20 unique pairs twice for each of the three performed actions. We computed a Pearson's r correlation separately for each of the three actions to examine the reliability of participant ratings across these two repetitions. The correlation between the animation chosen as inconsistent was significant for

throwing, $r(478) = 0.18$, $p < 0.001$, but neither for pushing, $r(478) = 0.07$, $p = 0.120$, nor for lifting, $r = 0.02$, $p = 0.606$.

3.3 Discussion

Participants struggled to identify which of the two stimuli was inconsistent. This was evidenced by low correct detection rates and generally unreliable choices. This was an unexpected finding given the large differences between the body mass of the two performers that contributed to the hybrid. The finding also contrasts those of Runeson and Frykholm [36], who found that observers could detect instances of deceptive performances in cases where performers attempted to move in ways that were typical of the opposite gender. On the other hand, the situation in our study differs slightly from that of Runeson and Frykholm [36]. In their study, all stimuli were consistent in the sense that motion and shape came from the same recording. Observers were asked to identify a deliberate deception by the actor rather than the editing of the recording by the experimenter. Identification of the deception might have also have been facilitated by the fact that the actors were laypeople rather than trained stage performers.

We had also expected that performance would be better when the shape of the performer remained constant across a stimulus pair. Instead, correct response rates were equally low for stimuli pairs where the shape varied compared to stimulus pairs between which the motion varied.

Even though observers were not able to explicitly identify which of the two stimuli was the hybrid, other perceptual measures may still be influenced by a disruption of the shape-motion integration process. Studies have shown that inconsistency can affect the perception of characteristics such as attractiveness and naturalness of point-light displays and animated characters [12, 33, 34]. This is relevant for entertainment media, where the affective responses toward animations is important. Therefore, in Experiment 2, we examined whether the effects of inconsistency on perception that were observed in past point-light display experiments are replicated with our realistic animated characters.

4 EXPERIMENT 2: EFFECT OF INCONSISTENCY ON JUDGEMENTS OF ANIMATION CHARACTERISTICS

In this experiment, we tested a prediction derived from the premise of the perceptual mismatch hypothesis [22]. We propose that inconsistent animations, created by hybridizing shape and motion from performers with significantly different body weights, could result in negative responses when compared to consistent animations. To study this question, we collected judgements of attractiveness, eeriness, and naturalness for consistent and inconsistent animations. These questions constituted a simplified version of a questionnaire that Ho and MacDorman [14] developed in an attempt to provide a quantitative approach to the “uncanny valley” phenomenon.

4.1 Methods

4.1.1 Participants. For this experiment, we recruited 24 new volunteers (13 females and 11 males) from Queen’s University, Kingston, Canada. Participants were between 18 and 26 years old ($M = 20.4$, $SD = 2.48$), and reported normal or corrected-to-normal vision. None had taken part in the motion capture sessions from which the stimuli were created or in Experiment 1.

4.1.2 Procedure. Participants were presented the same 60 stimuli used in Experiment 1 (see Section 3.1.2) in random order. After each stimulus, participants were prompted to rate their perception of the stimulus according to one of three characteristics: attractiveness (unattractive to attractive), eeriness (reassuring to eerie), or humanness (artificial to natural). Participants indicated their response on a continuous rating scale anchored between -3 and $+3$. We presented the three questions in a blocked design, for a total of 180 trials, and counterbalanced the order of the three blocks across participants (Figure 6).

Fig. 6. Experimental design of the judgement of animation characteristics rating task. Following the presentation of each stimulus, participants rated the attractiveness, eeriness, and perceived humanness on semantic differential items, using a continuous response scale anchored between -3 and $+3$. All three actions were presented with trials in random order.

Fig. 7. Mean attractiveness, eeriness, and humanness for the animation characteristics rating task. Participants rated either consistent or inconsistent stimuli, who lifted a box, pushed a sled, or threw a beanbag. The error bars represent ± 1 SEM. Follow-up two-tailed paired-sample t -tests were performed. $N = 24$.

4.2 Results

The results of the animation characteristics rating task are presented in Figure 7. There was no effect of consistency for any of the rated animation characteristics: Consistent and inconsistent versions of the animations were rated similarly in terms of their attractiveness, eeriness, and humanness. However, there was a strong effect of the action on all three rating questions: Box lifts and sled pushes were perceived as more attractive, less eerie, and more human than the object throws.

Table 2. 2×3 Repeated-measures ANOVAs Conducted Separately for the Three Judgements of Animation Characteristics

Source of Variance	SS	df	MS	F	p	η_p^2	Obs. power
Attractiveness							
Consistency	0.078	1	0.078	0.643	0.431	0.027	0.120
Error	2.77	23	0.121				
Action	8.67	2	4.34	7.36	0.002	0.242	0.924
Error	27.1	46	0.589				
Consistency * Action	0.202	2	0.101	1.12	0.335	0.046	0.235
Error	4.15	46	0.090				
Eeriness							
Consistency	0.114	1	0.114	1.41	0.248	0.058	0.206
Error	1.86	23	0.081				
Action	9.86	2	4.93	8.00	0.001	0.258	0.944
Error	28.3	46	0.616				
Consistency * Action	0.187	2	0.093	1.48	0.237	0.061	0.300
Error	2.89	46	0.063				
Humanness							
Consistency	0.062	1	0.062	0.0454	0.507	0.019	0.099
Error	3.15	23	0.137				
Action	29.6	2	14.8	23.9	<0.001	0.510	1.00
Error	28.5	46	0.620				
Consistency * Action	0.195	2	0.098	1.10	0.343	0.046	0.231
Error	4.10	46	0.089				

These observations were confirmed by two-way repeated-measures ANOVAs that were conducted separately for the three rating questions (Table 2). There was no main effect of consistency of animation, suggesting that inconsistency did not reduce perceived attractiveness, increase eeriness, or change the perceived humanness of the animations. However, there was a main effect of action in all three rating questions. Follow-up comparisons of rated animation characteristics pairs of actions, illustrated in Figure 7, indicated that ratings of box lifts and sled pushes were similar to each other on all three rating questions, while box lifts and sled pushes were perceived as more attractive, less eerie, and more human than beanbag throws. There was no interaction between consistency of animations and actions for any of the three judgements of animation characteristics.

We also examined the relationship between the three ratings obtained on individual animations ($N = 60$) for attractiveness ($M = -0.0125$, $SD = 0.559$), eeriness ($M = -0.260$, $SD = 0.600$), and perceived humanness ($M = 0.282$, $SD = 0.818$). Perceived humanness was positively correlated with attractiveness ($r(58) = 0.867$, $p < 0.001$) and negatively correlated with eeriness ($r(58) = -0.676$, $p < 0.001$), and attractiveness was negatively correlated with eeriness ($r(58) = -0.651$, $p < 0.001$).

4.3 Discussion

The predictions of the perceptual mismatch hypothesis [22] were not confirmed by our data. Introducing inconsistency between the shape and motion of the animations did not reduce perceived attractiveness, did not make them appear less human, and did not increase the eeriness of the stimuli. This was the case for all three of the actions presented in this experiment. Overall, the results do not support the idea that shape-motion inconsistency results in negative affect toward the animations.

Interestingly, judgments of animation characteristics depended on the viewed action. The object throws were perceived as more eerie and more artificial than either lifting or pushing. This might be due to a limitation of the animation technique, which did not include recordings of the motion of individual fingers. We assume that this becomes more obvious in the throwing stimuli than in the other two for two reasons. First, we assume that attention rests more on the hand for the throwing movement as compared to the two others. Deriving object weight in sled pushing and box lifting requires attention to the whole pose of the body, while throwing is mainly a matter of arm and hand motion. Second, the ball is a much smaller object than the box or the sled. It is perfectly plausible that the box was lifted with stretched out fingers, but since observers were told that the thrown object was a small beanbag, they would assume that the hand should be in a grasping pose.

Finally, in contrast to some other studies that had reported effects of inconsistencies between shape and motion [23, 49], the actions that were presented in the current study were all transitive actions; that is, all of them involved the manipulation of an object. That allows for additional degrees of freedom in the way the situation can be interpreted. Since the object itself was not visible, its properties could only be inferred from the way the actor was interacting with it. Changes in the relations between body shape and body motion may be interpreted by the observer in terms of changes of object properties. If that was the case, then we would not expect participants to perceive the hybrid stimuli as inconsistent or to attribute them with negative affect, but we may observe effects on how the objects and the interaction with them is perceived. We address this possibility in Experiment 3.

5 EXPERIMENT 3: PERCEPTION OF ACTION OUTCOMES

Past research into human action perception has often employed point-light displays [18] rather than animated characters. Point-light displays convey only limited information about body shape and mainly provide body kinematics. Kinematic information plays a critical role in inferring action outcomes, but, depending on the action, body shape may influence how this information is interpreted by the visual system.

Information from movement kinematics alone is often sufficient to support the perception of action outcomes to some degree. For example, the throwing distance of invisible objects can be inferred from the arm kinematics of point-light displays [31, 36]. In other actions, body shape is more likely to influence how kinematic information is interpreted by the visual system. For instance, although kinematics support the perception of a lifted weight [2, 8, 39], the weight of the performer may modulate how kinematics are used or interpreted. For instance, light performers who lift increasingly heavy boxes adjust their posture to stabilize themselves as a counterweight to the box, while heavier performers with a more stable base of support do not require postural adjustments [35]. In addition, while there are no studies that have examined the perception of the weight of a pushed sled, recent evidence [16] suggests that there might also be a strong interaction between shape and motion information for this type of action: Heavy performers can push a training sled at greater peak velocities than light performers, and this is likely to result from an ability to produce greater forces and to do so more quickly than lighter performers. Together this evidence suggests that there may be an action-specificity in the effect of shape-motion inconsistency on perception of action outcomes.

In Experiment 3, we used the full set of stimuli as described in Section 2.1 and asked participants to assess the outcome of the observed actions. They were instructed to estimate the weight of the sleds pushed by the actor, to estimate the distance of the thrown objects, and to estimate the weight of the lifted boxes.

5.1 Participants

We recruited 63 new volunteers from Queen's University, 42 women and 21 men. Twenty-one participants were assigned to each action viewing condition: throwing, pushing, or lifting. Participants were between 17 and 34 years of age ($M = 21.4$ years old, $SD = 3.97$). None had taken part in the motion capture sessions or in Experiments 1 or 2. All reported normal or corrected-to-normal vision.

Fig. 8. Following the presentation of each stimulus, participants indicated the perceived action outcomes using a continuous response scale. Consistent and inconsistent stimuli were randomly presented. Actions were varied as between-subjects factors.

5.2 Methods

5.2.1 Stimuli. As in Experiments 1 and 2, stimuli were five light and five heavy male performers who had been recorded while pushing, lifting, and throwing objects. However, contrary to the previous experiments, where performers were only shown manipulating one object, each performer was shown with multiple objects weights and throwing distances (see Section 2). Performers pushed six weighted sleds (23–123kg), threw beanbags toward five targets (2–6m), and lifted boxes with six different weights (1–11kg).

5.2.2 Procedure. To familiarize participants with the scenarios used in the experiment, they were shown typical examples of the items that were used during the motion capture session: a 6kg box that was used for the lifting, a sled loaded with 73kg, and a beanbag that was used for the throwing.

Participants viewed either the pushing, throwing, or lifting stimulus sets. Stimuli were presented one at a time in a randomized order. After presentation of each stimulus, participants were prompted to input the perceived action outcomes on a continuous rating scale (Figure 8). Weights were described and reported in pounds (1lb = 0.454kg), because participants in pilot studies indicated a greater familiarity with the imperial units. The scales ranged from 0lbs to 24lbs for the lifted weight, from 0lbs to 320lbs for the weight of the pushed sled, and from 1m to 7m for the distance of the thrown beanbag. No feedback was given.

5.3 Results

We conducted independent 2 (shape) $\times$ 2 (motion) repeated-measures ANOVAs on perceived action outcomes for each of the different actions. The levels of the factors indicated whether that aspect of the stimulus was based on a heavy or a light performer. The effects of shape and motion were action specific and are reported in the subsections below (Figure 9).

5.3.1 Pushing. There was a significant effect of shape, $F(1, 20) = 43.4, p < 0.001, \eta_p^2 = 0.684$. Characters with the shape of a heavy performer were perceived as pushing heavier sleds ($M = 76.2\text{kg}, SEM = 2.83\text{kg}$) than characters with light performer shapes ($M = 66.2\text{kg}, SEM = 2.40\text{kg}$). There was also a significant effect of motion, $F(1, 20) = 103, p < 0.001, \eta_p^2 = 0.838$. The sleds pushed by characters animated with the motion of a heavy performer were perceived as lighter ($M = 64.4\text{kg}, SEM = 2.52\text{kg}$) than sleds pushed by characters animated with light performer motions ($M = 78.0\text{kg}, SEM = 2.68\text{kg}$). Finally, there was also a significant interaction between shape and motion, $F(1, 20) = 18.0, p < 0.001, \eta_p^2 = 0.486$. When observers judged the action outcomes of characters whose shape and motion were consistent, observers perceived similar sled weights, irrespective of whether the shape and motion came from a light or a heavy performer, paired t -test, $t(20) = -1.99, p = 0.061$. However, sled

Fig. 9. Regression lines representing the relationship between perceived and real weights and distances. The shaded areas represent ± 1 SEM.

weights were perceived as lighter if the characters were created from a combination of light performer shapes and heavy performer motions ($M = 60.3\text{kg}$, $SEM = 2.46\text{kg}$), and sled weights were perceived as heavier if the characters were created from heavy performer shapes and light performer motions ($M = 83.9\text{kg}$, $SEM = 3.10\text{kg}$). Note that ratings for the consistent performers are very similar, no matter whether they were light or heavy. Observers apparently take body shape of performers into account when interpreting their poses and kinematics.

The slopes of the regression lines were influenced by both shape and motion. A 2 (shape) $\times 2$ (motion) repeated-measures ANOVA on the slope values revealed a significant effect of shape, $F(1, 20) = 10.1$, $p = 0.005$, $\eta_p^2 = 0.335$, with heavy body shapes resulting in steeper slopes ($M = 0.374$, $SEM = 0.030$) than those for light body shapes ($M = 0.324$, $SEM = 0.324$). There was also a significant effect of motion, $F(1, 20) = 10.2$, $p = 0.005$, $\eta_p^2 = 0.338$, with the light body motion resulting in steeper slopes ($M = 0.384$, $SEM = 0.035$) than those for heavy body motion ($M = 0.314$, $SEM = 0.029$). There was no interaction between shape and motion, $F(1, 20) = 0.335$, $p = 0.569$, $\eta_p^2 = 0.016$.

5.3.2 Throwing. The pattern of results observed is different in the object throwing condition. Here, the differences in perceived thrown distances mostly depend on the source of the motion information and, to a much lesser extent, on the source of the shape information. There was a strong effect of motion, $F(1, 20) = 2.20$, $p < 0.001$, $\eta_p^2 = 0.811$. Characters that were animated with the motion of a heavy performer were perceived as throwing farther ($M = 4.35\text{m}$, $SEM = 0.115\text{m}$) than characters animated with motion from light performers ($M = 3.59\text{m}$, $SEM = 0.140\text{m}$). There was also a small, yet significant, effect of shape, $F(1, 20) = 9.97$, $p = 0.005$, $\eta_p^2 = 0.333$. Characters with the shape of a heavy performer were perceived as throwing slightly farther ($M = 4.03\text{m}$, $SEM = 0.127\text{m}$) than characters based on light performer shapes ($M = 3.91\text{m}$, $SEM = 0.112\text{m}$). There was no interaction between shape and motion, $F(1, 20) = 0.847$, $p = 0.368$, $\eta_p^2 = 0.041$.

The slopes of the regression lines were not influenced by shape or motion. A 2 (shape) $\times 2$ (motion) repeated-measures ANOVA on the slope values revealed that there was no significant effect of shape, $F(1, 20) = 2.59$, $p = 0.123$, $\eta_p^2 = 0.115$, with heavy body shapes resulting in similar slopes ($M = 0.437$, $SEM = 0.029$) to those for light body shapes ($M = 0.407$, $SEM = 0.030$). There was also no significant effect of motion, $F(1, 20) = 0.116$, $p = 0.775$, $\eta_p^2 = 0.006$, with the heavy body motion resulting in similar slopes ($M = 0.426$, $SEM = 0.031$) to those for light body motion ($M = 0.418$, $SEM = 0.030$). The interaction between shape and motion was not significant, $F(1, 20) = 3.39$, $p = 0.080$, $\eta_p^2 = 0.145$.

5.3.3 Lifting. The box lift pattern is more complex. There was a significant effect of shape, $p < 0.001$, $\eta_p^2 = 0.582$. Characters with heavy performer shapes were perceived to lift heavier boxes ($M = 5.90\text{kg}$, $SEM = 0.780\text{kg}$) than characters with light performer shapes ($M = 5.44\text{kg}$, $SEM = 0.168\text{kg}$). There was no significant effect of motion, $F(1, 20) = 0.290$, $p = 0.596$, $\eta_p^2 = 0.014$. There was also no interaction between shape and motion, $F(1, 20) = 0.570$, $p = 0.459$, $\eta_p^2 = 0.028$. All effects on perceived box weight seemed to be carried by the shape of the performer.

However, further examination of the results provided evidence that observers had difficulty estimating box weights, as evidenced by the shallow slopes of the regression lines. There is some evidence that motion from light performers helps observers perform the task. A 2 (shape) $\times$ 2 (motion) repeated-measures ANOVA on the slope values revealed that there was no significant effect of shape, $F(1, 20) = 0.181$, $p = 0.675$, $\eta_p^2 = 0.009$. However, there was a significant effect of motion, $F(1, 20) = 8.46$, $p = 0.009$, $\eta_p^2 = 0.297$. Stimuli created with light performer motions resulted in slightly steeper slopes ($M = 0.073$, $SEM = 0.018$) than stimuli created with heavy performer motions ($M = -0.006$, $SEM = 0.015$). The interaction between shape and motion was not significant, $F(1, 20) = 3.31$, $p = 0.084$, $\eta_p^2 = 0.142$.

5.4 Discussion

The results of this experiment demonstrate that the changes introduced by hybridizing the shape of one performer with the motion of another were consequential: Observers clearly respond to these manipulations. However, contrary to our initial expectations, the changes are not interpreted as inconsistencies between shape and motion or as abnormalities in the way a person moves. The changes are absorbed by the interpretation of the object or the way the actor interacts with it. The additional degrees of freedom, which the unknown specifics of the object offers, provide enough room for a new, plausible, and credible interpretation. The weight of the object (in the case of the sled and the box) and the distance thrown (in the case of the beanbag) are absorbing the manipulation introduced by hybridizing shape and motion from different performers. This explains why questions asked about the animated CG characters, rather than the objects they manipulated, had not revealed any indication that participants were able to identify and evaluate the experimental manipulations.

We observed different patterns of shape and motion contributions to the perceived action outcomes. When observers reported their perception of the weight of a pushed training sled, the results depended strongly on the internal consistency of the stimuli. Replacing the veridical kinematics of a heavy actor with the kinematics of a light actor increases perceived weight of the sled. For a lighter-weight person, moving the sled requires more effort that is probably expressed in the posture and motion of the pushing character. The same logic explains why replacing the veridical body shape of the same heavy actor with a lighter-weight body shape decreases the perceived weight of the sled. When replacing both shape and motion of a heavy actor with the ones of a light actor, thus generating a coherent animation again, the two effects cancel each other. Consequently, when observers judged stimuli whose shape and motion were consistent, perceived weight was independent of the weight of the actor.

This interpretation is consistent with work that suggest that the body weight of an individual strongly modulates the kinematics of the pushing motion [16], as well as with the Kinematic Specification of Dynamics principle that suggests that the motion of light individuals is more strongly impacted by changes in object weight than the motion of heavy individuals [35, 36].

In the case of beanbag throws, we observe a very different pattern. The perceived distance of the throw seems to mainly depend on the kinematics with only little contribution from body shape. The bean bag was relatively light (about 400g) and the throwing distance can probably be controlled by the kinematics of the arm alone. The postural changes responsible for differences between sled pushes of light and heavy actors do not have an equivalent in the throwing movements. This result is consistent with previous literature using point-light displays. Munzert and colleagues [31] have shown that distance estimation from arm kinematics can be achieved in the absence of an explicit representation of the rest of the body.

The results of the box lift condition are harder to interpret. Participants appeared unable to respond accurately on the perception task. The slight effect of body shape on slopes is consistent with the idea that it is easier to perceive differences in box weights with light body shapes than with heavy body shapes. Observers could not discriminate the weights of the boxes when the motion was taken from a heavy performer, regardless of the consistency of the animation, as evidenced by a regression coefficient that did not differ from zero. In a way, this is consistent with the Runeson and Frykholm's [35] hypothesis that heavy individuals do not need to significantly

adjust their posture when lifting weights because of the high stability of their center of moment during the box lift, whereas light performers must adjust their posture more for larger weights to account for the destabilization incurred by the box lift. Such a hypothesis predicts a better ability to perceive the box weights for light performers. However, the slope values for these performers were too small to draw reasonable conclusions. The above interpretation therefore remains speculative and calls for further experiments with heavier box weights that contribute more strongly to postural readjustments.

Note that the action outcome perception task was relatively difficult, as evidenced by the relatively shallow slopes obtained for the three actions. Some earlier studies that used point-light displays or video recordings found more accurate ratings [31, 35]. The task of the current study, although ecologically valid, may be more difficult than other methods used in previous box lifting studies.

In this version of MoSh, hands were depicted with open palms and static, extended fingers. We can speculate that observers interpreted these hand configurations as inefficient and that it would only be possible for them to throw objects at shorter distances and lift lighter objects. The usage of CG characters in the current study as compared to point-light displays that have been employed in earlier ones might also affect search patterns and the guidance of attention. Even in simple tasks such as gender identification from walking stimuli, eye gaze fixation patterns can differ significantly depending on the rendering of the stimulus: Fixations on point-light displays are directed mostly toward the hips and feet [37], while fixations on animated CG characters are mostly on the upper body and faces [7]. The explicit surface of the CG characters might influence which regions of the body the observer attends to, incidentally reducing the ability to extract information from other regions of the visual display that may contain the most relevant information, such as the hands.

6 GENERAL DISCUSSION AND CONCLUSIONS

We demonstrated that observers were barely able to explicitly detect shape-motion inconsistency (Experiment 1) and that judgments of animations characteristics were not affected by inconsistency either (Experiment 2). However, inconsistency in shape and motion had a systematic, action-specific effect on the perception of action outcomes (Experiment 3).

The visual system has undeniably access to vast knowledge regarding the role of shape and motion information in action production. Past research has shown that knowledge about relationships between shape and motion is used for a number of different person perception tasks [2, 3, 8, 26, 31, 35, 36, 39, 40, 44, 45, 50]. Most of this previous work, however, was conducted using point-light displays that provide only very limited information about body shape and leave a large part of the person's appearance to the imagination of the observer.

In the current study, we used realistic character renderings, which combine a person's individual body kinematics with detailed individual body shape. These displays therefore provide the observer with much more information and realism about a person's individual appearance. Our experimental results suggest that when the performers that contribute to a hybrid have very different overall body mass, hybridization affects the perceived dynamics of the actions, but are neither identified as inconsistent animations nor are they attributed with negative connotations such as eeriness or reduced attractiveness.

Our results seem to contrast previous work that used point-light displays of walking people and showed that inconsistency between the geometry of the body and its kinematics reduced attractiveness and realism of walker displays [23, 49]. However, unlike the movements that we used in the current study, walking is not a transitive action and does not involve manipulation of an object. In our study, the objects and their perceived properties provide additional degrees of freedom that allow the observer to come up with a new valid and consistent reinterpretation of the observed event. It is the objects and their properties that absorb the changes introduced to the stimulus by combining shape and motion from different performers.

If that interpretation is valid, then our results would be specific to transitive actions. Locomotion actions such as walking, running, jumping, and many other activities that do not involve the direct manipulation of objects

may be more likely to reveal to the observer inconsistencies that result from retargeting. Similar reasoning has been invoked by researchers who manipulated motion patterns during throwing actions [53]. The authors demonstrated that these movements can be deformed to a very large extent before the visual system starts to interpret them as unnatural. The visual system seeks valid interpretations of the world whenever possible. That may be easier for transitive actions, which offer more opportunities to discount for inconsistencies introduced by the retargeting process.

It seems that the human visual system, while knowledgeable about the natural relations between kinematics and body shape in human motion, is willing to accept reinterpretations of retargeted animations, if at all possible. However, the animator has to be aware of potential changes in the perception of manipulated objects, which may have unintended effects on the narrative of an event. In our experiments, the manipulated objects were invisible, which provided additional degrees of freedom for the imagination of the observer to adjust object properties to augment the available visual data and generate a plausible event. Even if objects were visible, some of their properties (such as the weight) may not be obvious. However, in other situations, even renderings of transitive actions may constrain object properties to a point at which they do not provide enough freedom for a consistent re-interpretation that avoids a sense of inconsistency.

In conclusion, the results have interesting implications for retargeting in computer animation. It seems that the visual system is relatively tolerant to inconsistencies, so long as the shape and motion of the animation are realistic. When the performer and avatar are matched for height, but have very different body weight, this inconsistency is not likely to introduce changes in animation judgements or to be noticeable. In this sense, researchers who carefully match the body shape of performers, recorded while walking, with the body shape of their virtual avatars, are unlikely to have introduced disruptive inconsistencies [17, 28]. However, if conveying the properties of manipulated objects to an observer is critical, the role of body shape should be carefully considered. This could even be used for artistic effect in animation retargeting: If the pushing motion of a light person was used to animate a heavy and strong-looking body shape, then the apparent weight of a pushed object would appear much heavier than it was in reality, yet still appear natural to an observer.

ACKNOWLEDGMENTS

The authors thank the anonymous referees for their valuable comments and helpful suggestions.

REFERENCES

- [1] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. 2005. SCAPE: Shape completion and animation of people. In *ACM Transactions on Graphics*, Vol. 24. 408–416.
- [2] Geoffrey P. Bingham. 1987. Kinematic form and scaling: Further investigations and the visual perception of lifted weight. *J. Exp. Psychol.: Hum. Percept. Perf.* 13, 2 (May 1987), 155–177.
- [3] Geoffrey P. Bingham. 1993. Scaling judgments of lifted weight: Lifter size and the role of the standard. *Ecol. Psychol.* 5, 1 (Jan. 1993), 31–64.
- [4] Harry Brenton, Marco Gillies, Daniel Ballin, and David J. Chatting. 2005. The uncanny valley: Does it exist? In *Proceedings of the 19th British HCI Group Annual Conference*. 1–4.
- [5] Claudia E. Cohen. 1981. Person categories and social perception: Testing some boundaries of the processing effect of prior knowledge. *J. Pers. Soc. Psychol.* 40, 3 (Jan. 1981), 441.
- [6] James D. Douketis, Gilles Paradis, Heather Keller, and Chantal Martineau. 2005. Canadian guidelines for body weight classification in adults: Application in clinical practice to screen for overweight and obesity and to assess disease risk. *Can. Med. Assoc. J.* 172, 8 (2005), 995–998.
- [7] Cathy Ennis, Ludovic Hoyet, and Carol O’Sullivan. 2015. Eye-tracktive: Measuring attention to body parts when judging human motions. In *Eurographics*. 47–40.
- [8] Antonia F. de C. Hamilton, Daniel W. Joyce, J. Randall Flanagan, Chris D. Frith, and Daniel M. Wolpert. 2007. Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychol. Res.* 71, 1 (Jan. 2007), 13–21.
- [9] Weidong Geng and Gino Yu. 2003. Reuse of motion capture data in animation: A review. In *Proceedings of the International Conference on Computational Science and Its Application (ICCSA’03)*. Lecture Notes in Computer Science, Vol. 2669. 620–629.
- [10] James J. Gibson. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

- [11] Michael Gleicher. 1998. Retargetting motion to new characters. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. 33–42.
- [12] Julie Grèzes, Chris. D. Frith, and Richard E. Passingham. 2004. Brain mechanisms for inferring deceit in the actions of others. *J. Neurosci.* 24, 24 (2004), 5500–5505.
- [13] David A. Hirshberg, Matthew Loper, Eric Rachlin, and Michael J. Black. 2012. Coregistration: Simultaneous alignment and modeling of articulated 3D shape. In *Proceedings of the European Conference on Computer Vision*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid (Eds.). 242–255.
- [14] Chin-Chang Ho and Karl MacDorman. 2010. Revisiting the uncanny valley hypothesis: Developing and validating an alternative to the Godspeed indices. *Comput. Hum. Behav.* 26 (2010), 1508–1518.
- [15] Jessica Hodgins, Sophie Jörg, Carol O’Sullivan, Sang Il Park, and Moshe Mahler. 2010. The saliency of anomalies in animated human characters. *ACM Trans. Appl. Percept.* 7, 4 (2010), 1–11.
- [16] James Jr. Hoffman. 2014. *An Investigation of the Sled Push Exercise: Quantification of Work, Kinematics, and Related Physical Characteristics*. Ph.D. Dissertation. Sport Physiology and Performance, East Tennessee State University.
- [17] Ludovic Hoyet, Kenneth Ryall, Katja Zibrek, Hwangpil Park, Jehee Lee, Jessica Hodgins, and Carol O’Sullivan. 2013. Evaluating the distinctiveness and attractiveness of human motions on realistic virtual bodies. *ACM Trans. Graph.* 32, 6 (2013), 1–11.
- [18] Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 2 (Jun. 1973), 201–211.
- [19] Kerri L. Johnson and Louis G. Tassinary. 2007. Compatibility of basic social perception determines perceived attractiveness. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12 (Ma. 2007), 5246–5251.
- [20] Mark H. Johnson. 2006. Biological motion: A perceptual life detector? *Curr. Biol.* 16, 10 (May 2006), R376–R377.
- [21] Daniel Jokisch and Nikolaus F. Troje. 2003. Biological motion as a cue for the perception of size. *J. Vis.* 3, 4 (May 2003), 252–264.
- [22] Jari Kätsyri, Klaus Förger, Meeri Mäkräinen, and Tapio Takala. 2015. A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Front. Psychol.* 6 (2015), 1–16.
- [23] Malte Klüver, Heiko Hecht, and Nikolaus F. Troje. 2016. Internal consistency predicts attractiveness in biological motion walkers. *Evol. Hum. Behav.* 37, 1 (Jan. 2016), 40–46.
- [24] Matthew Loper, Naureen Mahmood, and Michael J. Black. 2014. MoSh: Motion and shape capture from sparse markers. *ACM Trans. Graph.* 33, 6 (2014), 1–10.
- [25] Karl F. MacDorman and Hiroshi Ishiguro. 2006. The uncanny advantage of using androids in cognitive and social science research. *Interact. Stud.* 7, 3 (2006), 297–337.
- [26] George Mather and Linda Murdoch. 1994. Gender discrimination in biological motion displays based on dynamic cues. *Proc. Roy. Soc. Lond. B* 258, 1353 (Dec. 1994), 273–279.
- [27] Daisuke Matsui, Takashi Minato, Karl F. MacDorman, and Hiroshi Ishiguro. 2005. Generating natural motion in an android by mapping human motion. *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 1–8.
- [28] Rachel McDonnell, Sophie Jörg, Jessica K. Hodgins, Fiona Newell, and Carol O’Sullivan. 2007. Virtual shapers and movers: Form and motion affect sex perception. In *Proceedings of the Symposium on Applied Perception in Graphics and Visualization (APGV’07)*. 7–11.
- [29] Rachel McDonnell, Michaël Larkin, Simon Dobbyn, Steven Collins, and Carol O’Sullivan. 2008. Clone attack! Perception of crowd variety. In *ACM Transactions on Graphics*, Vol. 27. 1–8.
- [30] Alberto Menache. 2011. *Understanding Motion Capture for Computer Animation and Video Games* (2nd ed.). Morgan Kaufmann.
- [31] Jörn Munzert, Tanja Hohmann, and Ernst-Joachim Hossner. 2010. Discriminating throwing distances from point-light displays with masked ball flight. *Eur. J. Cogn. Psychol.* 22, 2 (Jan. 2010), 247–264.
- [32] Lukasz Piwek, McKay Lawrie S., and Frank E. Pollick. 2014. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition* 130, 3 (2014), 271–277.
- [33] Clément Raïevsky and François Michaud. 2009. Emotion generation based on a mismatch theory of emotions for situated agents. In *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*. SciTech Book News, 247–266.
- [34] Sverker Runeson. 1977. On the possibility of ‘smart’ perceptual mechanisms. *Scand. J. Psychol.* 18, 1 (1977), 172–179.
- [35] Sverker Runeson and Gunilla Frykholm. 1981. Visual perception of lifted weights. *J. Exp. Psychol.: Hum. Percept. Perf.* 7, 4 (Aug. 1981), 733–740.
- [36] Sverker Runeson and Gunilla Frykholm. 1983. Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *J. Exp. Psychol.: General* 112, 4 (Dec. 1983), 585–615.
- [37] Daniel R. Saunders, David K. Williamson, and Nikolaus F. Troje. 2010. Gaze patterns during perception of direction and gender from biological motion. *J. Vis.* 10, 11 (Sep. 2010), 1–10.
- [38] Ayse Pinar Saygin, Thierry Chaminade, Hiroshi Ishiguro, Jon Driver, and Chris Frith. 2012. The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Soc. Cogn. Affect. Neurosci.* 7, 4 (2012), 413–422.
- [39] Jaeho Shim and Les G. Carlton. 1997. Perception of kinematic characteristics in the motion of lifted weight. *J. Motor Behav.* 29, 2 (Mar. 1997), 131–146.

- [40] Jaeho Shim, Les G. Carlton, and Jitae Kim. 2004. Estimation of lifted weight and produced effort through perception of point-light display. *Perception* 33, 3 (Mar. 2004), 277–291.
- [41] James C. Thompson, J. Gregory Trafton, and Patrick McKnight. 2011. The perception of humanness from the movements of synthetic agents. *Perception* 40, 6 (2011), 695–704.
- [42] Angela Tinwell. 2014. *The Uncanny Valley in Games and Animation*. CRC Press.
- [43] Angela Tinwell, Deborah Abdel Nabi, and John P. Charlton. 2013. Perception of psychopathy and the Uncanny Valley in virtual characters. *Comput. Hum. Behav.* 29, 4 (2013), 1617–1625.
- [44] Nikolaus F. Troje. 2002. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.* 2, 5 (2002), 371–387.
- [45] Nikolaus F. Troje. 2002. The little difference: Fourier based gender classification from biological motion. In *Dynamic Perception*, R. P. Würtlitz and M. Lappe (Eds.). AKA Press, Berlin, 115–120.
- [46] Nikolaus F. Troje. 2008. Retrieving information from human movement patterns. In *Understanding Events: How Humans See, Represent, and Act on Events*, T. F. Shipley and J. M. Zacks (Eds.). Oxford University Press, 308–334.
- [47] Nikolaus F. Troje. 2013. What is biological motion? Definition, stimuli and paradigms. In *Social Perception: Detection and Interpretation of Animacy, Agency and Intention*, M. D. Rutherford and V. A. Kuhlmeier (Eds.). MIT Press, 13–36.
- [48] Nikolaus F. Troje and Dorita H. F. Chang. 2013. Shape-independent processes in biological motion perception. In *People Watching: Social, Perceptual, and Neurophysiological Studies of Body Perception*, K. L. Johnson and M. Shiffrar (Eds.). Oxford University Press, Oxford, 82–100.
- [49] Nikolaus F. Troje and Stephanie Lau. 2012. Perceived naturalness of human motion depends on internal consistency. *J. Vis.* 12, 9 (2012), 466.
- [50] Nikolaus F. Troje, Cord Westhoff, and Mikhail Lavrov. 2005. Person identification from biological motion: Effects of structural and kinematic cues. *Percept. Psychophys.* 67, 4 (Febr. 2005), 667–675.
- [51] Richard Wages, Stefan M. Grünvogel, and Benno Grützmacher. 2004. How realistic is realism? Considerations on the aesthetics of computer games. *Lecture Notes in Computer Science* 3166 (2004), 216–225.
- [52] Cord Westhoff and Nikolaus F. Troje. 2007. Kinematic cues for person identification from biological motion. *Percept. Psychophys.* 69, 2 (Febr. 2007), 241–253.
- [53] Qin Zhu and Geoffrey P. Bingham. 2014. Seeing where the stone is thrown by observing a point-light thrower: Perceiving the effect of action is enabled by information, not motor experience. *Ecol. Psychol.* 26 (Oct. 2014), 229–261.

Received February 2018; revised September 2018; accepted September 2018