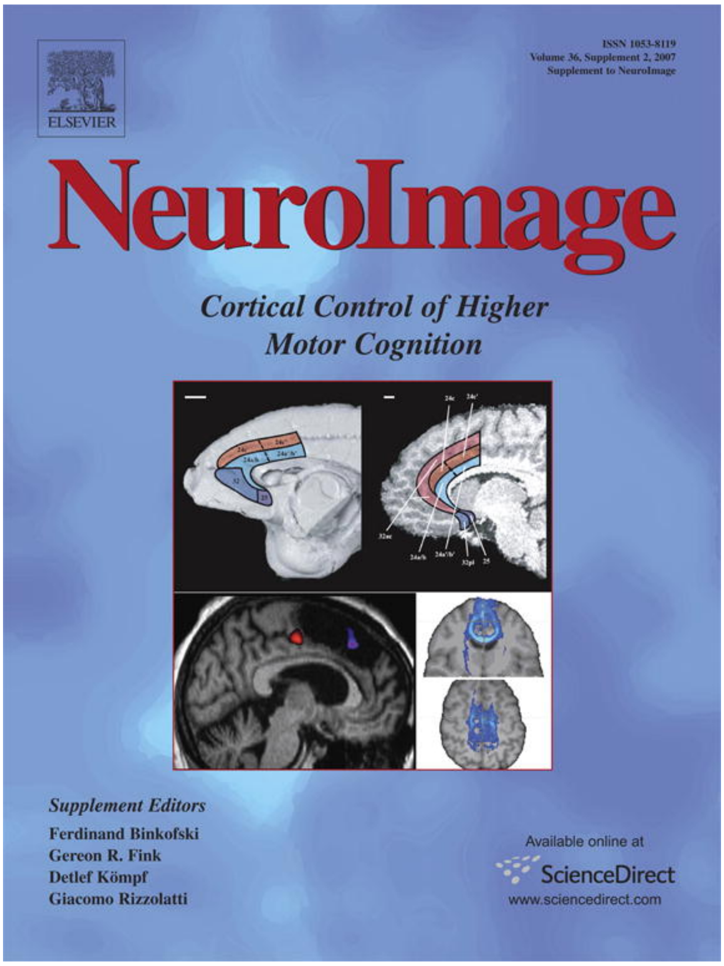

ISSN 1053-8119
Volume 36, Supplement 2, 2007
Supplement to NeuroImage

ELSEVIER

NeuroImage

*Cortical Control of Higher
Motor Cognition*

Supplement Editors
Ferdinand Binkofski
Gereon R. Fink
Detlef Kömpf
Giacomo Rizzolatti

Available online at
ScienceDirect
www.sciencedirect.com

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 36 (2007) T22 – T32

Predicting point-light actions in real-time

Markus Graf,[a,*] Bianca Reitzner,[a,b] Caroline Corves,[a,c] Antonino Casile,[d]
Martin Giese,[d] and Wolfgang Prinz[a]

[a] *Max-Planck-Institute for Human Cognitive and Brain Sciences, Department of Psychology, Amalienstraβe 33, D-80799 Munich, Germany*
[b] *University of Eichstätt, Department of Psychology, Eichstätt, Germany*
[c] *University of Munich, Department of Psychology, Munich, Germany*
[d] *Laboratory for Action Representation and Learning, Dept. of Cognitive Neurology, Hertie Institute for Clinical Brain Research, University Clinic Tübingen, Germany*

Received 11 August 2006; accepted 20 March 2007
Available online 27 March 2007

There is convincing evidence for a mirror system in humans which simulates actions of conspecifics. One possible purpose of such a simulation system is to support action prediction in real-time. Our goal was to study whether the prediction of actions involves a real-time simulation process. We motion-captured a number of human actions and rendered them as point-light action sequences. Observers perceived brief videos of these actions, followed by an occluder and a static test posture. We independently varied the occluder time and the movement gap (i.e., the time between the endpoint of the action and the test posture). Observers were required to judge whether the test stimulus depicted a continuation of the action in the same depth orientation. Prediction performance was best when occluder time and movement gap corresponded, i.e., when the test posture was a continuation of the sequence that matched the occluder duration (Experiments 1, 2 and 4). This pattern of results was destroyed when the sequences and test images were flipped around the horizontal axis (Experiment 3). Overall, our findings suggest that action prediction involves a simulation process that operates in real-time. This process can break down when the actions are presented under viewing conditions for which observers have little experience.

Keywords: Forward model; Real-time prediction; Point-light action; Simulation

Introduction

Visual perception is not simply a post-hoc reconstruction of the visual input, but is a predictive activity. Perceptual anticipation has been demonstrated in a number of different fields. For instance, it has been argued that the flash lag effect involves the extrapolation of predictable movements (e.g., de Valois and de Valois, 1991; Khurana and Nijhawan, 1995; Nijhawan, 1994, 1997; but see Eagleman and Sejnowski, 2000). In representational momentum, movement or implied movement of an object results in perceptual extrapolation (or prediction) of a trajectory beyond the actual movement presented on the screen (e.g., Freyd and Finke, 1984; Finke and Freyd, 1985; Bertamini, 1993). In a typical experiment, an object is shown in three successive orientations, implying rotation in a particular direction. Observers are probed immediately after with a fourth presentation of the object, and asked whether the probe is in the same orientation as in the last presentation. It is much harder for observers to reject a probe rotated further beyond the last presentation than a probe rotated backwards, which suggests that perceptual representations are extrapolated into the immediate future. It seems that perceptual representation is based on anticipation, and is not just some representational analogue of physical momentum (Verfaillie and d'Ydewalle, 1991). Neuroimaging studies support the idea that some kind of perceptual extrapolation is performed. Brain areas that process real motion, such as area MT, are activated when stimuli are presented that imply motion (like the picture of a starting rocket), or moving stimuli that become occluded (Kourtzi and Kanwisher, 2000; Olson et al., 2003; Senior et al., 2000). Disruption of these brain areas by transcranial magnetic stimulation eliminates representational momentum (Senior et al., 2002).

Predictive activity has been found not only for object motion, but also for more complex movement patterns, like for human motion stimuli. Representational momentum has been demonstrated for point-light action stimuli (Jarraya et al., 2005), and for human actions in complex scenes (Thornton and Hayes, 2004). Evidence that observers automatically anticipate the future posture of actors has been provided in a priming study (Verfaillie and Daems, 2002). Brief sequences of human actions were shown in an initial priming phase. In a later test phase, reaction times (RTs) were faster for test postures that would succeed a primed sequence, as compared to (a) test postures preceding a primed sequence, and (b) postures unrelated to previous action sequences (for related

---

[*] Corresponding author. Fax: +49 341 9940 030.
E-mail addresses: markus.graf@cbs.mpg.de, markus.graf@gmail.com (M. Graf).
URL: http://www.cbs.mpg.de/~graf (M. Graf).

1053-8119/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2007.03.017

findings see Verfaillie et al., 1994; Kourtzi and Shiffrar, 1999). Accordingly, neurons tuned to specific actions in monkey superior temporal sulcus (STS) also respond to articulated static end postures of these actions, but not to non-articulated starting postures (Jellema and Perrett, 2003a). Thus, STS codes biological motion also when implied from static postures. Moreover, predictive coding of the behavior of others was reported (Jellema and Perrett, 2003b). Neurons in STS responded vigorously to the sight of a posture that followed a particular body action, but not when it followed other actions. In contrast, the effective action or posture presented in isolation or in different sequences failed to produce a response. Thus, STS neurons seem to encode specific action-posture sequences.

All these findings suggest that perception is predictive. However, not much is known about the timing of these predictive aspects. In this paper we argue that more detailed hypotheses about timing can be derived from research on visuomotor control. It has been proposed that the motor system can be used to emulate observed action, helping to project the future course of an event (Grush, 2004). The motor system could be providing vision with a general framework within which to represent and process time, or change-over-time (Thornton and Knoblich, 2006; Wilson and Knoblich, 2005).

In accordance with this proposal, perceptual and motor systems seem to be tightly coupled. There is plenty of evidence that perception and action planning are coded in a common representational medium (e.g., Prinz, 1990, 1997; Hommel et al., 2001). Moreover, observing, imagining or representing an action excites the motor programs used to execute that same action. Evidence for this kind of covert imitation by the motor system comes from different literatures. Behavioral experiments demonstrated that people have an unconscious tendency to mimic others' behavior. This is known as the chameleon effect (e.g., Chartrand and Bargh, 1999), or, in the context of emotion-related behaviors, emotional contagion (e.g., Hatfield et al., 1994). This kind of imitation seems to be a fast and automatic process (see Prinz, 2002, 2005, for reviews). Further evidence for covert imitation has been provided in neurophysiology. The most prominent examples are mirror neurons, found in monkey premotor and parietal cortices. These neurons respond when the monkey performs an object-directed action, and when the monkey perceives the same action performed by someone else (e.g., di Pellegrino et al., 1992; Gallese et al., 2002; Umiltà et al., 2001). Thus, mirror neurons code the execution of actions, and the perception of the same actions — even when the monkey does not execute overt behavior. Neuroimaging studies indicate that there is also a mirror system in humans. Motor planning areas are activated when humans simply observe actions of others (e.g. Buccino et al., 2001, 2004). In humans, motor activations occur also when complex, skilled whole-body movements are observed. The activation is greater for subjects who are experts in performing that type of movement (Calvo-Merino et al., 2005). In accordance, patients with frontal lobe lesions (which reduce the ability for inhibition) may sometimes show an automatic and uncontrollable tendency to imitate (De Renzi et al., 1996).

Why does the brain covertly imitate actions? Simulation theory suggests that we perceive or imagine the actions of conspecifics by internally simulating them with a network including the motor system (Jeannerod, 2001; Gallese, 2005). More specifically, observed actions are simulated with one's own motor repertoire (Casile and Giese, 2006; Calvo-Merino et al., 2005). Several possible *functions* of such simulation have been proposed, like the facilitation of overt imitation (Iacoboni et al., 1999), the understanding of actions in motor terms (Rizzolatti et al., 2001), and the understanding of others' intentions (Blakemore and Decety, 2001; Fogassi et al., 2005; Iacoboni et al., 2005). Furthermore, the mirror system may subserve working memory (Wilson, 2001), and even play a role in the evolutionary development of language (Arbib, 2005). Most important for the present study, it has been proposed that an important function of the mirror system lies in the *prediction* of others' actions (Blakemore and Frith, 2005; Prinz, 2006; Wilson and Knoblich, 2005; Wolpert and Flanagan, 2001).

Several studies point to a predictive function of the motor system in action perception. First, motor activity is found prior to observing someone else's actions when the action can be predicted (Kilner et al., 2004; Haueisen and Knösche, 2001). Second, when subjects observe an actor performing a task, the coordination between their gaze and the actor's hand is predictive, rather than reactive (Flanagan and Johansson, 2003). Third, a subset of mirror neurons respond even when the final part of an action is hidden by an occluder and therefore can only be inferred (Umiltà et al., 2001). Moreover, premotor areas are activated in serial prediction tasks, even for purely perceptual stimuli that cannot be mapped onto the human body (Schubotz and von Cramon, 2002, 2004). Thus, (pre-) motor areas seem to be generally involved in predictive tasks, i.e. not only in action perception.

But why should the motor system be involved in predictive tasks? For accurate visuomotor control it is fundamental to know our body state. However, sensory signals that convey information about body states are subject to significant delays due to receptor transduction, neural conduction and central processing. Using sensory information to estimate the body state can lead to large errors, especially for fast movements. An alternative is to estimate state using prediction based on motor commands (i.e. by using efference copies). The estimate can be made ahead of the movement, which solves the time delay problem. In order to fulfill this function, predictive (forward) models need to be time locked to the currently ongoing real event, i.e. need to be online or real-time simulations. Thus, to control one's own actions requires forward models which predict the sensory consequences of actions in real-time (e.g., Wolpert and Flanagan, 2001).[1] It has been proposed that these already existing real-time forward models for the control of one's own actions may be used also to predict the actions of *others* (Grush, 2004; Blakemore and Frith, 2005; Prinz, 2006; Wolpert and Flanagan, 2001; Wolpert et al., 2003). However, to our knowledge, evidence for real-time prediction of others' actions has not yet been published so far.

The aim of the present study is to investigate whether the prediction of others' actions involves real-time simulation processes. We studied real-time processes based on the following logic: Observers perceived brief sequences of point-light actions, followed by an occluder and then a test posture. We independently varied the duration of the occluder (occluder time of 100, 400 or 700 ms) and the distance between the end of the movie and the test posture in terms of the number of frames in the sequence

---

[1] Predictive models provide a number of further advantages, like canceling out the effect of sensory changes induced by self-motion (reafference), filter sensory information, i.e. attenuating unwanted information or highlighting information critical for control. In addition, prediction allows us to determine whether motion of our bodies has been generated by us, or by an external agent (Wolpert and Flanagan, 2001).

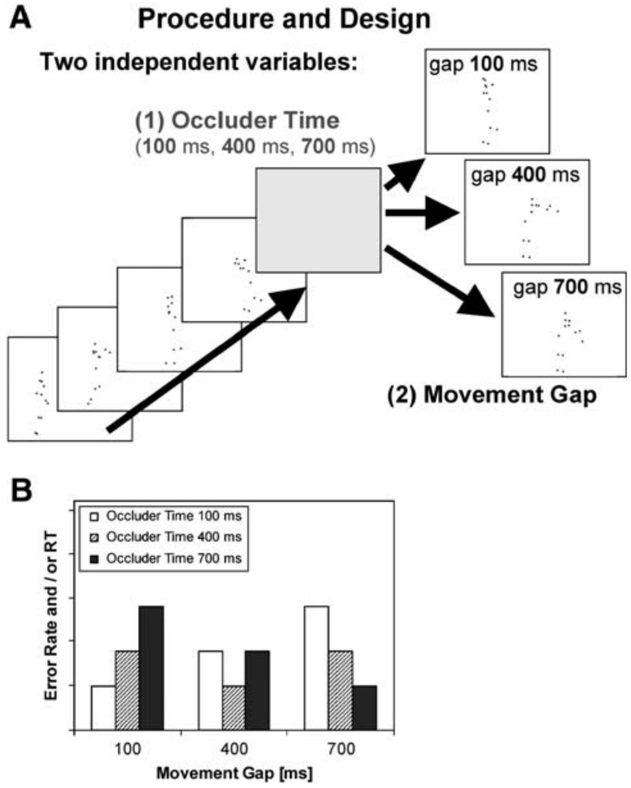

Fig. 1. (A) Every trial started with fixation dot (not depicted) followed by a brief action sequence. Then an occluder was presented, followed by a test posture. Both occluder time and movement gap are varied independently. (B) Results expected according to the real-time prediction hypothesis. According to the real-time prediction hypothesis, performance should deteriorate with increasing distance between occluder time and movement gap. More precisely, performance should be best when occluder time and movement gap correspond. For instance, error rate should be lowest when the occluder is on the screen for 700 ms, and the test posture is extrapolated into the future for 700 ms.

(corresponding to movement gaps of 100, 400 or 700 ms), as illustrated in Fig. 1A. The test posture was either a continuation in the same orientation, or was rotated in depth. Participants were instructed to decide whether the test posture was a continuation in the same orientation, or in a different orientation. According to the real-time prediction hypothesis, the internal action simulation should run time-locked to the real action, even when it is covered by an occluder. Performance in Same trials (demanding a 'same' response) should be best when occluder time and movement gap correspond, because then the action representation (updated in real-time) corresponds to the test posture. As performance should depend on the similarity between the updated internal model and the test posture, performance is predicted to deteriorate with increasing distance between occluder time and movement gap. More specifically, the real-time hypothesis predicts an interaction between occluder time and movement gap, with best performance when both variables correspond (see Fig. 1B). In contrast, in Different trials (calling for a 'different' response) there are two error sources in the stimulus, a temporal error (when occluder time and movement time do not correspond) and an additional spatial error (due to depth rotation of the test posture). As we have no definite hypothesis about the interaction between these two error sources, clear-cut predictions for Different trials cannot be derived.[2] Therefore we focus on Same trials, for which the temporal factor can be investigated in its pure form, while basic findings for Different trials are reported in Appendix A.

We used a prediction task in which explicit judgments about the timing of the actions were not requested, to avoid that potential real-time effects are induced by the task instruction. To accentuate this point, subjects were explicitly instructed to decide whether the test posture was a continuation at any point in time. We used point-light action sequences for our experiments (Johansson, 1973, 1975; for reviews see Thornton, 2006; Vanrie and Verfaillie, 2004), because these stimuli emphasize the motion information, and minimize the feasibility of other strategies for the recognition and prediction of actions.

Experiment 1

Materials and methods

Participants

Sixteen right handed subjects participated in the experiment (12 female; age range: 21–36 years; mean age = 26.9). Participants reported normal or corrected-to-normal vision and were naïve with respect to the hypotheses under investigation. Subjects were paid for their participation.

Stimuli and material

A number of familiar actions were recorded with a Vicon 612 motion capture system (Vicon Motion Systems Ltd., Oxford, UK). The 3-D positions of 41 reflecting markers were recorded with a temporal sampling rate of 120 Hz, with a spatial error less than 1 mm. Seven infrared cameras were positioned as follows in a room of about 10 × 6 m: four at the four corners, two at the mid-points of the two short sides of the rectangle, one at the midpoint of one of the two long sides. All actions were performed by a male right-handed agent. Marker trajectories were pre-processed using commercial software from Vicon, and a kinematic model was fitted to the individual actions. Nine actions were used for the test phase (lift something from floor, push something away, knee-bends, waving both arms, stand up from chair, embrace someone, put on one's boot, bowling, throwing ball overhead with both arms). The practice phase consisted of three actions that were not used in the test phase (throw basketball, leapfrog, stand up from floor). Action segments were chosen as to guarantee that their duration was long enough to allow for partial occlusion, and to ensure that they could be recognized without specific prior expertise. Cyclic actions were included only when a half cycle was longer than the prediction period (to avoid ambiguities). Unpredictable accidental movements at the end of the recorded actions were eliminated based on the judgments of two independent raters. Point-light actions were rendered with 30 Hz (i.e. we presented every fourth frame of the recorded sequence), which was fully sufficient to produce smooth action sequences. The point-light displays consisted of 13 black dots on white background, located at the major joints of the body

---

[2] It should be noted that action representations are viewpoint dependent, i.e. recognition performance deteriorates with increasing depth rotation of the stimulus (e.g., Daems and Verfaillie, 1999; Verfaillie, 1993). However, viewpoint dependency does not interfere with our experimental design, as viewpoint variations only appear in Different trials, and are the same for all possible combinations of occluder time and movement gap.

(centre of the head, shoulders, elbows, wrists, sternum, center of pelvis, knees, and ankles). These points were either original marker positions, or were computed by averaging between several original marker points (for instance, the four markers on the head were averaged to one central point). All points were located on the actor; objects used by the actor (e.g., a ball) were not visible in the point-light animations. The actions could be presented from different viewpoints. Stimuli were rendered by parallel projection from the 3-D marker positions, assuming a view direction with an altitude of 5° above the horizontal plane. The azimuth was determined for every action individually, such that the main direction of movements being towards the right side of the visual field. Global body translations were not eliminated, in order to keep the movements as natural as possible.

The experiments were conducted in a dimly lit room. Observers sat approximately 74 cm from the monitor. The point-light character was about 7 cm in height (visual angle of 5.4°) and could move within an area of 340 pixels width and 312 pixels height (about 12 cm × 11 cm, corresponding to a visual angle of 9.3° × 8.6°), located at the center of the screen. An occluder of the same size was rendered in a light gray. The stimulus dots were approximately 2 mm in diameter (10 min visual angle). Stimuli were presented on a Fujitsu Siemens 19T2 19-inch color monitor, with resolution set to 1024 × 768 pixels and a refresh rate of 120 Hz. A Fujitsu Siemens computer (3.00 GHz processor, 512 MB RAM, NVIDIA NV 44 graphics card) was used to control stimulus presentation and data collection, running Matlab 6.5 and Psychtoolbox (Brainard, 1997; Pelli, 1997).

Procedure and design

A trial started with a central fixation dot (500 ms), followed by a smooth movie sequence depicting the beginning of an action. Then an occluder was presented for either 100, 400 or 700 ms, immediately followed by a static test posture (see Fig. 1A). The test posture showed a continuation of the action, either in the same orientation as the action in the movie sequence, or in a different depth orientation (i.e. either rotated in depth to the right or to the left side). Subjects were instructed to decide whether the test posture was a continuation of the action in the same orientation — at any point in time. They should respond by pressing a key with their right or left index finger. Subjects were instructed to respond as quickly as possible while keeping the number of errors low.

The duration of the action sequences in the test phase varied from 1299 to 4162 ms (43–129 frames, with 30 Hz frame rate). The test image was presented for maximally 1000 ms, or until response. When subjects did not respond within 2500 ms after the onset of the test stimulus, the trial was aborted and was not included in the data analysis. The next trial started 1100 ms after the subject's response (respectively after trial abortion). The experiment consisted of a brief familiarization phase, followed by a practice and a test phase. In the familiarization phase all actions were presented twice, together with the name of the action. The practice phase consisted of 54 trials, in which subjects received feedback about prediction accuracy after every trial. The test phase, in which no feedback was given, consisted of 648 trials (9 actions × 3 movement gaps × 3 occluder times × 2 response versions × 4 repetitions). Within each repetition, trial order was randomized differently for each subject. There was a self-timed break after 72 trials (overall 9 blocks). The whole experiment lasted about 75 min.

Two independent variables were varied within subjects. The first was the duration of the occluder, which was on the screen for 100, 400, or 700 ms. The second variable was the movement gap, i.e. the distance between the end of the action sequence and the test posture (in terms of the number of frames in the action sequence), again corresponding to 100, 400, or 700 ms. For every correct continuation two distractor postures were created by depth-rotating the posture to the left or to the right. The rotation angles were adjusted for every action and movement gap individually, such that the dissimilarity between correct postures and distractors was comparable for different actions, as rated by three judges. The distractor stimuli, i.e. the depth rotated postures, were similar to the correct continuations, so that overall task difficulty was high. Consequently, data analysis focused on the error data. RTs were analyzed only for correct responses. Analyses of variance (ANOVAs) for repeated measurements were used for data analysis. We will discuss only Same trials, because there are no concise predictions for Different trials (basic results for Different trials are reported in Appendix A).

Results and discussion

Error rates

The results for Same trials (i.e. trials demanding a 'same' response) nicely fitted the predicted pattern under the real-time hypothesis. In order to investigate whether performance deteriorated with increasing distance between occluder time and movement gap, we averaged over cells with the same distance level, i.e. over distances of 0 ms, 300 ms, and 600 ms respectively. Errors (i.e. misses) increased monotonically with increasing distance between occluder time and movement gap (Fig. 2A), as predicted by the real-time hypothesis. A 1-way ANOVA indicated a reliable effect of distance ($F(2,30) = 15.76$; $p < 0.001$).[3] The distance function showed a highly significant linear trend, suggesting a monotonic increase, $F(1,15) = 19.30$; $p = 0.001$, and a significant quadratic trend, $F(1,15) = 4.90$; $p = 0.043$. Overall, error rates increased monotonically with increasing distance between occluder time and movement gap, in accordance with the distance prediction of the real-time hypothesis. The real-time pattern was confirmed in more specific further analyses. Performance was best when movement gap and occluder time corresponded (100–100 ms, 400–400 ms, 700–700 ms), i.e. was best when the test posture corresponded to the real-time updated action representation (see Fig. 2B). This was confirmed by a 3 (occluder time) × 3 (movement gap) ANOVA indicating a significant interaction between occluder time and movement gap ($F(4,60) = 6.03$; $p < 0.001$). Neither the main effect of occluder time was significant ($F(2,30) < 1$, n.s.), nor the effect of movement gap ($F(2,30) = 2.91$; $p = 0.070$, n.s.). We performed separate 1-way ANOVAs for each of the three movement gaps. Significant effects of occluder time were found for movement gap 100 ms ($F(2,30) = 6.20$; $p = 0.006$) and movement gap 700 ms ($F(2,30) = 4.30$; $p = 0.023$), but not for movement gap of 400 ms ($F(2,30) = 1.35$; $p = 0.274$, n.s.).

Reaction times

Similar to the error rates, RTs in Same trials increased with increasing distance between occluder time and movement gap. A 1-way ANOVA for RTs indicated a reliable distance effect ($F(2,30) = 14.44$; $p < 0.001$). The distance function showed a highly significant linear trend, $F(1,15) = 35.79$; $p < 0.001$, and a significant quadratic

---

[3] This distance analysis corresponds to an analysis of variance with contrasts set according to the hypothesis of real-time prediction.

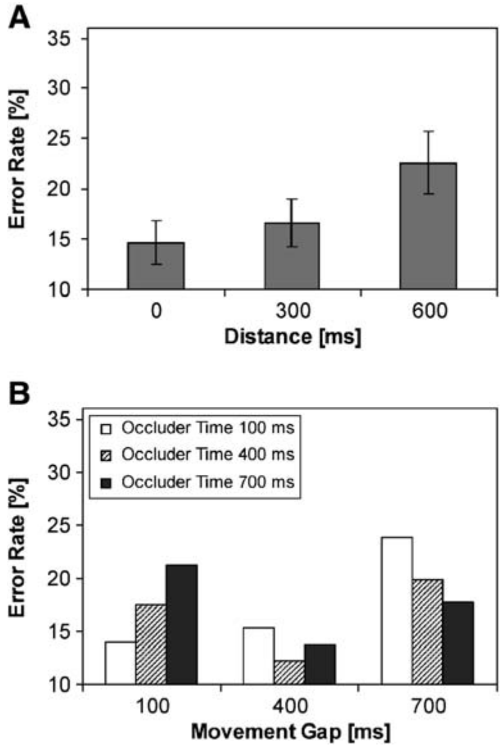

Fig. 2. (A) Error rates in Experiment 1 plotted as a function of the distance between occluder time and movement gap. Error bars represent standard errors of the mean. (B) Error rates are plotted for the different movement gaps and occluder times. Error rates correspond well to the hypothesized real-time pattern (see Fig. 1B). Performance is best when occluder time and movement gap correspond.

trend, $F(1,15)=8.08$, $p=0.012$. This monotonic distance effect for RTs is consistent with the real-time hypothesis. There was no indication for a speed-accuracy trade-off. The more specific analysis, however, showed a somewhat different pattern of results than for the error data. RTs were longer for brief occluder times, and especially for occluder times of 100 ms (see Table 1). A 3 (occluder time) × 3 (movement gap) ANOVA confirmed this pattern, with a highly significant main effect of occluder time ($F(2,30)=33.51$; $p<0.001$). Also the main effect of movement gap was significant ($F(2,30)=3.51$; $p=0.043$), while the interaction between occluder time and movement gap was only marginally significant ($F(4,60)=2.18$; $p=0.082$, n.s.).

Thus, while error rates for Same trials conformed to the expectations, reaction times did only partially. We suspect that this is due to the high degree of temporal uncertainty entailed in our task. Reaction times are known to be sensitive to temporal uncertainty, i.e. to the predictability of the preparatory interval elapsing between a warning signal – that precedes the imperative stimulus – and the imperative stimulus itself (e.g., Bertelson, 1967; Bertelson and Tisseyre, 1968; Requin et al., 1991). In the present task the imperative stimulus was provided by the test posture to which participants replied, and the warning signal was provided by action offset. Since occluder times were completely randomized, preparation time was unpredictable on each trial. As is well known from the literature, response preparation is under such conditions particularly poor for 'early' trials, i.e. trials on which the imperative stimulus is presented earlier than on average (e.g. Bertelson, 1967; Bertelson and Tisseyre, 1968; Sanders, 1998, p. 172). This is exactly the pattern our data show: RTs were exceptionally high for short occluder times. In retrospect we may thus conclude that RTs are partly inconclusive with respect to our hypotheses, because in this measure the predicted effects of real-time simulation are superimposed, and in fact masked, by temporal uncertainty effects.

Overall, Experiment 1 provided reasonable evidence for real-time updating in an action prediction task. In Same trials, subjects made more errors with increasing distance between occluder time and movement gap. More specifically, error rates were lowest when occluder time and movement gap corresponded, and the pattern of interaction matched the real-time hypothesis. Thus, accuracy was highest when the test posture corresponded to an assumed internal action model updated in real-time. In general, RTs were more difficult to interpret than errors, because the RT pattern seems to be superimposed by a foreperiod effect. Nevertheless, there was a monotonic distance effect for RTs. Together, the findings of Experiment 1 confirm the hypothesis of real-time

Table 1
Performance for same trials (Experiments 1–4)

RTs Experiment 1 (same trials) [ms]				
Distance		0 ms	300 ms	600 ms
		741 (35)	733 (31)	773 (36)
Full pattern		Movement Gap		
		100 ms	400 ms	700 ms
Occluder time	100 ms	765 (39)	760 (32)	816 (36)
	400 ms	728 (39)	723 (34)	737 (27)
	700 ms	733 (39)	710 (30)	739 (35)
RTs Experiment 2 (same trials) [ms]				
Distance		0 ms	300 ms	600 ms
		813 (41)	816 (36)	847 (39)
Full pattern		Movement gap		
		100 ms	400 ms	700 ms
Occluder time	100 ms	831 (56)	859 (37)	902 (43)
	400 ms	805 (45)	803 (38)	837 (38)
	700 ms	810 (43)	782 (35)	817 (37)
RTs Experiment 3 (same trials) [ms]				
Distance		0 ms	300 ms	600 ms
		756 (39)	757 (38)	768 (40)
Full pattern		Movement gap		
		100 ms	400 ms	700 ms
Occluder time	100 ms	786 (35)	800 (34)	817 (44)
	400 ms	734 (40)	729 (48)	761 (42)
	700 ms	724 (39)	738 (42)	757 (40)
Error rates Experiment 4 (same trials) [%]				
Distance		0 ms	300 ms	600 ms
		15.9 (2.3)	16.6 (2.4)	22.1 (2.8)
Full pattern		Movement gap		
		100 ms	400 ms	700 ms
Occluder time	100 ms	13.8 (3.7)	13.9 (2.2)	21.8 (2.6)
	400 ms	20.7 (4.9)	13.9 (2.8)	19.5 (2.3)
	700 ms	22.6 (4.4)	12.4 (2.2)	20.0 (2.5)

Numbers in parentheses are standard errors of the means.

simulation. The data fit well with the predicted real-time pattern, considering that subjects may use alternative strategies for some actions (e.g., by focusing onto relations between specific points, for instance the legs, and testing whether this relation remains identical or similar).

Experiment 2

In Experiment 2 we investigated whether we can replicate this pattern of results, using an extended set of actions. Predictions were identical to Experiment 1.

Materials and methods

Participants

Sixteen right handed subjects participated in the experiment (10 female; age range: 19–34 years; mean age $=25.3$).

Stimuli, procedure and design

The same methods were used as in Experiment 1, except for the following changes. Twelve actions were used for the test phase (lift something from floor, push something away, knee-bends, waving both arms, stand up from chair, stand up from floor, stand up from knees, embrace someone, jumping forward, leapfrog, put on one's boot, throwing ball overhead with both arms). The practice phase consisted of three actions that were not used in the test phase (throw basketball, bowling, throwing ball sideways). The duration of the movie sequences for the actions in the test phase varied from 899 to 4162 ms (31–129 frames). The test phase consisted of 648 trials (12 actions × 3 movement gaps × 3 occluder times × 2 response versions × 3 repetitions). The practice phase consisted of 36 trials; it was repeated when subjects still had problems with the task.

Results and discussion

Error rates

As in Experiment 1, the results fitted well with the hypothesized real-time pattern. The distance analysis showed a deterioration of error rates with increasing distance between occluder time and movement gap (Fig. 3A). A 1-way ANOVA indicated a significant effect of distance ($F(2,30)=12.23$; $p<0.001$). Moreover, trend analyses showed a significant linear trend, $F(1,15)=24.48$, $p<0.001$, while a quadratic trend was not reliable, $F(1,15)<1$. In accordance with the prediction, performance deteriorated in a monotonic way with increasing distance. In a more specific analysis, performance was best when occluder time and movement gap corresponded (Fig. 3B). As predicted, a 3 (occluder time) × 3 (movement gap) ANOVA indicated a significant interaction ($F(4,60)=7.23$; $p<0.001$). Both main effects were not reliable (occluder time: $F(2,30)=1.68$; $p=0.204$; movement gap: $F(2,30)<1$). Separate 1-way ANOVAs showed significant effects of occluder time for all three movement gaps (for movement gap 100 ms, $F(2,30)=4.97$; $p=0.014$; for movement gap 400 ms, $F(2,30)=4.29$; $p=0.023$; for movement gap 700 ms, $F(2,30)=5.96$; $p=0.007$). Thus, the error pattern corresponded well to the predicted real-time pattern.

Reaction times

In a similar way as for errors, RTs increased with increasing distance between occluder time and movement gap (Table 1), confirmed by a significant distance effect ($F(2,30)=5.64$, $p=0.008$). The distance function showed a significant linear trend ($F(1,15)=7.15$, $p=0.017$), while the quadratic trend was not reliable ($F(1,15)=2.77$, $p=0.117$). There was no indication for a speed-accuracy trade-off. As in Experiment 1, RTs were higher for brief occluder times (see Table 1). A 3 (occluder time) × 3 (movement gap) ANOVA confirmed this finding, indicating a significant main effect of occluder time, $F(2,30)=8.79$, $p=0.001$. The effect of movement gap was not significant, $F(2,30)=1.58$, $p=0.224$. There was a reliable interaction between occluder time and movement gap, $F(4,60)=2.76$, $p=0.036$, as predicted by the real-time hypothesis.

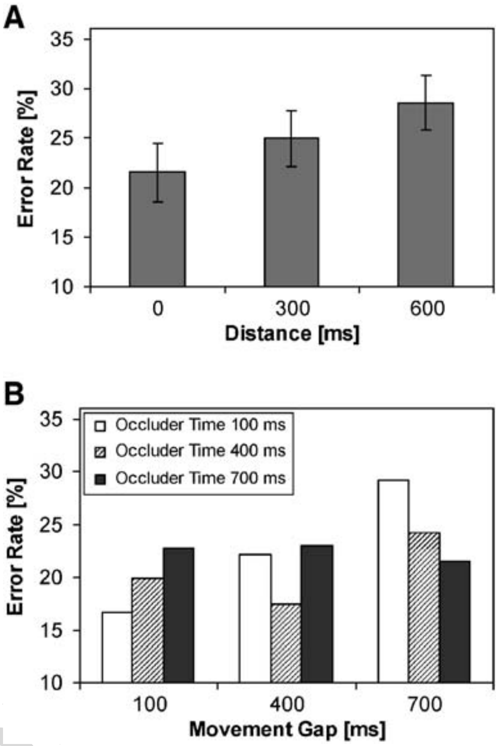

Fig. 3. (A) Error rates in Experiment 2 plotted as a function of distance between occluder time and movement gap. Error bars represent standard errors of the mean. (B) Error rates plotted for the different movement gaps and occluder times. As predicted by the real-time hypothesis, performance is best when occluder time and movement gap correspond.

To sum up, the results of Experiment 2 replicated the findings of Experiment 1. Errors and RTs in Same trials increased with increasing distance between occluder time and movement gap. Error rates were lowest when occluder time and movement gap corresponded, i.e. when the test posture corresponded to the real-time updated internal model. Together, the results confirm the real-time hypothesis.

Experiment 3

In Experiment 3 we investigated how stable the real-time pattern is, and whether it can be compromised by changes in stimulus presentation. It is well established that action perception is impaired for inverted actions, i.e. for actions presented upside down (e.g., Pavlova and Sokolov, 2000; Troje, 2003; Loula et al., 2005), although the low-level visual and motion features are highly similar for upright and inverted actions. Inverted actions violate gravitational constraints (Troje, 2003), and we have little experience for actions performed upside down, neither motor nor visual (e.g. Grossman and Blake, 2001; Hiris et al., 2005; Jastorff et al., 2006; Loula et al., 2005). As action perception and the activation in the mirror system depend on expertise (Calvo-Merino et al., 2005; Casile and Giese, 2006), real-time prediction should be impaired for inverted actions. Thus, we predicted that the real-time pattern would break down for inverted actions.

Materials and methods

Participants

Seventeen right-handed individuals participated in the experiment against payment. The data of one participant were excluded because accuracy for Same trials was at chance level (50.2%). Data analysis was based on the data of 16 participants (11 female; age range: 20–35 years; mean age=26.5).

Stimuli, procedure and design

The same methods were used as in Experiment 2, except that the images were turned upside down by flipping them around the horizontal axis, so that the direction of movement remained unchanged. Thus, both the action sequence and the test image were upside down. Also in the familiarization phase the actions were presented upside down.

Results and discussion

Error rates

Errors in Different trials were much higher in Experiment 3 than in Experiment 2 (Experiment 2: $M=35.1$, $SE=3.4$; Experiment 3: $M=48.3$, $SE=2.6$), $t(15)=-2.70$, $p=0.016$ (two-tailed), reflecting higher task difficulty. The overall Same trial error rates in Experiment 3 did not differ significantly from Experiment 2 (Experiment 2: $M=22.3$, $SE=2.1$; Experiment 3: $M=23.2$, $SE=1.9$), as indicated by a paired-samples t test, $t(15)=-.33$, $p=0.744$ (two-tailed). Nevertheless, the pattern of Same trial errors was very different from that of the previous experiments. There was a significant effect of distance between occluder time and movement gap ($F(2,30)=5.20$; $p=0.011$), but the effect was small (see Fig. 4A). More important, a linear trend was not reliable ($F(2,30)=3.87$, $p=0.068$), while a quadratic trend was significant ($F(2,30)=7.10$, $p=0.018$). Thus, in contrast to the previous experiments, there was no monotonic distance effect for errors. The absence of the real-time pattern is more obvious in the specific analyses. In a 3 × 3 ANOVA, occluder time was not significant ($F(2,30)<1$). In contrast to the previous experiments, there was a reliable effect of movement gap ($F(2,30)=7.65$, $p=0.002$). Importantly, the interaction between occluder time and movement gap was not significant, $F(4,60)=2.04$, $p=0.100$. This pattern of results clearly deviates from the real-time pattern. Performance was not best when occluder time and movement gap corresponded, but simply depended on the distance of the test posture to the last frame of the action sequence (see Fig. 4B). None of the 1-way ANOVAs for the different movement gap conditions was significant (movement gap 100 ms: $F(2,30)=1.18$, $p=0.322$, n.s.; movement gap 400 ms: $F(2,30)<1$, n.s.; movement gap 700 ms: $F(2,30)=2.79$, $p=0.077$, n.s.).

Reaction times

Overall Same trial RTs were shorter in Experiment 3 than in Experiment 2 (Experiment 2: $M=825$, $SE=38$; Experiment 3: $M=760$, $SE=39$), but this difference was not significant, $t(15)=1.05$, $p=0.311$ (two-tailed). The distance effect for RTs was not significant ($F(2,30)=1.41$, $p=0.259$), confirming the break-down of the real-time pattern for inverted actions. There was neither a reliable linear trend ($F(1,15)=1.97$, $p=0.180$, n.s.), nor a quadratic trend ($F(1,15)<1$, n.s.). RTs were again higher for brief occluder times (see Table 1). A 3 (occluder time) × 3 (movement gap) ANOVA indicated a significant effect of occluder time, which suggests the presence of a foreperiod effect, $F(2,30)=14.78$, $p<0.001$. The main effect of movement gap was not reliable, $F(2,30)=2.39$, $p=0.108$, and there was no evidence for an interaction between occluder time and movement gap ($F(4,60)<1$, n.s.). Thus, there was no real-time pattern in the RTs when the actions were inverted.

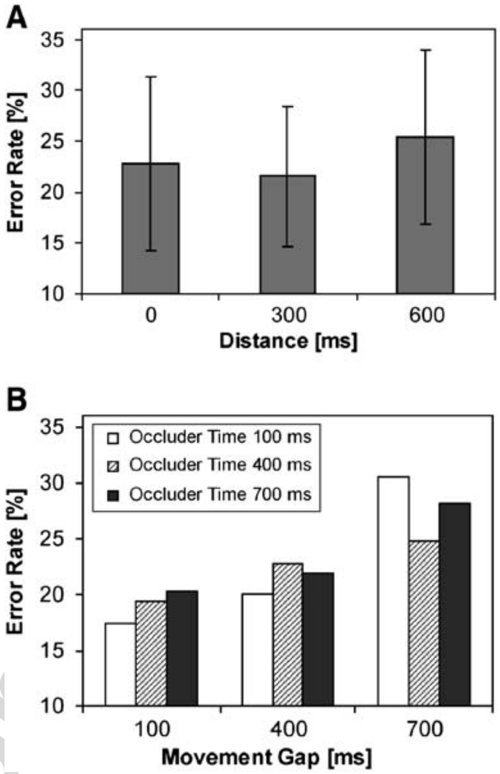

Fig. 4. (A) Error rates in Experiment 3 with inverted actions, plotted as a function of distance between occluder time and movement gap. Error bars represent standard errors of the mean. (B) Errors for the different movement gaps and occluder times. There is no benefit when occluder time and movement gap correspond. Instead, performance deteriorates with increasing distance to the end of the action sequence.

Overall, the pattern of results for inverted actions did not correspond to the real-time pattern. Prediction accuracy depended now on the distance to the last frame of the sequence, and there was no benefit when occluder time and movement gap corresponded. Correspondingly, there was no interaction between occluder time and movement gap, neither for errors nor for RTs. As predicted, the real-time pattern broke down for inverted actions, i.e. when subjects had little expertise with the action stimuli.

Experiment 4

In previous experiments RTs were difficult to interpret because response latencies were superimposed by an additional foreperiod effect (e.g., Bertelson, 1967; Bertelson and Tisseyre, 1968), i.e., RTs were longer for brief occluder times. In Experiment 4 we investigated whether we can reduce the temporal uncertainty which causes the foreperiod effect, and obtain a more stable real-time pattern for RTs. By blocking occluder times participants should be able to build up an expectation about how long the occluder would be on the screen, which in turn may reduce the foreperiod effect. Thus, we predicted a real-time pattern also for RTs when occluder times are blocked.

Materials and methods

Participants

Seventeen right-handed individuals participated in the experiment against payment. The data of one participant had to be excluded because the debriefing after the experiment indicated that she misunderstood the instruction. Data analysis was based on the data of 16 participants (11 female; age range: 20–39 years; mean age = 26.2).

Stimuli, procedure and design

The same methods (and actions) were used as in Experiment 1, except for the following changes. Trials were blocked by occluder time, both in the practice and in the test phase. Trials were divided into 12 blocks with 54 trials with identical occluder times. Subjects were informed before each block about the occluder duration (short, middle, or long). The sequence of blocks was varied pseudorandomly, and was counterbalanced over subjects. In the practice phase a block consisted of 18 trials, starting with occluder time of 400 ms, followed by 100 ms and 700 ms. Given the aim of the experiment, our focus was on the RT data.

Results and discussion

Error rates

As in Experiments 1 and 2, the error rates in Same trials fitted well with the hypothesized real-time pattern. The distance analysis showed a deterioration of error rates with increasing distance between occluder time and movement gap (Table 1). A 1-way ANOVA indicated a significant effect of distance ($F(2,30) = 13.45$; $p < 0.001$). Moreover, trend analyses showed a significant linear trend, $F(1,15) = 18.02$, $p = 0.001$, and a significant quadratic trend, $F(1,15) = 5.85$, $p = 0.029$. In a more specific analysis, performance tended to be best when occluder time and movement gap corresponded (Table 1). As predicted, a 3 (occluder time) × 3 (movement gap) ANOVA indicated a significant interaction ($F(4,60) = 4.19$; $p = 0.005$). Both main effects were not reliable (occluder time: $F(2,30) = 1.06$; $p = 0.358$; movement gap: $F(2,30) = 2.98$, $p = 0.066$). Separate 1-way ANOVAs for the different movement gaps showed significant effects of occluder time for movement gap of 100 ms ($F(2,30) = 5.15$; $p = 0.012$), but not for movement gap 400 ms ($F(2,30) < 1$, n.s.) and movement gap 700 ms ($F(2,30) = 1.41$; $p = 0.260$, n.s.). Thus, the error rates corresponded to the predicted real-time pattern, although not as perfectly as in Experiments 1 and 2.

Reaction times

More important for this experiment, which focused on reaction times, RTs increased with increasing distance between occluder time and movement gap (Fig. 5A). This was confirmed by a significant distance effect ($F(2,30) = 14.08$, $p < 0.001$). The distance function showed a significant linear trend ($F(1,15) = 28.87$, $p < 0.001$), while the quadratic trend was not reliable ($F(1,15) = 2.05$, $p = 0.172$). There was no indication for a speed-accuracy trade-off. In contrast to the previous experiments, RTs for brief occluder times were not slowest, but were in a similar range as RTs for long occluder times. A paired-samples t test showed no significant difference between RTs for brief (100 ms) and long (700 ms) occluder times, $t(15) = 0.335$, $p = 0.742$ (two-tailed), n.s. This indicates that our blocking manipulation was successful in diminishing the foreperiod effect (see Fig. 5B). A 3 (occluder time) × 3 (movement gap) ANOVA still showed a significant main effect of occluder time, $F(2,30) = 3.58$, $p = 0.040$, but this seemed to be due to somewhat faster RTs for intermediate occluder times. The effect of movement gap was not significant, $F(2,30) = 2.60$, $p = 0.091$. The interaction between occluder time and movement gap was highly significant ($F(4,60) = 5.70$, $p = 0.001$), as predicted by the real-time hypothesis, and the specific pattern of interaction corresponded well with the real-time pattern (Fig. 5B). Separate 1-way ANOVAs for the different movement gaps showed significant effects of occluder time for movement gap of 100 ms ($F(2,30) = 5.61$; $p = 0.009$), for movement gap 400 ms ($F(2,30) = 7.62$, $p = 0.002$), but not for movement gap 700 ms ($F(2,30) = 1.63$; $p = 0.213$). Thus, when occluder times were blocked, the real-time pattern was present also in the RT data.

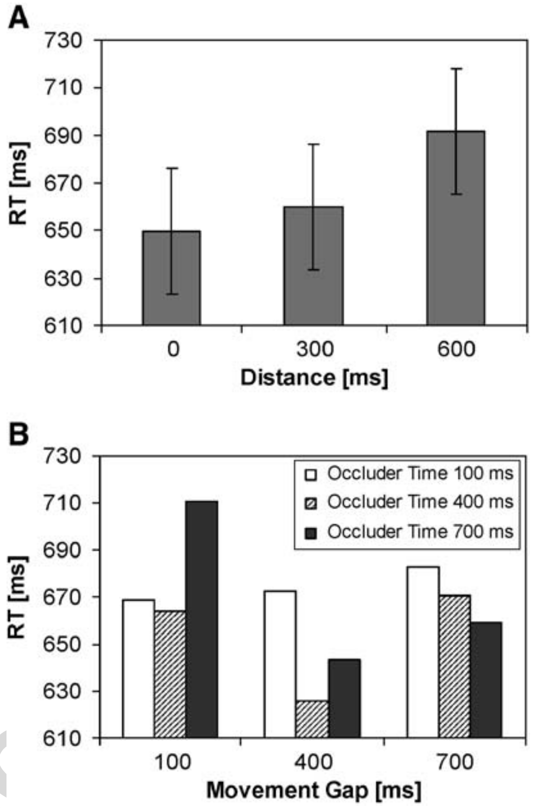

Fig. 5. (A) RTs in Experiment 4 plotted as a function of the distance between occluder time and movement gap. Error bars represent standard errors of the mean. (B) RTs are plotted for the different movement gaps and occluder times. Error rates correspond well to the hypothesized real-time pattern. Performance is best when occluder time and movement gap correspond.

To sum up, the results of Experiment 4 showed the full real-time pattern also for RTs. When blocked for occluder times, RTs increased with increasing distance between occluder time and movement gap. More specifically, the RT pattern matched the hypothesized real-time pattern of interaction. RTs were fastest when occluder time and movement gap corresponded, i.e. when the real-time updated internal model corresponded to the test posture. The real-time pattern was somewhat noisier for the error rates than in previous experiments. Nevertheless, the distance effect and the interaction between occluder time and movement gap were significant also for errors. Basically, Experiment 4 extends the findings of Experiments 1 and 2 to RT data. Thus, when the timing of the occluder was predictable, the real-time pattern was not covered by the foreperiod effect, showing a real-time pattern also for RTs.

General discussion

Four experiments provided evidence for real-time simulation in the prediction of point-light actions. Internal models used for prediction seem to run time-locked to ongoing perceptual events in the environment, even when the action is occluded. This is supported by the following findings. First, performance in Same trials varied systematically with the distance to the hypothesized internal model updated in real-time. Second, accuracy was best when the test image corresponded to the updated internal model, and there was a significant pattern of interaction between occluder time and movement gap (Experiments 1, 2 and 4). For RTs, this pattern was superimposed by a foreperiod effect (Experiments 1–2). However, when occluder times were blocked, and thus easier to predict, this interaction pattern was found also for RTs (Experiment 4). Third, real-time simulation broke down for inverted actions (Experiment 3), i.e. when subjects have little or no experience with the action stimuli. This suggests that real-time prediction requires visual and/or motor expertise. The break-down for inverted actions may result from costs due to compensation processes required for the processing of the inverted actions (e.g., Murray, 1997; for review see Graf, 2006), i.e. for orientations for which participants have little expertise.

Note that we found evidence for real-time processes, although we used a task in which judgments about the timing of the actions were not explicitly requested. Even more, subjects were explicitly instructed to decide whether the test posture was a continuation at *any* point in time. In addition, the real-time pattern appeared although the relative trial frequency did not induce a real-time strategy. The design of the experiments did not encourage subjects to expect a correspondence between occluder time and movement gap (and thus to engage in a real-time strategy), because both variables corresponded in only one third of the trials.

Our findings accord with previous evidence for a predictive function of simulation in action perception (e.g., Flanagan and Johansson, 2003; Verfaillie and Daems, 2002). Moreover, our experiments extend these previous studies by specifying the timing of action prediction, suggesting that action representations are updated in real-time. Thus, our results provide important insights about the time course of predictions in action perception.

However, we want to make clear that our findings have two limitations. First, the present results do not allow to conclude that the real-time pattern is specific for actions. Real-time effects may extend to non-biological motion patterns with artifact objects, as typically investigated in the apparent motion literature, or to non-human motion stimuli based on artificial (i.e. non-human) skeleton models that cannot be mapped to our action repertoire — as long as subjects have sufficient expertise with the stimuli (Jastorff et al., 2006).

Second, although the real-time hypothesis has been derived from research on visuomotor control, our results do not allow to decide whether the motor system is involved in real-time prediction. The pattern of results may be due to purely perceptual real-time processes, without incorporating the motor system. Note that a current perceptual model of action recognition involves predictive aspects on a purely perceptual basis (Giese and Poggio, 2003; Jastorff and Giese, 2004; Xie and Giese, 2002). Neurophysiological evidence suggests predictive coding in STS, a visual area which is selective for biological motion (Jellema and Perrett, 2003b). However, many predictive cells in STS had long onset latencies (~500 ms), which gives ample time for top-down input from the mirror system. Accordingly, recurrent feedback from fronto-parietal mirror areas to STS has been demonstrated in fMRI studies with humans (Iacoboni et al., 2001; Iacoboni, 2005). An involvement of the (pre-) motor cortex in real-time simulation seems likely, given the evidence that premotor cortex is activated in purely perceptual serial prediction tasks, even with stimuli that cannot be mapped to the human body (Schubotz and von Cramon, 2002, 2004). Beyond such purely perceptual predictive simulation a variety of results suggest that, at least for imitable actions, the mirror neuron system might be involved in such internal simulation (e.g., Buccino et al., 2001, 2004; Calvo-Merino et al., 2005; Iacoboni et al., 2005). In addition, it has been shown that apparent motion perception with human body stimuli is influenced by biomechanical movement limitations of the body, suggesting an involvement of action-relevant information in a perceptual apparent motion task (Shiffrar and Freyd, 1990, 1993).

The inversion of point-light action displays may affect both motor and visual information, as indicated by a study investigating the role of motor and visual experience in the visual analysis of the identity of human point-light actors (Loula et al., 2005). Participants viewed point-light descriptions of themselves, a friend, and a stranger performing various actions. In action identification and discrimination tasks, sensitivity to one's own motion was highest. Visual sensitivity to friends', but not strangers' actions, was above chance. The improved performance for self-action suggests that motor experience defines sensitivity to human movement, because we have much motor but little visual experience with our own actions. In contrast, we have visual but little or no motor experience with our friends' actions. In a control study with inverted displays, performance was at chance in all three conditions (Loula et al., 2005). This suggests that both motor-related and visual processes may be disturbed by inverting the actions. Consequently, the break-down of the real-time pattern for inverted actions may be due to an impairment of perceptual and motor processes. Alternatively, it is conceivable that the break-down has a purely perceptual basis.

In order to decide whether the motor system is involved in real-time prediction, at least three research strategies seem helpful. First, functional MRI may help to disentangle perceptual and motor aspects in real-time prediction tasks. Second, dual task paradigms can be used to investigate whether the real-time pattern breaks down when a motor dual task has to be performed simultaneously, leading to a possible interference in the motor system. Third, experiments using motor learning paradigms, or including different levels of motor expertise, may help to clarify this issue.

Acknowledgments

The project was funded by the Max Planck Society. We thank Jennifer Schiller for help with smoothing the Vicon data and creating the point light stimuli, and Markus Goldbach for occasional support with programming.

Appendix A. Effects of distance for Different trials in Experiments 1–4

Experiment 1 (Different trials)	Distance		
	0 ms	300 ms	600 ms
Error rates [%]	31.1 (3.3)	30.2 (3.1)	26.0 (3.3)
RTs [ms]	797 (31)	807 (33)	803 (29)

Experiment 2 (Different trials)	Distance		
	0 ms	300 ms	600 ms
Error rates [%]	36.8 (3.8)	36.0 (3.3)	32.4 (3.3)
RTs [ms]	864 (39)	862 (41)	848 (38)

Experiment 3 (Different trials)	Distance		
	0 ms	300 ms	600 ms
Error rates [%]	49.7 (2.7)	49.8 (2.7)	45.3 (2.9)
RTs [ms]	837 (42)	836 (43)	865 (51)

Experiment 4 (Different trials)	Distance		
	0 ms	300 ms	600 ms
Error rates [%]	43.2 (5.2)	43.1 (4.7)	40.5 (4.6)
RTs [ms]	763 (19)	761 (22)	770 (23)

Errors in Different trials are false alarms. Numbers in parentheses are standard errors of the means.

References

Arbib, M.A., 2005. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. Behav. Brain Sci. 28, 105–124.

Bertamini, M., 1993. Memory for positions and dynamic representations. Mem. Cogn. 21, 449–457.

Bertelson, P., 1967. The time course of preparation. Q. J. Exp. Psychol. 19, 272–279.

Bertelson, P., Tisseyre, F., 1968. The time course of preparation with regular and irregular foreperiods. Q. J. Exp. Psychol. 20, 297–300.

Blakemore, S.-J., Decety, J., 2001. From the perception of action to the understanding of intention. Nat. Rev., Neurosci. 2, 561–567.

Blakemore, S.-J., Frith, C., 2005. The role of motor contagion in the prediction of action. Neuropsychologia 43, 260–267.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur. J. Neurosci. 13, 400–404.

Buccino, G., Vogt, S., Ritzi, A., Fink, G.R., Zilles, K., Freund, H.-J., Rizzolatti, G., 2004. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. Neuron 42, 323–334.

Calvo-Merino, B., Glaser, D.E., Grèzes, J., Passingham, R.E., Haggard, P., 2005. Action observation and acquired motor skills: an fMRI study with expert dancers. Cereb. Cortex 15, 1243–1249.

Casile, A., Giese, M.A., 2006. Non-visual motor learning influences the recognition of biological motion. Curr. Biol. 16, 69–74.

Chartrand, T.L., Bargh, J.A., 1999. The chameleon effect: the perception–behavior link and social interaction. J. Pers. Soc. Psychol. 76, 893–910.

Daems, A., Verfaillie, K., 1999. Viewpoint-dependent priming effects in the perception of human actions and body postures. Vis. Cogn. 6, 665–693.

De Renzi, E., Cavalleri, F., Facchini, S., 1996. Imitation and utilization behavior. J. Neurol. Neurosurg. Psychiatry 61, 396–400.

de Valois, R.L., de Valois, K.K., 1991. Vernier acuity with stationary moving Gabors. Vision Res. 31, 1619–1626.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. Exp. Brain Res. 91, 176–180.

Eagleman, D.M., Sejnowski, T.J., 2000. Motion integration and postdiction in visual awareness. Science 287, 5460.

Finke, R.A., Freyd, J.J., 1985. Transformations of visual memory induced by implied motions of pattern elements. J. Exper. Psychol., Learn., Mem., Cogn. 11, 780–794.

Flanagan, J.R., Johansson, R.S., 2003. Action plans used in action observation. Nature 424, 769–771.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G., 2005. Parietal lobe: from action organization to intention understanding. Science 308, 662–667.

Freyd, J.J., Finke, R.A., 1984. Representational momentum. J. Exper. Psychol., Learn., Mem., Cogn. 10, 126–132.

Gallese, V., 2005. Embodied simulation: from neurons to phenomenal experience. Phenomenol. Cogn. Sci. 4, 23–48.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 2002. Action representation and the inferior parietal lobule. In: Prinz, W., Hommel, B. (Eds.), Common mechanisms in perception and action. Attention and Performance XIX. Oxford Univ. Press, pp. 334–355.

Giese, M.A., Poggio, T., 2003. Neural mechanisms for the recognition of biological movements. Nat. Rev., Neurosci. 4, 179–192.

Graf, M., 2006. Coordinate transformations in object recognition. Psychol. Bull. 132, 920–945.

Grossman, E.D., Blake, R., 2001. Brain activity evoked by inverted and imagined biological motion. Vision Res. 41, 1475–1482.

Grush, R., 2004. The emulation theory of representation: motor control, imagery, and perception. Behav. Brain Sci. 27, 377–442.

Hatfield, E., Cacioppo, J.T., Rapson, R.L., 1994. Emotional contagion. Cambridge Univ. Press, New York.

Haueisen, J., Knösche, T.R., 2001. Involuntary motor activity in pianists evoked by music perception. J. Cogn. Neurosci. 13, 786–792.

Hiris, E., Krebeck, A., Edmonds, J., Stout, A., 2005. What learning to see arbitrary motion tells us about biological motion perception. J. Exp. Psychol. Hum. Percept. Perform. 3, 1096–1106.

Hommel, B., Müsseler, J., Aschersleben, G., Prinz, W., 2001. The theory of event coding (TEC): a framework for perception and action planning. Behav. Brain Sci. 24, 849–937.

Iacoboni, M., 2005. Understanding others: imitation, language, and empathy. In: Hurley, S., Chater, N. (Eds.), Perspectives on imitation. From neuroscience to social science—Volume 1: Mechanisms of imitation and imitation in animals (77–100). MIT Press, Cambridge, MA.

Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., Rizzolatti, G., 1999. Cortical mechanisms of human imitation. Science 286, 2526–2528.

Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.C., Mazziotta, J.C., Rizzolatti, G., 2001. Reafferent copies of imitated actions in the right superior temporal cortex. Proc. Natl. Acad. Sci. U. S. A. 98, 13995–13999.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., 2005. Grasping the intentions of others with one's own mirror neuron system. PLoS Biol. 3 (3), e79 (0529–0535).

Jarraya, M., Amorim, M.-A., Bardy, B.G., 2005. Optical flow and viewpoint change modulate perception and memorization of complex motion. Percept. Psychophys. 67, 951–961.

Jastorff, J., Giese, M.A., 2004. Time-dependent hebbian rules for the learning of templates for visual motion recognition. In: Ilg, U., Bülthoff, H.H., Mallot, H. (Eds.), Dynamic Perception. Infix, Berlin, pp. 151–156.

Jastorff, J., Kourtzi, Z., Giese, M.A., 2006. Learning to discriminate complex movements: biological versus artificial trajectories. J. Vis. 6, 791–804.

Jeannerod, M., 2001. Neural simulation of action: a unifying mechanism for motor cognition. NeuroImage 14, S103–S109.

Jellema, T., Perrett, D.I., 2003a. Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? Neuropsychologia 41, 1728–1737.

Jellema, T., Perrett, D.I., 2003b. Perceptual history influences neural responses to face and body postures. J. Cogn. Neurosci. 15, 961–971.

Johansson, G., 1973. Visual perception of biological motion and a model for it's analysis. Percept. Psychophys. 14, 201–211.

Johansson, G., 1975. Visual motion perception. Sci. Am. 232, 76–88.

Khurana, B., Nijhawan, R., 1995. Extrapolation or attention shift? Nature 378, 565–566.

Kilner, J.M., Vargas, C., Duval, S., Blakemore, S.J., Sirigu, A., 2004. Motor activation prior to observation of a predicted movement. Nat. Neurosci. 7, 1299–1301.

Kourtzi, Z., Kanwisher, N., 2000. Activation in human MT/MST by static images with implied motion. J. Cogn. Neurosci. 12, 48–55.

Kourtzi, Z., Shiffrar, M., 1999. Dynamic representations of human body movement. Perception 28, 49–62.

Loula, F., Prasad, S., Harber, K., Shiffrar, M., 2005. Recognizing people from their movements. J. Exp. Psychol. Hum. Percept. Perform. 31, 210–220.

Murray, J.E., 1997. Flipping and spinning: spatial transformation procedures in the identification of rotated natural objects. Mem. Cogn. 25, 96–105.

Nijhawan, R., 1994. Motion extrapolation in catching. Nature 370, 256–257.

Nijhawan, R., 1997. Visual decomposition of color through motion extrapolation. Nature 386, 66–69.

Olson, I.R., Gatenby, J.C., Leung, H.-C., Skudlarski, P., Gore, J.C., 2003. Neuronal representation of occluded objects in the human brain. Neuropsychologia 42, 95–104.

Pavlova, M., Sokolov, A., 2000. Orientation specificity in biological motion perception. Percept. Psychophys. 62, 889–899.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–442.

Prinz, W., 1990. A common coding approach to perception and action. In: Neumann, O., Prinz, W. (Eds.), Relationships between perception and action: Current approaches. Springer-Verlag, Berlin, pp. 167–201.

Prinz, W., 1997. Perception and action planning. Eur. J. Cogn. Psychol. 9, 129–154.

Prinz, W., 2002. Experimental approaches to imitation. In: Meltzoff, A.N., Prinz, W. (Eds.), The imitative mind: Development, evolution, and brain bases. Cambridge Univ. Press, New York, pp. 143–162.

Prinz, W., 2005. An ideomotor approach to imitation. In: Hurley, S., Chater, N. (Eds.), Perspectives on imitation. From neuroscience to social science—Volume 1: Mechanisms of imitation and imitation in animals. MIT Press, Cambridge, pp. 141–156.

Prinz, W., 2006. What re-enactment earns us. Cortex 42, 515–517.

Requin, J., Brener, J., Ring, C., 1991. Preparation for action. In: Jennings, J.R., Coles, M.G.H. (Eds.), Handbook of cognitive Psychophysiology: central and autonomic nervous system approaches. Wiley, Chichester, pp. 357–448.

Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. Nat. Rev., Neurosci. 2, 661–670.

Sanders, A.F., 1998. Elements of human performance: Reaction processes and attention in human skill. Lawrence Erlbaum, Mahwah, NJ.

Schubotz, R.I., von Cramon, D.Y., 2002. Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. NeuroImage 15, 787–796.

Schubotz, R.I., von Cramon, D.Y., 2004. Sequences of abstract non-biological stimuli share ventral premotor cortex with action observation and imagery. J. Neurosci. 24, 5467–5474.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E.T., Brammer, M., David, A.S., 2000. The functional neuroanatomy of implicit-motion perception or 'representational momentum'. Curr. Biol. 10, 16–22.

Senior, C., Ward, J., David, A.S., 2002. Representational momentum and the brain: an investigation into the functional necessity of V5/MT. Vis. Cogn. 9, 81–92.

Shiffrar, M., Freyd, J.J., 1990. Apparent motion of the human body. Psychol. Sci. 1, 257–264.

Shiffrar, M., Freyd, J.J., 1993. Timing and apparent motion path choice with human body photographs. Psychol. Sci. 4, 379–384.

Thornton, I.M., 2006. Point-light walkers and beyond. In: Knoblich, G., Thornton, I.M., Grosjean, M., Shiffrar, M. (Eds.), Human body perception from the inside out. Oxford Univ. Press, New York, pp. 271–304.

Thornton, I.M., Hayes, A.E., 2004. Anticipating actions in complex scenes. Vis. Cogn. 11, 341–370.

Thornton, I.M., Knoblich, G., 2006. Action perception: seeing the world through a moving body. Curr. Biol. 16, R27–R29.

Troje, N., 2003. Reference frames for orientation anisotropies in face recognition and biological-motion perception. Perception 32, 201–210.

Umiltà, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., Rizzolatti, G., 2001. I know what you are doing. A neurophysiological study. Neuron 31, 155–165.

Vanrie, J., Verfaillie, K., 2004. Perception of biological motion: a stimulus set of human point-light actions. Behav. Res. Methods Instrum. Comput. 36, 625–629.

Verfaillie, K., 1993. Orientation-dependent priming effects in the perception of biological motion. J. Exp. Psychol. Hum. Percept. Perform. 19, 992–1013.

Verfaillie, K., Daems, A., 2002. Representing and anticipating human actions in vision. Vis. Cogn. 9, 217–232.

Verfaillie, K., d'Ydewalle, G., 1991. Representational momentum and event course anticipation in the perception of implied periodical motions. J. Exper. Psychol., Learn., Mem., Cogn. 17, 302–313.

Verfaillie, K., De Troy, A., Van Rensbergen, J., 1994. Transsaccadic integration of biological motion. J. Exper. Psychol., Learn., Mem., Cogn. 20, 649–670.

Wilson, M., 2001. The case for sensorimotor coding in working memory. Psychon. Bull. Rev. 8, 44–57.

Wilson, M., Knoblich, G., 2005. The case for motor involvement in perceiving conspecifics. Psychol. Bull. 131, 460–473.

Wolpert, D.M., Flanagan, J.R., 2001. Motor prediction. Curr. Biol. 11, R729–R732.

Wolpert, D.M., Doya, K., Kawato, M., 2003. A unifying computational framework for motor control and social interaction. Philos. Trans. R. Soc. Lond., B 358, 593–602.

Xie, X., Giese, M.A., 2002. Nonlinear dynamics of direction-selective recurrent neural media. Physics Review E 65, 051904. DOI:10.1103/PhysRevE.65.051904.