

# Predicting others' actions: evidence for a constant time delay in action simulation

Peggy Sparenberg · Anne Springer ·  
Wolfgang Prinz

Received: 20 April 2010 / Accepted: 14 February 2011 / Published online: 2 March 2011  
© Springer-Verlag 2011

**Abstract** Recent evidence indicates that humans can precisely predict the outcome of occluded actions. It has been suggested that these predictions arise from a mental simulation which might run in real-time. The present experiments aimed to specify the time course of this simulation process. Participants watched transiently occluded point-light actions and the temporal outcome after occlusion was manipulated. Participants were instructed to judge the temporal coherence of the action after a short (Experiment 1) and a long occlusion period (Experiment 2). Both experiments revealed a comparable negative point of subjective equality (PSE), indicating that action simulation took constantly longer than the observed action itself. Such a temporal error was not present when inverted actions were used, (Experiment 3) ruling out a pure visually driven effect. The results suggest that the temporal error is due to costs arising from a switch from action perception to an internal simulation process involving motor representations.

## Introduction

The visual environment of humans contains thousands of situations in which people are partially or completely occluded from vision for a certain time. Although the other person is not visible for a short moment, observers are able to perceive the action in a fluent manner. In fact,

observers seem to internally substitute the perceptual gap with a mental representation of the unseen parts of the action that is equivalent to the visual representation during visual perception. This internal substitution is often referred to as action simulation (Brass, Schmitt, Spengler, & Gergely, 2007; Gallese & Goldman, 1998; Graf et al., 2007; Hamilton, Wolpert, & Frith, 2004; Jeannerod, 2001; Prinz & Rapinett, 2008). It is proposed that humans use common mechanisms for executing and observing motor actions (Hommel, Musseler, Aschersleben, & Prinz, 2001; Prinz, 1990, 1997) which may enable them to understand and to predict others' ongoing behavior (Blakemore, Wolpert, & Frith, 2000; Prinz, 2006; Wilson & Knoblich, 2005; Wolpert & Flanagan, 2001). More precisely, it is suggested that humans perceive actions by simulating them using their own motor repertoire (Jeannerod, 1999, 2001).

This assumption is supported by mirror neuron research, with evidence from single-cell recording in macaque monkeys (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996) and imaging studies in humans (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Grezes, Armony, Rowe, & Passingham, 2003; Rizzolatti & Craighero, 2004) showing overlapping brain areas involved in action production and action observation. Furthermore, behavioral studies showed an influence of action observation on action execution (Brass, Bekkering, Wohlschläger, & Prinz, 2000; Brass, Bekkering, & Prinz, 2001; Kilner, Paulignan, & Blakemore, 2003; Stürmer, Aschersleben, & Prinz, 2000) and visa versa (Jacobs & Shiffrar, 2005; Reed & Farah, 1995; Reed & McGoldrick, 2007). However, although these findings provide insight into the underlying representations of action perception, little is known about the exact timing of the proposed action simulation process.

---

P. Sparenberg (✉) · A. Springer · W. Prinz  
Department of Psychology,  
Max-Planck-Institute for Human Cognitive and Brain Sciences,  
Stephanstrasse 1a, 04103 Leipzig, Germany  
e-mail: sparenberg@cbs.mpg.de

Recently, it has been shown that action simulation may run in real-time (Graf et al., 2007). Graf et al. (2007) presented point-light characters performing common actions, with the actions temporally occluded. After occlusion, a static test posture was presented. The participants had to decide whether the test posture was a continuation in the same visual angle, or whether it was rotated in depth. The authors manipulated the duration of occlusion (i.e., occluder time) and the test posture time (i.e., the time between the last seen test frame prior to occlusion and the test posture) in three time steps (100, 400 and 700 ms). The combination of each occluder time with each test posture time resulted in three different time distances of 0, 300 and 600 ms (i.e., time mismatch between occluder time and test posture time). Accordingly, occluder time and test posture time could either match (i.e., time distance of 0 ms) or mismatch (i.e., time distance of 300 and 600 ms, respectively). A time distance of 0 ms corresponds to a real-time action outcome, because the internal action representation would match the actual test posture if it was dynamically updated in real-time. Results showed that performance was best when occluder time and test posture time matched, indicating that the mental action representation corresponds to the action outcome presented. Moreover, with increasing time distance between occluder time and test posture time, error rates increased due to the increasing difference between the mental action representation and the action outcome presented. The authors took these results as evidence of a real-time process involved in action simulation.

However, another series of experiments conducted by Prinz and Rapinett (2008) investigated the timing of action simulation using a different paradigm. The authors presented video clips showing a person sitting behind a box (functioning as a spatial occluder). The person performed arm movements to displace a teapot and only the arm was seen. While the starting and the final part of the transport action were visible, the middle part was occluded by the box. The authors presented temporally correct and incorrect action outcomes after occlusion. Incorrect outcomes were either temporally earlier (resulting from an apparently faster action execution behind the occluder) or temporally later outcomes (resulting from an apparently slower action execution behind the occluder). The participant's task was to judge whether the reappearance of the action after occlusion was either too early, on time, or too late. The results showed that the participants judged action outcomes which were presented 40–120 ms later than temporally coherent outcomes were judged as being on time. This indicates that action simulation takes longer than a real-time process.

Prinz and Rapinett's (2008) finding that action simulation takes more time than the observed action itself,

does not necessarily contradict the finding of Graf et al. (2007). In fact, the temporal resolution used by Graf et al. (2007) (minimum 300 ms time distance between the occluder time and the test posture time) is too small to detect a temporal error of 40–120 ms, as was found by Prinz and Rapinett (2008). Therefore, we focused on the question of whether one can replicate the temporal error found by Prinz and Rapinett using a modified version of the paradigm of Graf et al. (2007). We conducted three experiments using temporally occluded point-light actions similar to those used by Graf and colleagues, combined with a finer temporal resolution according to Prinz & Rapinett's (2008) study. Furthermore, in contrast to Graf et al., we assessed the time course of action simulation directly by instructing the participants to judge the temporal dimension of the action outcome after occlusion.

Experiment 1 investigated the time course of action simulation using a relative short occlusion period (i.e., short occlusion time of 300 ms). Experiment 2 used a relatively long occlusion period (i.e., short occlusion time of 500 ms). Using different occluder times allowed investigating the time course of action simulation by comparing the timing of action simulation in short- and long-occluder times. Data showed a constant temporal error in Experiment 1 and 2. In order to explore the underlying representations of this temporal error, we conducted a further experiment in which we used inverted stimuli (Experiment 3). By using inverted stimuli, we can rule out the possibility of a pure visual extrapolation effect and investigate the role of motor representations in action simulation.

## Experiment 1

### Participants

Twenty right-handed participants (mean age 25 years; range 19–34 years; 10 female) completed a single session lasting about 60 min. All participants reported normal or corrected-to-normal vision and were naive with respect to the purpose of the study. They were paid for their participation.

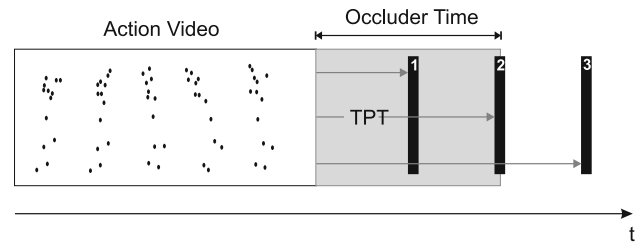
### Material

We used five video sequences showing a point-light character performing several actions (lifting something from the floor, leapfrog, pushing something away, waving with both hands and bowling). Point-light stimuli were used (instead of real video pictures) in order to isolate motion information from form information (Johansson, 1973,

1975) and to decrease social information (like sex, eye gaze, physical attractiveness, etc.). The videos were taken from a stimulus set provided by Graf et al. (2007) and showed a right-handed, male agent recorded using a motion capture system (Vicon Motion Systems Ltd., Oxford, UK) with a temporal sampling rate of 120 Hz. Each point-light display consisted of 13 black dots on a white background. Dots were located at the major joints of the actor's body (center of the head, shoulders, elbows, wrists, sternum, and center of pelvis, knees, and ankles) and were approximately 2 mm in diameter. The actions were rendered with 30 Hz (i.e., every fourth frame of the recorded sequence was shown). All actions could be easily recognized by the participants without any specific prior expertise. The point-light character was about 7 cm in height. The actions were performed within an area of 340 pixels width and 312 pixels height (about  $12 \times 11$  cm) at the center of the screen (for further details see Graf et al., 2007). An occluder of the same size was presented as a white square with a light green frame.

#### Procedure and design

A trial started with the presentation of a fixation cross (500 ms) followed by an action sequence performed by a point-light character (1,100–3,950 ms). After this action sequence, the point-light character disappeared and instead the green-framed white square (functioning as an occluder) was presented for a fixed duration of 300 ms (i.e., a temporal occluder). This occluder was followed by a static test posture out of the continuation of the action sequence. The timing within the action (i.e., the timepoint in the test posture) was manipulated in 13 time steps. These test posture times (TPT) referred to the time between the last test frame seen prior to occlusion and the static test posture (i.e., the time which would go on by without any occluder) and ranged from 100 ms after occlusion onset to 500 ms in 33 ms steps (i.e., 100, 133, 167, 200, 233, 267, 300, 333, 367, 400, 433, 467 and 500 ms). Taking into account occluder duration of 300 ms, a real-time condition was included (i.e., both occluder time and TPT lasted 300 ms) as well as temporally incorrect conditions. Early test postures referred to the test posture time shorter than the presented occluder time (TPT < 300 ms); late test postures referred to test posture times longer than the presented occluder time (TPT > 300 ms) (Fig. 1). Participants were asked whether the static test posture, which reappeared after occlusion, was temporally earlier or later than the actual action outcome that would be expected after occlusion (forced choice). Participants were asked to immediately respond when the static test posture appeared and a time out for their response was set for 2,000 ms (time out



**Fig. 1** Schema of a trial: Each trial started with a fixation cross (not depicted) which was followed by an action sequence. After that an occluder was presented. Occluder time was fixed on 300 and 500 ms, respectively, followed by a static test posture. The factor test posture time (TPT) was defined as the time between the last seen test frame prior to the occluder and the one presented as the static test posture. The TPT was either (i) shorter than the occluder time (1 negative values), or (ii) longer than the occluder time (3 positive values), or (iii) it could correspond to the occluder time (2 value of 0 ms; represents the actual real-time action outcome)

trials were excluded from data analysis). Responses were given via foot pedals.<sup>1</sup>

It should be noted that, in contrast to Prinz and Rapinett (2008), in this study the timepoint of the presentation of the test posture was fixed (i.e., fixed occlusion duration) and we presented action outcomes which represented a different timing within the action (i.e., variable test posture time). We measured the test posture time, which is judged as the correct continuation after the fixed temporal occlusion. In contrast, Prinz and Rapinett (2008) presented one and the same action outcome (i.e., fixed test posture time) and they manipulated the timepoint of test posture presentation (i.e., variable occlusion duration). They measured the timepoint, which was judged as the correct continuation after a fixed spatial occlusion. Therefore, in our experiments, the judgment of earlier action outcomes as the correct continuation indicates that action simulation takes longer than the observed action (in contrast to the study by Prinz and Rapinett (2008)).

The experiment consisted of 390 trials (5 actions  $\times$  13 test posture times  $\times$  6 repetitions) divided into ten blocks. The factor test posture time was completely randomized. Prior to the experiment, the participants received an initial familiarization phase where all actions were presented twice. This was followed by a practice phase containing different actions than in the experiment (knee-bends, throwing a ball with both hands, throwing a basketball). The practice phase consisted of 45 trials and did not include a real-time condition. This was because the participants received feedback on the forced choice task (“earlier” or “later”) and there was no correct response

<sup>1</sup> We decided on a foot response because follow-up studies are planned in which the presented paradigm is combined with secondary motor tasks involving both hands. Earlier studies in our laboratory gave no indication that foot responses differed from hand responses in either error rates or RTs.

alternative in the real-time condition. An experimental session lasted about 1 h. No feedback was given to the participants during the experimental phase.

The data analyses focused on the response distribution (i.e., percentage of “later” judgments) across different TPTs. TPTs are reported relative to the occluder time. Negative values represent TPTs shorter than the occluder time presented and positive values TPTs longer than the occluder time presented. The proportion of “later” responses was analyzed by using an iterative least-squares fitting to an exponential logistic function. This allows the analysis of two parameters of the individual functions: the point of subjective equality (PSE) and the just noticeable difference (JND) (Gescheider, 1997). The PSE is the point at which the function yields a probability of 0.5 (i.e., a TPT is judged at chance level and it refers to the subjective real-time outcome (i.e., the presented action outcome at this point in time matches the dynamically updated mental action representation of the participant). The JND is defined as half the difference between the TPT, which is classified as “later” on 25% and the TPT, which is classified as “later” on 75% of the trials. It is a measure of the steepness of the function and indicates the smallest possible physical difference that can be detected reliably.

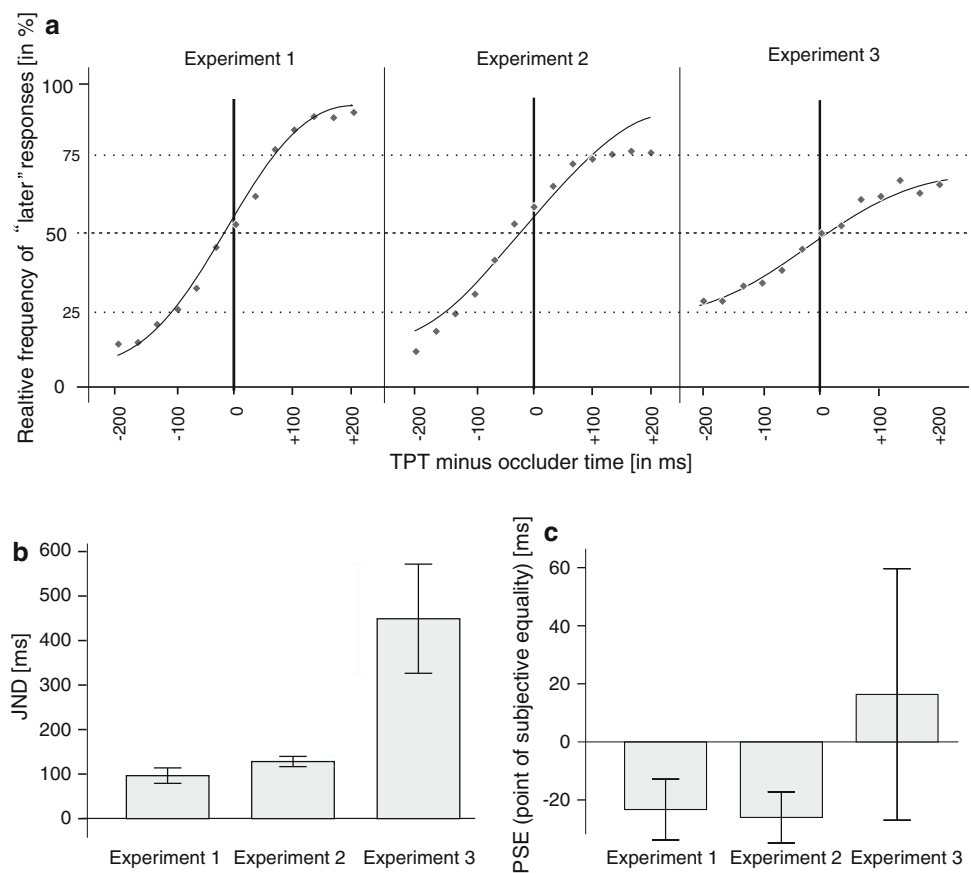
If a real-time simulation process takes place, this would imply that the participant’s mental representation corresponds to an actual real-time outcome after occlusion (due to the assumed internal real-time update of the perceived action during occlusion). In this case, we should find a PSE of zero (TPT = 0 ms), because the participant’s mental action representation corresponds to the test posture times presented.

On the other hand, if the action simulation process takes longer than the simulated action itself, we should find a negative PSE. A negative PSE would imply that the participants perceive earlier action outcomes as real time, because their mental action representation is behind relative to a real-time outcome.

## Results and discussion

First, we analyzed the overall response behavior across the 13 different TPTs in a one-way ANOVA for repeated measurements. Data showed a significant effect of TPT ( $F(12, 228) = 146.227$ ;  $p < .001$ ;  $\eta^2 = .885$ ), but more importantly, the linear trend analysis was reliable ( $F(1, 19) = 275.334$ ;  $p < .001$ ;  $\eta^2 = .935$ ). As expected, the overall response behavior showed a significant linear increase in “later” responses with increasing TPT (Fig. 2a,

**Fig. 2** Results of the Experiments: **a** proportion of “later” responses dependant on the TPT value presented (depicted as the deviance between the occluder time minus the test posture time) for each single experiment. Bars represent standard error of the mean. The slope of the TPT pattern decreased from Experiment 1 to 3, indicating a decreasing task performance. **b** JND for the separate experiments: **c** PSE for the separate experiments



Experiment 1). This confirms that the participants could solve the task although our paradigm combined a direct measurement of the time course of action simulation with a relatively high temporal resolution in contrast to previous work (Graf et al., 2007).

Second, we found a negative PSE of  $-23$  ms, which differed significantly from zero ( $p < .05$ ). This supports the assumption that the action simulation process takes longer than the simulated action itself (Fig. 2c). The JND was 97 ms (Fig. 2b).

Finally, for the reaction times, the ANOVA across the 13 different TPTs showed a significant effect of TPT ( $F(12, 228) = 10.781; p < .001; \eta^2 = .362$ ). More importantly, the quadratic trend test was significant ( $F(1, 19) = 29.591; p < .001; \eta^2 = .609$ ) (Table 1). RTs were significantly higher in the real-time condition as compared to the “later” conditions ( $p < .01$ ; Bonferroni corrected). The increased RTs in the real-time condition could be an indication of participants’ increased uncertainty.

Overall, we found evidence that action simulation takes more time than the simulated action itself in relatively short occlusion conditions (i.e., occluder time of 300 ms). In extension to earlier studies (Prinz & Rapinett, 2008), we applied an iterative fitting on the individual raw data, which allows us to exactly estimate the size of the temporal error.

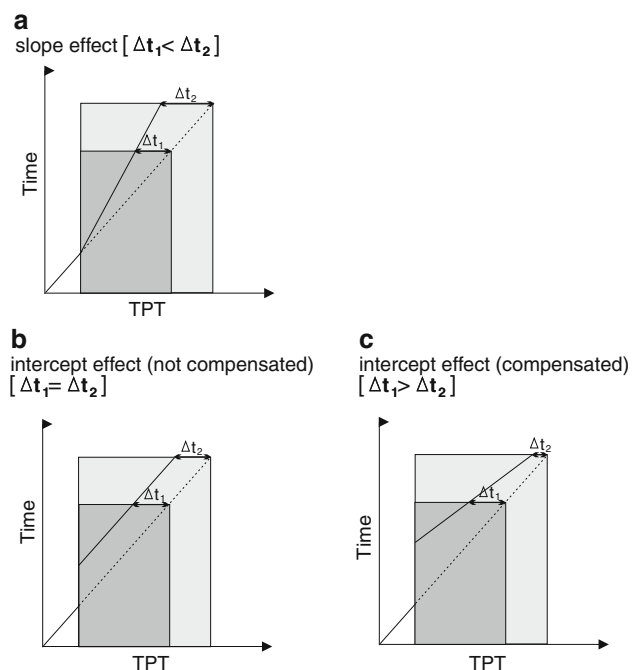
However, the results do not answer the question of how this temporal error arises. In accordance with Prinz and Rapinett, we considered two sources for the temporal error: the slope and the intercept of a linear extrapolation of the occluded action part. A slope effect would imply that action simulation is in itself slower than the observed action. In contrast, an intercept effect would imply that

action simulation involves a constant time error. In this case two possibilities can be considered: either action simulation runs in real-time with the constant temporal error or the initial temporal error is compensated by an internal simulation process running faster than the real action.

In order to test these alternatives, Experiment 2 used an occlusion period of 500 ms instead of 300 ms as in Experiment 1. If the temporal error reflects a slope effect, the temporal error should be larger (i.e., more negative PSE) for long occluder durations (Experiment 2) than short occlusion durations (Experiment 1). In contrast, if the temporal error reflects an intercept effect with action simulation running in real-time with the constant temporal error, the temporal errors should be comparable between short and long occlusion durations. On the other hand, if the initial temporal error arises from an intercept effect which is compensated by an action simulation process running faster than the real action, we should find a larger temporal error in short occluder durations (Experiment 1) as compared to a long occlusion durations (Experiment 2), because in the long occlusion period more time is available for compensating the initial temporal error (Fig. 3).

**Table 1** Mean reaction times and standard deviation (parenthesized) for each experiment

TPT relative to occluder time	Experiment 1	Experiment 2	Experiment 3
-200	821 (181)	870 (176)	931 (154)
-167	842 (187)	902 (191)	958 (169)
-133	859 (192)	903 (173)	959 (174)
-100	858 (216)	925 (179)	940 (178)
-67	890 (209)	935 (167)	978 (163)
-33	921 (203)	929 (186)	999 (166)
0	909 (189)	931 (195)	956 (178)
+33	897 (180)	937 (178)	986 (164)
+67	876 (199)	903 (175)	974 (177)
+100	826 (180)	891 (177)	965 (180)
+133	795 (166)	895 (188)	942 (175)
+167	789 (166)	884 (158)	932 (173)
+200	787 (145)	873 (181)	949 (139)



**Fig. 3** Possible sources of the negative temporal error and different predictions for the temporal error ( $\Delta t$ ) in short (dark gray bar) and long occlusion durations (light gray bar): slope (a) and intercept assumption (b + c). If the temporal error arises from an intercept effect, the error can either remain stable (b) or decrease with ongoing occluder duration (c)

## Experiment 2

### Participants

Twenty-two right-handed participants (mean age 26 years; range 20–33 years; 10 female) completed a single session lasting about 60 min. Two participants were excluded from data analysis because they did not follow the instructions. Thus, data analysis was based on a total number of 20 participants. All participants reported normal or corrected-to-normal vision and were naive with respect to the purpose of the study. They were paid for their participation.

### Material

The same stimuli were used as in Experiment 1.

### Procedure and design

Trials were similar to the trials in Experiment 1, except for the fact that the occluder time was fixed at 500 ms. Accordingly, the test posture time was manipulated in the following 13 time steps: 300, 333, 367, 400, 433, 467, 500, 533, 567, 600, 633, 667 and 700 ms). As in Experiment 1, a real-time condition was included (i.e., both occluder time and test posture time lasted 500 ms). Procedure and data analysis were the same as in Experiment 1.

### Results and discussion

First, overall performance was significantly lower in Experiment 2 as compared to Experiment 1 (72% correct responses and 79% correct responses, respectively) ( $p < .05$ ). This decreased performance can be explained by the longer occlusion period, because the visual system receives no visual input for a longer time. Nevertheless, task performance in Experiment 1 and 2 differed significantly from chance level ( $p < .001$ ).

Second, an analysis of the overall response behavior across the 13 different TPTs showed, as expected, a significant main effect of TPT ( $F(12, 228) = 129.142$ ;  $p < .001$ ;  $\eta^2 = .872$ ) and, more importantly, the linear trend of “later” responses with increasing test posture time was reliable ( $F(1, 19) = 358.814$ ;  $p < .001$ ;  $\eta^2 = .950$ ) (Fig. 2a, Experiment 2).

As in Experiment 1, the PSE and the JND was analyzed. We found a significant negative PSE of  $-26$  ms ( $p < .05$ ) which supports the assumption that action simulation takes longer than the simulated action itself during relatively long occlusion periods (Fig. 2c). A comparison between the PSEs in Experiment 1 and 2 showed no significant difference between both ( $p > .80$ ) (Fig. 2c). This argues for the idea that the temporal error is caused by a stable

intercept effect. The JND was 128 ms. A comparison between the JNDs in Experiment 1 and 2 showed no significant difference between both ( $p > .40$ ).

In line with Experiment 1, RTs showed a significant effect of TPT ( $F(12, 228) = 3.263$ ;  $p < .001$ ;  $\eta^2 = .147$ ). More importantly, the quadratic trend test was significant ( $F(1, 19) = 17.838$ ;  $p < .01$ ;  $\eta^2 = .484$ ) (Table 1). Again, RT increase in the real-time condition could be an indication of increased uncertainty in the participants.

Taken together, these findings replicate those from Experiment 1. The PSE was negative and differed significantly from zero, which supports the assumption that action simulation takes longer than the simulated action. Apart from this, the size of the PSE did not differ from those in Experiment 1, which supports the idea that the temporal error arises from an intercept effect (rather than from a slope effect). Apart from this, the data speak for a constant temporal error, which is not compensated over the course of the occlusion period (as suggested by Prinz and Rapinett, 2008).

## Experiment 3

Experiment 3 aimed to investigate how stable the observed temporal error is. Furthermore, it was intended to explore the representations involved in the action prediction task. Although Experiment 1 and 2 provide an insight into the time course of action simulation, we cannot make a claim about the underlying representations. It has been suggested that humans use not only visual representations during action observation, but also motor representations in order to understand and anticipate the actions observed in others (cf. “Introduction”). Accordingly, we considered that the prediction of visually occluded actions (in extension to direct action observation) does also involve motor representations. More precisely, Experiment 3 aimed to explore whether the temporal error found in Experiment 1 and 2 reflects a pure visual effect or whether it is related to the involvement of motor representations. This was tested by using an inverted version of our point-light stimuli, which are known to impair action perception (Loula, Prasad, Harber, & Shiffrar, 2005; Pavlova & Sokolov, 2000; Troje, 2003), although low-level visual input remains the same as in Experiment 1 and 2 in complexity and movement speed. Humans do have no/little visual expertise and no motor expertise with inverted actions (Loula et al., 2005) and it has been argued that action perception and the activation of motor representations might be highly related to motor expertise (Calvo-Merino et al., 2005; Casile & Giese, 2006). Theoretically, it has been proposed that inverted actions decrease the involvement of motor representations, because observers have no motor experience with inverted actions and humans are not able to map the observation of

inverted actions onto their motor repertoire (for further details see Graf et al., 2007, “Experiment 3”). Accordingly, if action simulation involves motor representations, performance in the prediction task should be impaired when inverted actions are presented. Furthermore, if the temporal error is related to motor representations, this temporal error should differ between inverted and uninverted actions.

Alternatively, one can argue that the task may be solved by a pure visual extrapolation process, meaning that the temporal error is a pure visual effect. Pure visual extrapolation is considered to rely on local motion cues and velocity (e.g., tracking of single dots). Such an effect would be present every time when a stimulus arrangement is occluded and needs to be extrapolated in space over a certain time period. If so, we should find no difference in the time course between inverted and uninverted point-light actions, because visual extrapolation should be applied equally to the inverted and uninverted actions.

### Participants

Twenty right-handed participants (mean age 24 years; range 20–31 years; 10 female) completed a single session lasting about 60 min. All participants reported normal or corrected-to-normal vision and were naive with respect to the purpose of the study. They were paid for their participation.

### Material

The same stimuli were used as in Experiment 1 and 2, except that the actions were presented upside down, with all actions and the static test postures flipped around the horizontal axis. All actions could be easily recognized by the participants although they were inverted. Familiarization and practice phase involved only inverted actions.

### Procedure and design

Trials were identical to the trials in Experiment 2, except for the fact that the actions were inverted. Again, occluder time was fixed at 500 ms and the test posture time was manipulated in the same 13 time steps (300, 333, 367, 400, 433, 467, 500, 533, 567, 600, 633, 667 and 700 ms).

If action simulation involves motor representations, we expect to find a decreased task performance in inverted as compared to uninverted actions and a smaller or even no significant TPT effect. Furthermore, if the temporal error is related to motor representations, we expect no significant negative (or positive) PSE.

Contrary, if the task is solved by a visual extrapolation process and the temporal error is a pure visual effect, we should find similar results in Experiment 3 as were found in

Experiment 1 and 2 (i.e., no difference in the general task performance, the TPT pattern and no difference in the PSE), because visual extrapolation can be performed equally in inverted and uninverted actions.

### Results and discussion

First, overall performance was not significantly lower in Experiment 3 compared to Experiment 2 (67% correct responses and 72% correct responses, respectively) ( $p = .054$ ), although a clear trend was present. Nevertheless, the task performance in Experiment 3 differed significantly from chance level ( $p < .001$ ), indicating that our participants understood and were able to complete the task although the stimuli were presented in an inverted manner.

Second, we found a significant effect of TPT ( $F(12, 228) = 29.472$ ;  $p < .001$ ;  $\eta^2 = .608$ ). More importantly, the linear trend analysis was reliable ( $F(1, 19) = 40.023$ ;  $p < .001$ ;  $\eta^2 = .678$ ), but it was much smaller than the TPT effect in Experiment 2 (Fig. 2a, Experiment 3). This was shown by a decreased  $\eta^2$  value (Experiment 2: 0.950; Experiment 3: 0.687) and by a significant TPT (13 TPTs)  $\times$  Experiment (Experiment 2, 3) interaction ( $F(12, 456) = 2.813$ ;  $p < .05$ ). A flatter increase in “later” responses with increasing TPT was found in Experiment 3. This indicated that the participants were less precise in judging a TPT as being earlier or later. This argues against the assumption that the task is solved by a pure visual extrapolation process because such a process can be applied to both uninverted and inverted actions equally.

The PSE in Experiment 3 was +16 ms and did not differ significantly from zero ( $p > .70$ ) (Fig. 2c). Although this PSE differed descriptively from the PSE in Experiment 2 (–26 ms), there was no significant difference between both ( $p > .30$ ). This is due to high variability between the participants in Experiment 3 (Fig. 2b).

The JND was 449 ms and a comparison between the JND in Experiment 2 and the JND in Experiment 3 showed a significant difference ( $p < .05$ ), which supports the assumption that the simulation of inverted actions is impaired relative to the simulation of uninverted actions.

As in Experiments 1 and 2, RTs showed a significant effect of TPT ( $F(12, 228) = 1.915$ ;  $p < .001$ ;  $\eta^2 = .092$ ). Again, the quadratic trend test was significant ( $F(1, 19) = 8.869$ ;  $p < .01$ ;  $\eta^2 = .318$ ) (Table 1). As in Experiments 1 and 2, the RT increase in the real-time condition could be an indication of participants’ increased uncertainty.

### General discussion

The present three studies aimed to investigate and quantify the time course of the mental simulation of temporally

occluded actions and to explore the underlying representations in this process. Our studies provide three major findings: First, data showed a significant negative point of subjective equality (PSE), indicating that action simulation took constantly longer than the observed action itself, supporting previous results (Prinz & Rapinett, 2008). Importantly, in contrast to previous findings, our results show that the size of temporal error can be exactly estimated and the size of the temporal error is 25 ms in average.

Second, comparable PSEs for short and long occlusion periods (Experiment 1:  $-26$  ms; Experiment 2:  $-23$  ms) were found. This supports the assumption that the temporal error arises from a stable intercept effect (rather than from a slope effect) that is not compensated by a simulation process which is faster than the observed action. The observed stable temporal error may reflect the cost of a switch from an action perception process to an action simulation process. Switching costs are stable and do not change with different occluder times.

Finally, results indicate that the prediction of inverted actions is different to the simulation of uninverted actions. The overall task performance was impaired by the inversion of the stimuli and the JND was significantly increased in inverted as compared to uninverted actions, which indicates that the participants were less sensitive to predict inverted actions. This excludes the possibility that our task is solved only by a simple visual extrapolation process, because pure visual extrapolation (i.e., tracking single dots) can be applied to both uninverted and inverted actions equally. The decrease in task performance when predicting inverted action sequences shows that pure visual extrapolation is not sufficient for solving the task. No significant PSE was present in Experiment 3 while this was the case in Experiments 1 and 2. Furthermore, the results showed a high interindividual variability. This indicates that the timing of action prediction was less systematically in inverted actions as compared to uninverted actions (Experiments 1 and 2). This finding supports the assumption that humans use not only visual representations for action simulation, but that action perception might highly involve motor representations (Calvo-Merino et al., 2005; Casile & Giese, 2006). We consider that inverted point-light stimuli should decrease the involvement of motor representations and that performance in internal action simulation should be impaired for inverted stimuli. This was exactly what our data showed. Moreover, the fact, that no significant PSE was present in inverted actions excludes the possibility that the temporal error is a pure visual effect, which should be always present when stimulus arrangements are extrapolated in space over a certain time period.

Although we can rule out that the temporal error is a pure visual one and although we suggest that the temporal

error is related to motor representations, we are aware of the fact that inverted actions confound a lack of both visual and motor expertise. It has been shown that motor-related and visual processes are disturbed in inverted actions (Loula et al., 2005). One might argue that the accuracy of visual extrapolation depends on the degree of visual expertise with the presented action and that the temporal error might be related to visual expertise. Although this is a plausible argument, a follow-up study of our labor showed no statistical reliable effect of visual expertise on the accuracy to extrapolate action sequences into the future. In that pilot study, participants were extensively familiarized with a pool of inverted action sequences (i.e., they watched each action sequence 50 times), while a pool of other inverted action sequences was not presented before. After that they performed exactly the same experiment as described in Experiment 3. The data showed that both the visually unfamiliar and visually familiar inverted action sequences showed a comparable PSE. This PSE was not negative, replicating the finding of Experiment 3. This finding argues against the possibility that visual expertise affects the temporal error. In addition to this first evidence, we suggest that further research needs to be done in order to further specify whether processes within the visual and/or the motor system are responsible for the temporal error.

The present data showing that action simulation takes longer than the simulated action itself seem to be counterintuitive to a predictive concept of action simulation (Blakemore et al., 2000; Blakemore & Frith, 2005; Graf et al., 2007; Wolpert & Flanagan, 2001). It has been proposed that the perception of another person's action leads to the anticipation of the progress of this action by an internal imitation of the observed action and an internal evocation of possible goals and intentions (Gallese & Goldman, 1998; Jeannerod, 2001). However, a second view of our data does not exclude the anticipation of the progress of an observed action. It is reasonable to assume that anticipation occurs because people skip the simulation of unimportant action parts and focus on the main markers of an action (e.g., the goal) and anticipate them. In our paradigm, we used action sequences where participants had to simulate the action in detail (including unimportant action parts). Furthermore, participants could not anticipate the occurrence of the occluder (i.e., they could not anticipate the switch between action perception and action simulation), which is a highly artificial situation. This could cause costs which manifest themselves as a slight temporal error in action simulation.

Finally, we would like to point out that the temporal error found (PSE of  $-26$  and  $-23$  ms, respectively) is small and still close to real-time. In a broader frame, this leads to the more functional question of how exact action simulation actually needs to be in everyday (inter-)active



settings. However, it should be noted that this functional aspect is not the focal point of our research. Rather, our focus is on tackling the underlying processes that enable humans to simulate unseen parts of others' actions. As the present findings indicate, action simulation can take longer as the simulated action itself under certain conditions. Addressing whether this is, for instance, fostered by internal feedback processes remains one of the interesting questions that needs to be addressed in future work.

**Acknowledgments** We thank Wiebke Berger for assistance in data acquisition and Mathias Lesche for support in programming. We are grateful to Bernhard Hommel and an anonymous reviewer for helpful comments on an earlier version of this manuscript.

## References

- Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, *43*(2), 260–267.
- Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, *11*(11), R11–R16.
- Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cognition*, *44*(2), 124–143.
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22.
- Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current Biology*, *17*(24), 2117–2121.
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cerebral Cortex*, *15*(8), 1243–1249.
- Casile, A., & Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology*, *16*(1), 69–74.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*(6), 2608–2611.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, *2*(12), 493–501.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(Pt 2), 593–609.
- Gescheider, G. A. (1997). *Psychophysics: the Fundamentals*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *Neuroimage*, *36*(Suppl 2), T22–T32.
- Grezes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to “mirror” and “canonical” neurones in the human brain: an fMRI study. *Neuroimage*, *18*(4), 928–937.
- Hamilton, A. F. D., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology*, *14*(6), 493–498.
- Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849–878. discussion 878–937.
- Jacobs, A., & Shiffrar, M. (2005). Walking perception by walking observers. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 157–169.
- Jeannerod, M. (1999). To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, *52*(1), 1–29.
- Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, *14*(1 Pt 2), S103–S109.
- Johansson, G. (1973). Visual-Perception of Biological Motion and a Model for Its Analysis. *Perception & Psychophysics*, *14*(2), 201–211.
- Johansson, G. (1975). Visual motion perception. *Scientific American*, *232*(6), 76–88.
- Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*(6), 522–525.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(1), 210–220.
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*(5), 889–899.
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action: Current approaches* (pp. 167–201). New York: Springer.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154.
- Prinz, W. (2006). What re-enactment earns us. *Cortex*, *42*(4), 515–517.
- Prinz, W., & Rapinett, G. (2008). Filling the Gap: Dynamic Representation of Occluded Action. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions* (pp. 223–236). Amsterdam: IOS Press.
- Reed, C. L., & Farah, M. J. (1995). The psychological reality of the body schema: a test with normal participants. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(2), 334–343.
- Reed, C. L., & McGoldrick, J. E. (2007). Action during body perception: processing time affects self-other correspondences. *Social Neuroscience*, *2*(2), 134–149.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1746–1759.
- Troje, N. F. (2003). Reference frames for orientation anisotropies in face recognition and biological-motion perception. *Perception*, *32*(2), 201–210.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460–473.
- Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732.