Performing hand actions assists the visual discrimination of similar hand postures

R.C. Miall [a,b,*], J. Stanley [a], S. Todhunter [b], C. Levick [b], S. Lindo [a], J.D. Miall [b,c]

[a] *Behavioural Brain Sciences, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
[b] *University Laboratory of Physiology, Oxford OX1 3PT, UK*
[c] *Cherwell School, Oxford OX2 7EE, UK*

Abstract

Recent theoretical work has suggested that internal predictive signals are used for motor control and coordination. The predictive signal – proposed to be the output of a forward model – would be a sensory representation of action. Hence, these sensory representations could potentially influence other sensory processes. We report here how performance of hand actions assisted the visual discrimination of target hand postures presented at random times within an on-going series of hand images. Reaction times to discriminate the targets were significantly shorter when the displayed images were both sequential and congruent with the action being performed. Hence, the planning or execution of action appears to allow better prediction of a displayed series of congruent images. In further control experiments, we show that the motor–visual priming effect is unlikely to be due to differential attentional demands and it is specific to a first person perspective display; it is short lasting, being lost if a 500 ms delay is introduced between successive stimulus presentations. The data are interpreted as evidence supporting the hypothesis that forward models in the motor system provide action-specific sensory predictions that are available to cognitive processes.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Forward model; Prediction; Motor planning; Visual discrimination; Reaction time

1. Introduction

The concept that the nervous system internally models the behaviour of the motor system has gained increasing prominence over recent years (Jordan, 1995; Kawato, Furukawa, & Suzuki, 1987; Kawato, 1999). Within this concept, forward models capture the forward or causal relationship between actions and the resultant change in the state of the motor system (Jordan & Rumelhart, 1992). These models estimate the next sensory state of the motor system based upon information about its current state, its dynamics and the motor command being issued to it.

Forward models can support sensorimotor control in many ways, including sensory confirmation, internal feedback, context estimation and state estimation (Haruno, Wolpert, & Kawato, 2001; Jordan et al., 1992; Miall & Wolpert, 1996; Wolpert, Ghahramani, & Jordan, 1995). There is also evidence that we distinguish the sensory consequences of our own actions from externally produced stimuli using a forward model (Wolpert et al., 1995; Wolpert & Ghahramani, 2000). Reafferent, self-induced sensations can therefore be cancelled or attenuated during movement, thus highlighting exafferent information critical for control.

It has been proposed that the process used by the forward model to predict the sensory consequences of one's own movements could also be available to cognitive processes (Decety & Grezes, 1999; Frith, Blakemore, & Wolpert, 2000; Jeannerod & Decety, 1995). Forward models providing sensory predictions within the motor system could provide a visual signal, or similar representation, which would assist in vision-based cognitive tasks. Psychophysical studies have suggested that internal representation of action is indeed used to solve motor-related tasks. When asked to judge the laterality of visually presented hands, subjects mentally rotate their own hand into the stimulus orientation for comparison (Gentilucci, Daprati, & Gangitano, 1998; Parsons, 1994; Parsons et al., 1995). Similarly, when subjects are asked to estimate the feasibility of grasping objects placed at different orientations, there is a correspondence between response time and the time taken to actually reach and grasp an object placed at the same orientation, suggesting that the subjects mentally move into the appropriate position in order to generate an

internal representation of the arm that then supports the decision process (Frak, Paulignan, & Jeannerod, 2001). However, these experiments cannot distinguish whether the mental rehearsal of the action is undertaken to generate a sensory representation, in order to solve the cognitive task, or whether it is undertaken to evoke a motor intention or plan, which is used without a specific sensory representation.

Recently, several visual cueing experiments (Brass, Bekkering, & Prinz, 2001; Brass, Bekkering, Wohlschlager, & Prinz, 2000; Craighero, Bello, Fadiga, & Rizzolatti, 2002; Craighero, Fadiga, Rizzolatti, & Umilta, 1999) have demonstrated that visual images can prime the motor system and thus lead to faster actions when the cue and the action are congruent (visuo-motor priming). Craighero et al. (1999, 2002) showed that initiation of a pre-specified reach to grasp hand action can be modulated by prior viewing of pictures of a hand that matched or did not match the planned hand orientation. They argued (Craighero et al., 2002) that the reduced response times were consistent with *motor–visual* priming, such that the motor plan led to more rapid visual processing of the cue to move. Unfortunately, as they acknowledged (p. 498), they could not eliminate the possibility that their results were in fact due to *visuo-motor* priming, and that the congruence of the cue image resulted in more rapid initiation of the planned action than did incongruent cues. We interpret their data as an effect of visuo-motor priming. More recently, Hamilton, Wolpert, and Frith (2004) showed that performing an action influenced judgement of observed actions, but found that the effect was one of interference: holding a heavy weight biased subjects to report that weights they observed being lifted were lighter than they really were. They suggest the forward model activated during the action cannot simultaneously be used for the perceptual task, and so the subjects' reports are biased away from their performed action. Other studies have tested the effects of action on perception, under the theory of event coding (Hommel, Musseler, Aschersleben, & Prinz, 2001), but without a specific, direct relationship between visual outcome of action and the perceived image—for example, Musseler and co-workers (Musseler & Hommel, 1997; Musseler, Steininger, & Wuhr, 2001) presented an arrowhead whose orientation reflects which of two buttons are struck, but the arrowhead has only an arbitrary relationship to finger shape or position. Actions can also influence spatial or temporal judgments about visual and somatosensory stimuli (Ross, Morrone, Goldberg, & Burr, 2001; Smith, Rorden, & Jackson, 2004; Yarrow, Haggard, Heal, Brown, & Rothwell, 2001). However, it is not clear that these phenomena are based on specific sensory predictions of the outcome of the planned action, rather than a more general remapping of the spatio-temporal relationships of any sensory stimuli (Duhamel, Colby, & Goldberg, 1992) around the time of action.

Hence, we do not know of any reports that have demonstrated that motor preparation or performance can selectively facilitate processes within the visual system, as might be expected from a forward model prediction of the action. We report evidence of this effect in a human visual discrimination task, in five related experiments. Visual stimuli comprised of images of a hand presented in a slow sequence on a computer screen. Oddball target images within the sequence showed a different hand posture, and the subjects vocally responded to these oddball images. During the task, subjects made slow hand actions that were congruent or incongruent with the displayed image series. We hypothesise that prediction of the observed sequence of images would allow more rapid discrimination of the oddball targets. Hence, we expect that if the observed image sequence is predictable, discrimination will be faster than if it is random. If the prediction of the observed sequence is facilitated by performance of a congruent action, then we expect yet faster discrimination. In contrast, if Hamilton et al. (2004) are correct, then the opposite effect should be seen: performance of congruent actions should impede discrimination.

2. Experiment 1: Congruency and prediction

2.1. Methods

Twenty-eight subjects (9 female, 19 male) participated in the first experiment, after giving informed consent; the experiments were approved by the Central Oxfordshire Research Ethics Committee. To motivate them, each subject received £5 payment; subjects with mean reaction times that were in the top third of the distribution received an additional £5 reward. The mean age of the subject group was 19.4 years, ranging from 18 to 22 years. All were right handed, had normal or corrected to normal vision and were naïve to the purpose of the experiment.

The experiment took place in a sound attenuated, dimly lit room. Participants sat at a table in front of a 44 cm computer monitor with a viewing distance of 60 cm; they wore a microphone headset and held their left hand either above their left leg, or rested it on the table (Fig. 1A), in the periphery of their vision.

The task consisted of discriminating a target image randomly presented within a sequence of static images (Fig. 2), vocally responding '*ta*' as fast as possible into a microphone that was connected to a voice-activated switch. All images were static views of a human left hand, rendered as 450×450 pixel bitmaps with the Poser Pro 4 animation package (Curious Labs Inc.), and displayed at a rate of one image per second using Presentation (Neurobehavioral Systems, USA). Two main sequences showed the hand in 15 positions ranging from a closed fist to full extension of the fingers (the '*fist*' sequence, Fig. 2), or from the open hand in pronation to supination (the '*flip*' sequence). Each image was presented to the left of a fixation pointer, in a panel 8 cm square, occupying approximately 2–10 degrees left of fixation.

Each trial consisted of 200 image presentations from the main sequence, either selected randomly from the 15 images, or cycling backwards and forwards through the sequence, selecting every fourth image in order. Thus, the displayed series of images was either unpredictable or showed a repeated hand action at a rate of two complete cycles every 15 s. Each image was presented for 1 s, without gaps; we chose to use slow presentation of static images so that the random sequence could be used.

Two oddball target images showed a hand with two extended fingers ('*peace*', used for the *fist* series, Fig. 2) or with opposed index finger and thumb (the '*OK*' signal, used for the *flip* series). These images were chosen to have some similarity with the
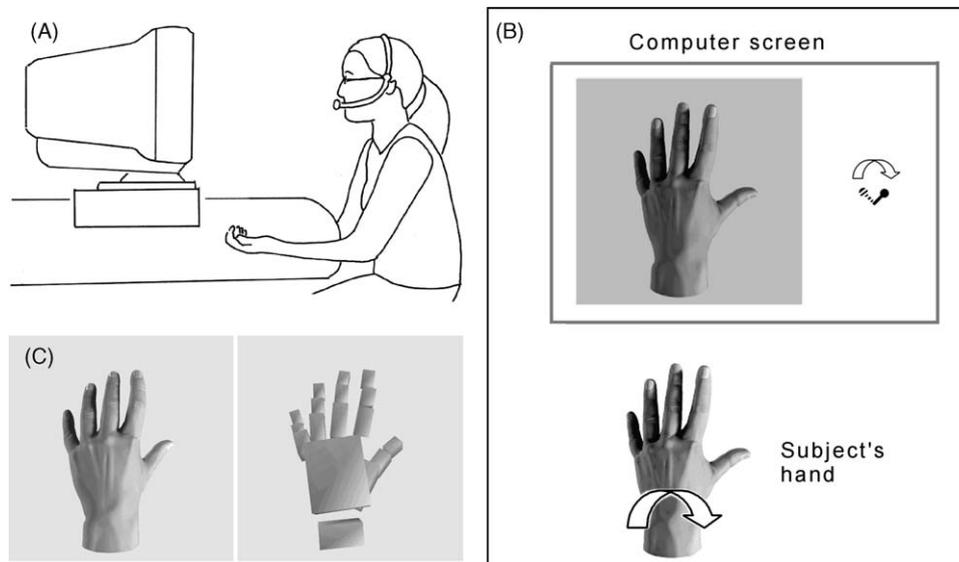
Fig. 1. An experimental setup: (A) subject is shown wearing the microphone with her left hand in the starting position. (B) Image and fixation point (right hand small black "lollipop" icon) are shown at the top, as seen on the computer monitor; hence, the image of the hand is presented to the left visual hemifield; the participant's left hand is shown in the foreground (bottom). The oscillating motion of fixation pointer was used to cue and synchronize the hand movements of the subject. (C) Examples of the realistic and box-rendered images used in Experiment 5.
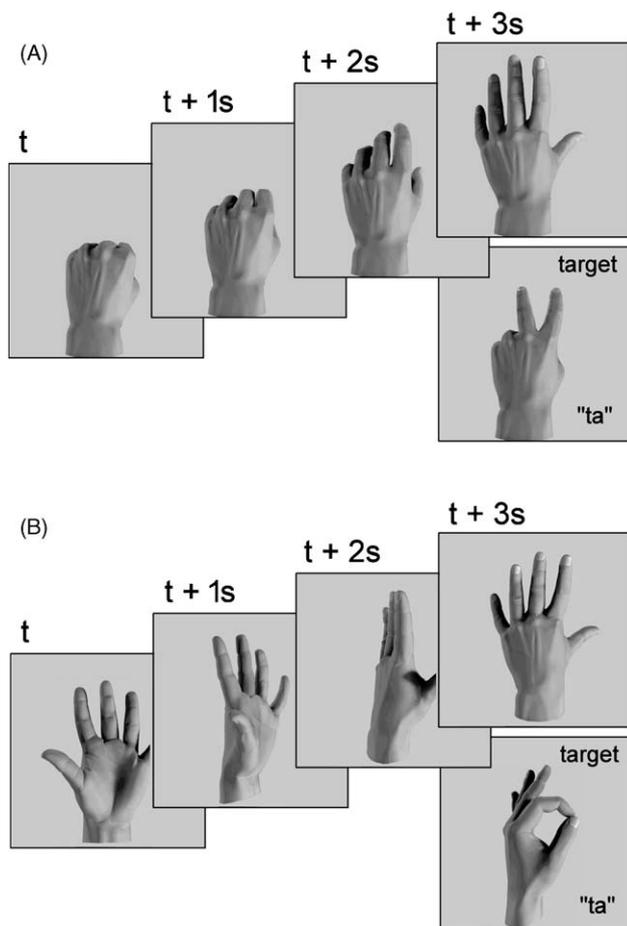

Fig. 2. Sequences of stimuli in Experiment 1: (A) images from the '*fist*' series were presented each second. On appearance of the randomly presented '*peace*' target image (bottom right), participants were to say '*ta*' into a microphone. The target image was randomly inserted into the main sequence. (B) Images from the '*flip*' series and the '*okay*' target image (right).

hand postures of the main series, so that subjects could not discriminate the targets on the basis of the overall image size or orientation, for example. The target image was inserted at infrequent random points within the main series of presentations, replacing the images at those points. Oddball targets were never presented within the first 10 s (10 stimulus presentations) of a trial; the inter-target interval was a minimum of 5 s; and otherwise their positions in the sequence were chosen with a random number generator with a probability of 0.1 for each sequence position. On average, targets were presented 8.5 times in the series of 200 images.

Reaction times (RT) to the target were measured from the subject's vocal responses. This avoided the inter-manual interference that had been detected in pilot experiments in which subjects reacted by clicking a computer mouse button. A voice-activated switch was used to generate a TTL pulse that was detected by the Presentation program. Prior to the start of the experiment, each subject practiced responding sufficiently loudly to activate the switch.

While discriminating these images, subjects also performed either the *fist* or *flip* action with their left hand. The images displayed on screen were therefore from actions that were the same as those being performed (*congruent* condition, both displayed images and performed action were *flip*, for example) or different (*incongruent* condition, performing *flip* but observing *fist*, or vice versa). Note that we use the term congruent to indicate the match between image set and action: in the random condition there is no temporal match between the performed action and observed images. A small lollipop-shaped fixation pointer was displayed 4 cm to the right of the stimulus panel (Fig. 1B), oscillating left and right in time with the presentation of the images (i.e. at a rate of two cycles every 15 s) and thus acting as a visual metronome for the hand movement the subject was to perform. Subjects were instructed to time their hand movement

using this metronome, rather than by attempting to match their hand movement to the displayed action.

At the start of each trial, subjects were visually instructed of the sequence order to be displayed (random or sequential), the action to perform (*flip* or *fist*), the target image to be detected (*peace* or *OK*), and whether the display and action sequences were congruent or incongruent. Verbal instructions were also given for initial trials within the experiment, to ensure that each subject was aware of the experimental conditions. Each condition was presented using both main image sequences (i.e. the sequence and target combinations: *flip* and *OK*, or *fist* and *peace*).

Each subject performed 18 trials each lasting 200 s, comprised of two practice conditions and eight test conditions repeated twice in succession. The order of tests was balanced across subjects to reduce any influence of learning or fatigue. Between trials, participants were permitted a small break (usually less than a minute) to rest their hands and eyes.

Two types of errors were detected and discarded from further analysis: late or missed responses (considered to be RTs longer than 1000 ms), and anticipation errors (defined as RTs shorter than 350 ms). False positive errors – responding to a non-target – were not logged by the program, but were very infrequent. Each subject's reaction times were then averaged across all remaining responses within each condition. Because of the random presentation of targets and occasional missed responses, between 15 and 22 responses were recorded per condition (typically 17). Subject mean responses were then averaged across the group, and differences tested with repeated measures ANOVA using SPSS. A 2×2 factorial within-subject design was used with the factors: congruency between the performed hand action and the displayed sequence (*congruent* versus *incongruent*) and the order of the displayed sequence (*sequential* versus *random*).

2.2. Results

We first tested whether predictability of the main sequence of images improved discrimination performance and in addition whether congruency between the hand actions that were performed and the images seen on screen affected discrimination. Subjects therefore viewed either random (unpredictable) or sequential series of hand images while performing the slow *fist* or *flip* hand actions, timed by the oscillating motion of a small fixation pointer (Fig. 1).

There were no anticipation errors. Late or missed responses to the targets (RTs > 1000 ms) accounted for 0.42% of all trials (5.1% of target trials). However, two subjects had error rates of 20 and 26% for target trials, falling outside the 99% confidence interval of the group mean, and were discarded from further analysis because of poor compliance with the task. Of the remaining 26 subjects, the mean error rate for target trials was 3.5% and the maximum was 6%. As these error rates were low, no attempt was made to analyse their occurrence with respect to task condition.

Mean reaction times for the four experimental conditions, averaged across the two image series are shown in Fig. 3. There was a highly significant effect of sequence order ($F(1,25) = 13.889; p = 0.001$) with participants responding faster to targets within a predictive main sequence than to targets
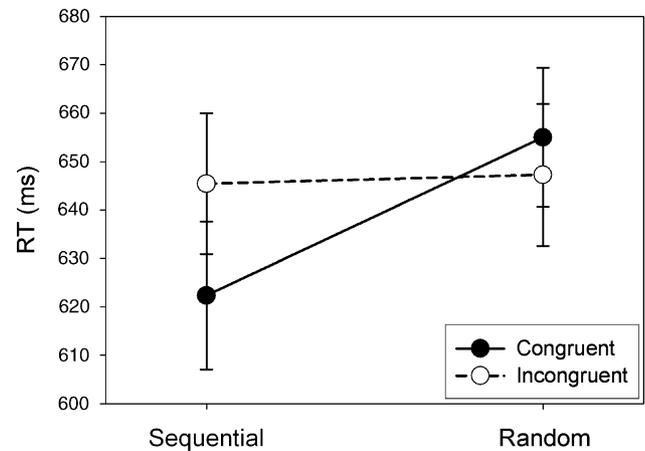

Fig. 3. Experiment 1: average reaction times (± 1S.E.M.) for discrimination of the target images during performance of congruent (filled circles, solid line) or incongruent actions (hollow circle, dashed line). The order of the main displayed image was sequential or random.

within a randomly presented main sequence (17.2 ms difference, S.E. = 4.9). However, most of this difference was driven by the large reduction in reaction times in the congruent movement condition. RTs were faster overall in the two congruent conditions than in the two incongruent conditions (mean difference: 7.7 ms, S.E. = 2.3; $F(1,25) = 10.997; p = 0.003$), and the interaction between congruency and sequence was highly significant ($F(1,25) = 13.889; p < 0.001$). That is, the shortest reaction times were for the sequential–congruent condition, 23.1 ms faster than the sequential-incongruent condition (S.E. of difference = 4.8 ms). Post hoc paired sample t-tests confirmed that responses in the sequential–congruent condition were significantly faster than in all three other conditions (26.9 ms faster than the mean of other three conditions, S.E. = 4.9 ms; statistically significant for both image series: d.f. = 25; $t > 3.95; p < = 0.001$). The 1.8 ms difference (S.E. = 6.7 ms) between the sequential and random series in the incongruent condition was not significant.

Unexpectedly, the image series also had an effect on reaction times, with faster mean response times in the '*fist*' sequence. Re-analysis of the data using a $2 \times 2 \times 2$ factorial design (including the two image series as the third factor), showed that the main effect of image series was significant ($F(1,25) = 16.414; p < 0.001$), and that interactions between series and congruency and between series and sequence and congruency were also significant ($F(1,25) = 5.44, 6.927; p = 0.028, 0.014$, respectively). However, crucially, the interaction between congruency and sequence remained highly significant ($F(1,25) = 10.997; p = 0.003$). Fig. 4 shows the data separated by image series; in both sets, the RTs in the sequential condition are lowest for the congruent condition, and in the random condition there is no difference between congruent and incongruent tasks.

2.3. Discussion

For this first experiment, we had hypothesised that prediction of the main sequence of displayed images would aid in the discrimination of the target images. Cued by this prediction, detection of the target images should be facilitated, as
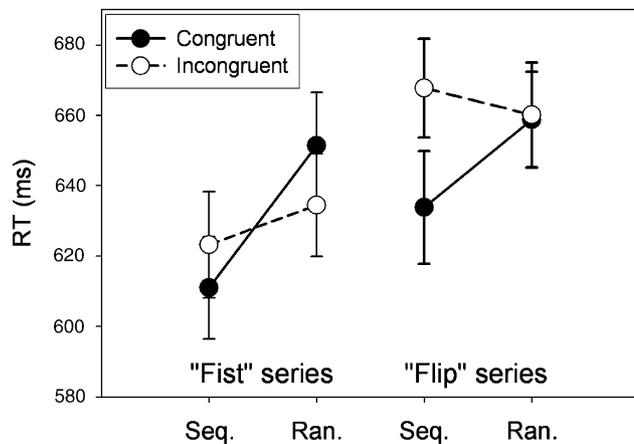
Fig. 4. Experiment 1: average reaction times (ms, ±S.E.M.) for detection of target stimuli, during sequential and random visual presentations of hand stimuli. Data are divided by displayed hand stimulus (fist or flip series) and whether the participant's performed hand movement was congruent (filled circles, solid line) or incongruent (hollow circles, dotted line) with this series.

the oddball targets do not match this anticipated visual representation. Reaction times were therefore expected to be faster when images were presented in a sequence rather than at random, and this main effect was found. This effect was small, however, and non-significant in the incongruent conditions. We further hypothesised that when a voluntary action is performed that is congruent with the displayed images, the corresponding prediction produced by a forward model should also assist in visual discrimination and further reduce reaction times. This second hypothesis was also supported by finding a significant interaction between the congruency and image sequence factors. The conjunction of these two predictive conditions – sequential image display and congruent action – produced reaction times significantly shorter than in the other three conditions, with a mean difference of 27 ms.

3. Experiment 2: Attention

Observing human actions while performing different actions has been shown to affect performance (Kilner, Paulignan, & Blakemore, 2003); it is easier to perform actions when they are congruent with the observed action. The same is true for congruence between actions and visual feedback (Poulton, 1974). Thus, one possible explanation for the reaction time savings seen during performance of congruent actions compared to incongruent actions is that there was an attentional difference between the tasks. To address this, we ran an experiment in which subjects were challenged to discriminate a change in the shape of the fixation pointer. If performance of congruent actions allows more attention to be paid to the discrimination task, then we would expect reaction time savings to be seen in this task as well.

3.1. Methods

Sixteen subjects participated in this experiment (5 male, 11 female); none had taken part in the previous experiment; age range was 18–30 years (mean 21.6 years).

All subjects performed two of the conditions presented in Experiment 1, namely the discrimination of the target hand images within the sequential "*fist*" and "*flip*" image series, while performing congruent and incongruent hand actions. They were also tested in a modified discrimination task (the "fixation task") in which they responded to the sudden change of the fixation cue from a rotating pointer (Fig. 1) to a cross, while again performing hand actions congruent or incongruent with the displayed image sequences.

The task, apparatus, stimuli and the experimental situation were the same as Experiment 1 with the exception that the fixation cue changed to a small cross, in the fixation target task. As in Experiment 1, the presentation of all images (whether the hand sequence, the hand target or the fixation cross) lasted for 1 s.

3.2. Results

For this new group of 16 subjects, there were no anticipation errors; late or missing responses (RT > 1000 ms) accounted for 0.17% of all trials and 2.44% of target trials. One of the 16 subjects performed poorly compared to the rest of the group, with RT standard deviations > 130 ms for six out of eight conditions, and was excluded from analysis. As expected from Experiment 1, responses were faster in the congruent condition than in the incongruent conditions (Fig. 5: 13.7 ms, S.E. = 7.1 ms; one-tailed t-test, $t = 1.916$, $p = 0.038$); however, the planned ANOVA did not reach statistical significance. The difference between reaction times when discriminating change of the fixation cue was negligible (2.4 ms, S.E. = 5.9 ms; two-tailed t-test, $t = 0.406$, $p = 0.691$).

3.3. Discussion

This experiment confirmed the reaction time savings for discrimination of hand images congruent with the performed action,
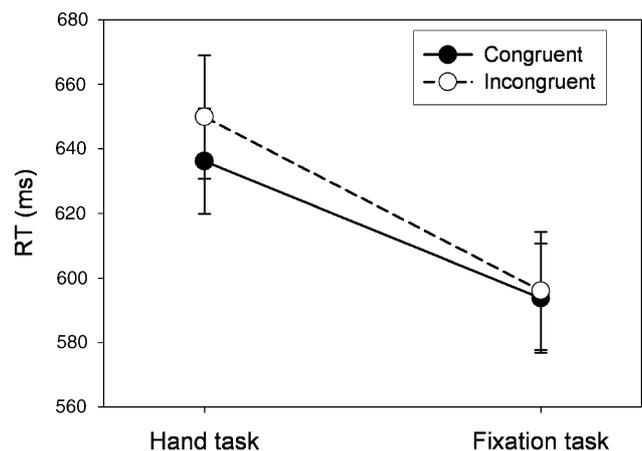

Fig. 5. Experiment 2: average reaction times (±1S.E.M.) for discrimination of the target images during performance of congruent (filled circles, solid line) or incongruent actions (hollow circle, dashed line). The task was to discriminate target hands within the displayed sequence of hand images ("hand task") or to discriminate change in the fixation pointer from a lollipop shape (Fig. 1) to a cross ("fixation task").

while finding no reaction time differences for discrimination of the pointer. It allows us to rule out differences in attentional resources between the congruent and incongruent conditions as the cause of the reaction time savings observed.

4. Experiment 3: Time course

We used the presentation of a slow series of static images, initially chosen so that comparisons could be made between the sequence and random conditions of Experiment 1. However, this then raises a question of when the predictions of each forthcoming image are generated and/or used: are the predictions continuously generated, while the subjects perform the slow, continuous hand movements, or are they generated or used only at the time of each visual presentation? Previous experiments on visuo-motor priming have shown that the effect of priming stimuli is short-lived, and have identified a time-window of 300–700 ms over which the onset of a congruent cue can prime responding (Vogt, Taylor, & Hopkins, 2003). We make a preliminary address to this question by modifying the display, so that each image was presented for 500 ms, separated by a neutral background for 500 ms.

4.1. Methods

Twelve (seven male, five female) of the original participants in Experiment 1 participated in the third experiment, after an interval of 5 months. Age range, 20–22 years (mean = 20.8).

A simple comparison was made between two conditions: congruent versus incongruent actions, in both cases for the sequential picture display only. Unlike Experiment 1, we did not use the random condition in this experiment. As before, both '*fist*' and '*flip*' image series were presented in each condition, creating a 2 × 2 factorial design. Again, the order of test presentations was manipulated between participants to minimize any order effects.

The task, apparatus, stimuli and the experimental situation were the same as Experiment 1 with the exception that the visual images shown on the screen were reduced in duration from 1000 to 500 ms, and were separated by a neutral grey screen for 500 ms. Hence, a short delay was introduced between presentations of each visual stimulus.

4.2. Results

Prior to analysing this experiment, and in order to confirm that this subset of 12 subjects was typical of the original group of subjects tested in Experiment 1, we correlated the performance of the two subgroups ($n = 12$ versus the remaining $n = 14$) across all eight conditions tested in Experiment 1 (two series, four task conditions). The correlation was highly significant ($r^2 = 0.699$, $p = 0.009$). The mean reaction time saving seen for this subgroup (24.3 ms, S.E. = 7.3 ms) in the congruent sequential condition of Experiment 1 versus the other three conditions was comparable to that found in the whole group (27 ms, S.E. = 4.9 ms). Hence, the subgroup retested in Experiment 3 was typical of the original group.

They showed no anticipation errors or late responses (RT > 1000 ms); missing responses accounted for 0.1% of all trials, 1.3% of target trials. As in Experiment 1, a significant effect of image series was found ($F(1,11) = 5.361$, $p = 0.041$), with shorter mean reaction times in the '*fist*' condition. However, the effect of congruency was not significant—the mean reaction times were only 3.7 ms faster in the congruent condition (S.E. = 4.6). Nor was there any significant interaction between the image sequence and the effect of congruency ($F(1,11) < 0.492$, $p > 0.247$).

4.3. Discussion

These results indicate that, with the introduction of a 500 ms interval between successive stimulus presentations, the reaction time savings seen in the first experiment were lost. While there were again response differences between the two series (with a small response advantage for the congruent condition over the incongruent condition in the '*flip*' series, and the reverse for the '*fist*' series), in neither case was the difference statistically significant.

In our experiment, we assume (but cannot yet demonstrate) that the forward model prediction of action is continuous, as the actions are themselves continuous. But our results show that the process that we assume underlies the integration of a continuous internal forward model prediction of action with the periodic process of visual discrimination appears to be short-lived and does not span the 500 ms interval. Vogt et al. (2003) reported a similar effect, with a time-window for effective priming of about 300–700 ms from prime cue onset to response onset. In other words, if the priming stimulus was present for longer durations, it was ineffective, strongly suggesting that it is the onset moment that is critical. Hence, in our task, the effect of the internal, action-related predictions on visual discrimination may also be time-locked to the visual stimulus onset—in our case each 1 s update of the static images of the hand. Clarifying this interaction will require additional experiments.

5. Experiment 4: Perspective

Visuo-motor priming appears to depend on perspective (Craighero et al., 2002; Vogt et al., 2003), with differences reported for movements cued by images of a hand shown in the first person perspective against cues in the third person perspective. A third person viewpoint advantage might reflect imitative experience, or experience of images seen in mirrors (Craighero et al., 2002); in contrast the first person effect observed for images of hands may reflect the action-relevance of the cue, allowing direct matching of the cue image with hand posture (Vogt et al., 2003). The objective of this experiment was thus to determine whether the motor–visual priming effects shown in Experiments 1 and 2 are dependent on the perspective of the image being displayed concurrently on the screen. Hence, we repeated the two main conditions of congruent and incongruent actions performed during a sequential image presentation (as tested in Experiment 1 and 2) in first versus third person perspective conditions, in another group of subjects.

5.1. Methods

Eighteen subjects participated (10 male, 8 female); none had taken part in the previous experiments; age range was 20–22 years.

A 2×2 factorial within-subject design was used with one factor being the congruency between subjects' own actions and the main sequence displayed, and the other factor being the perspective view of the images (first versus third). The main sequence was presented only in sequential mode, as used in Experiment 2, showing a predictably cyclic action. As before, the main sequence presented was one of two hand actions ('*fist*' or '*flip*'). Again, the order of test presentations was manipulated between participants to minimize any order effects.

The task, apparatus, stimuli and the experimental situation were the same as Experiments 1 and 2 with the exception that the visual images used in the third person view were top-bottom inverted from the first person view, as if one's own hand was viewed in a mirror.

5.2. Results

While 5 of the 18 subjects performed relatively poorly compared to the previous groups, missing 10–14% of target trials, none were identified as outliers. There were no anticipation errors; overall, late or missing responses (RT > 1000 ms) accounted for 0.4% of all trials and 5.3% of target trials.

The results for Experiment 4 are displayed in Fig. 6. For this group, the main effect of congruence was not significant and the main effect of the view was just outside significance ($p = 0.074$) with a trend toward faster responses in the first person view (8 ms mean difference, S.E. = 3.8 ms). However, there was a significant interaction between the perspective view and congruency ($F(1,17) = 9.935$, $p = 0.006$). As expected from Experiments 1 and 2, responses were faster in the first person congruent condition than in the first person incongruent conditions (9.4 ms mean difference, S.E. = 4.6 ms; one-tailed t-test, $t = 2.03$, $p = 0.029$).
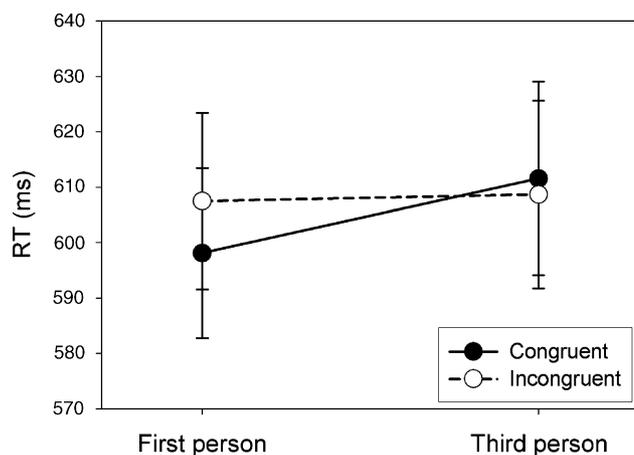

Fig. 6. Experiment 4: average reaction times (± 1S.E.M.) for discrimination of the target images during performance of congruent (filled circles, solid line) or incongruent actions (hollow circle, dashed line). The perspective of the displayed images was either upright ("first person", see Figs. 1 and 2) or inverted ("third person").

The difference caused by the first versus third person perspective view within the congruent condition was highly significant (13.5 ms, S.E. = 4.55 ms, two-tailed t-test, $t = 2.957$, $p = 0.009$), while the difference between the first and third person incongruent conditions was negligible (<1 ms, S.E. = 4.04 ms, $t = 0.3$).

As in the other experiments, there was a difference between the two image series used (*fist* versus *flip*; $2 \times 2 \times 2$ ANOVA $F(1,17) = 8.149$, $p = 0.011$); however, the interaction of series, view and congruency was not significant.

5.3. Discussion

Experiment 4 indicates that the previously observed RT advantage for congruent movements is lost when the images are put into the third person perspective. From Experiments 1 and 2, we expected and found a significant difference between first person congruent and incongruent conditions; that this difference was smaller than in Experiment 1 (9.4 ms versus 23.1 ms) may reflect the smaller subject group ($n = 18$ versus 26), of which several performed the task relatively poorly. It is therefore important that despite these factors, the difference between first person and third person view was highly significant in the congruent condition, while it was insignificant in the incongruent condition. Experiment 4 result is also consistent with Experiment 2: the congruency effect was seen only when discriminating the first person hand images, and not the unrelated change in fixation cue shape. This suggests the forward model prediction is advantageous only for discrimination of first person perspective images of the congruent hand action.

6. Experiment 5: Realistic biological rendering

Kilner et al. (2003) suggested that the conflict caused by visuo-motor incongruence affected movement performance only when the observed actor was a biological (human) agent, as no conflict was seen when the copied actor was a robot arm. In the same vein, imaging studies suggest separate processing of biological and non-biological rendered action images (Perani et al., 2001), and realistic and cartoon action sequences (Han, Jiang, Humphreys, Zhou, & Cai, 2005). However, in Kilner's study the kinematics differed between biological and non-biological agents, hence it is not clear whether the important congruency difference was due to the actor's kinematics or their biological nature. Many studies of biological motion show that the kinematics are important, as biological motion is readily attributed to moving lights (Johansson, 1973) as long as they obey the normal kinematic patterns (Ahlstrom, Blake, & Ahlstrom, 1997; Grossman et al., 2000). Recent work by Aymoz and Viviani (2004) has suggested that observation of biological agents performing actions allows anticipation within visual processing streams, and this effect was absent for non-biological agents. Again, their conditions changed both the agency and the kinematic parameters of the observed actions, and moreover, their task did not involve action performance, so involvement of the motor planning system, while likely, was not obligatory.

We repeated the two main conditions of congruent and incongruent actions performed during a sequential image presentation

(as tested in Experiments 1, 2 and 4) with the observed images rendered as photo-realistic biologically rendered hands, or as non-biological boxes. The kinematics of the slowly displayed sequences were however identical.

6.1. Methods

Ten subjects participated in the final experiment (six male, four female); none had taken part in previous experiments; age range was 21–60 years (median 21).

A 2×2 factorial within-subject design was used with one factor being the congruency between subjects' own actions and the main sequence displayed, and the other factor being the graphical rendering of the images (photo-realistic versus box-rendered; Fig. 1C). The main sequence was presented only in sequential mode, as used in Experiment 2, showing a predictably cyclic action. As before, the main sequence presented was one of two hand actions ('fist' or 'flip'). Again, the order of test presentations was manipulated between participants to minimize any order effects.

The task, apparatus, stimuli and the experimental situation were the same as Experiments 1, 2 and 4 with the exception that all images (main sequence and targets) used in the unrealistic condition were rendered as a set of cuboids (using the Poser "box-rendered" option; Fig. 1C).

6.2. Results

All subjects performed well; there were nine anticipation errors: 0.03% of all trials and only 0.69% of target trials. Overall, late or missing responses (RT > 1000 ms) accounted for 0.48% of all trials and 4.6% of target trials.

The results for Experiment 5 are given in Fig. 7. For this group, the main effect of congruence was significant (ANOVA $F(1,9) = 10.114$, $p = 0.011$), with faster responses in the congruent conditions, while the effect of photo-realistic rendering was not significant ($F(1,9) = 0.3$, $p > 0.5$). There was no significant interaction between the factors ($F(1,9) = 0.56$, $p > 0.4$).
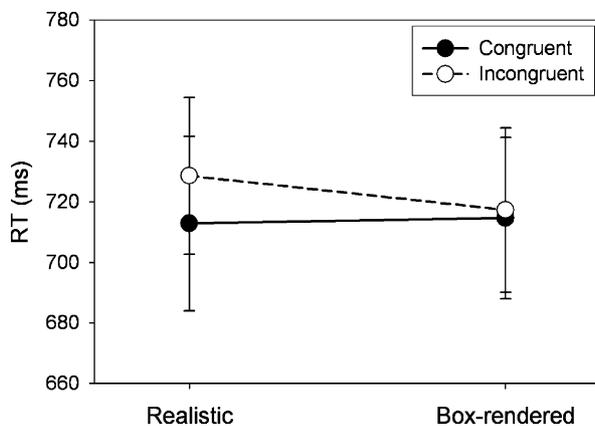

Fig. 7. Experiment 5: average reaction times ($\pm$1S.E.M.) for discrimination of the target images during performance of congruent (filled circles, solid line) or incongruent actions (hollow circle, dashed line). The rendering of the displayed images was either photo-realistic or unrealistic ("box-rendered", see Fig. 1C).

6.3. Discussion

This final experiment indicates that the visual representation of the biological agent performing the observed hand movements does not significantly affect the discrimination task, although there was a weak trend toward a greater effect for the photo-realistically rendered condition (Fig. 7). It replicates Experiments 1, 2 and 4 by showing reaction time advantages when there is congruence between the observed images and the actions being performed.

7. General discussion

We aimed to investigate whether the reaction time (RT) to discriminate target images among a sequence of visual stimuli could be reduced by performing actions congruent with the visual images of hands presented on the screen. We argue that if this effect was seen, it would suggest that enhanced prediction of the main sequence of images allowed more rapid discrimination of the oddball targets. RTs were indeed faster in Experiment 1 when the main images were presented in sequence, allowing the participants to predict the visual image series, but this effect was small. RTs were also faster when the images were congruent with the hand action being performed. But, crucially, reaction times were significantly shorter in the congruent and sequential condition than in the other three, suggesting that performance of the congruent action was the most important factor in predicting the image series, over and above the effect of an ordered visual sequence. This advantage for the congruent condition was also replicated in Experiments 2, 4 and 5, using different subject groups for each experiment.

Our second experiment indicates that the RT savings were not due to reduced attentional load in the congruent condition, as might be expected if controlling the hand action were easier in that condition. There was no difference in subjects' ability to discriminate a change of the fixation cue, suggesting that equal attention was available to the discrimination task, when performing congruent and incongruent actions. Two caveats need to be mentioned here regarding the fixation pointer discrimination task. First, the target stimuli are presented in a different spatial location to the hand stimuli (see Fig. 1B); second, the nature of the target stimulus (a pointer rather than the hand) differs considerably from the hand stimuli. Hence, one might argue that demands of this fixation task may be too far removed from the original to provide a sensitive test of attentional load. However, all dual task experiments incorporate, by intention, wide differences between the two tasks, and the attentional demand of one task are seen in the second. Hence, we believe that if attentional differences were important, they would show some effect on the fixation task as well as on the hand discrimination task.

The use of third person perspective stimuli (Experiment 4) provides additional evidence against differences in attentional load causing the reaction time differences we have observed. The third person discrimination task has identical task demands to the main, first person task. Furthermore, the stimuli are identical in all aspects to the first person task, except for perspective.

Given that humans have a preference for specular imitation, we suggest that performing congruent movements in this condition should be of equivalent difficulty, or even easier than in the first person perspective task. Hence, attentional load should be equal or possibly even lower. The lack of any priming while viewing the third person perspective stimuli suggests that the congruency advantage observed in Experiment 1 is not simply due to reduced attentional load in the congruent condition, thus freeing greater attentional resources for the discrimination task.

However, our third experiment showed that, with the introduction of a 500 ms interval between successive stimulus presentations, the reaction time savings were lost. Hence, the priming effect of action-related predictions assumed to underlie these performance differences appears to be short-lived and does not span the 500 ms inter-stimulus interval. This result is not easily integrated with the idea of a forward model based on the on-going action. Since the executed hand actions were continuous, despite the intermittent visual display, we would expect that the forward model predictions would also be continuous. Instead, it appears that the link between predictions based on motor execution and the visual discrimination processes is short-lived. One possibility is that visual stimuli engage the motor–visual predictive system, so that the output of the forward model is used by visual areas in a process time-locked to visual events. In this case, the presentation of the blank screen for 500 ms may act to disengage the system. This suggestion of a visual priming or engagement of a continuous motor prediction may be consistent with the 300–700 ms window for priming effects reported by Vogt et al. (2003); they found that the priming process was time-locked to onset of the visual priming stimulus. It suggests that the visual predictions from the forward model are only available in the context of visually relevant tasks.

However, it does not seem possible to explain the totality of our results in terms of visuo-motor priming. The major result seen in Experiments 1, 2, 4 and 5 – that congruency between observed and executed actions facilitates visual discrimination – can only be explained by visuo-motor effects if one assumes that these visual–motor effects alter the neural resources available to the discrimination task. This possibility has been discounted by the results of Experiments 2 and 4, in which no evidence was found for facilitated performance in the fixation task or the third person perspective conditions.

Another important question is why we find a facilitating effect of action on visual discrimination whereas Hamilton et al. (2004) found a contrastive effect. Again, it may be because of the short time scale of the integration between action and perception. In their experiments, subjects lifted a weight some 1–2 s before observing a video clip of a similar action, and kept the weight elevated until after the clip was finished. Hence, judgment of the video clips was quite separate in time from the performed lifting action. Additional experiments will be needed to better understand these temporal relationships.

In sum, we interpret the results of these experiments as evidence for the idea of motor–visual priming. They suggest that the internal representation of an action during its motor execution can influence the visual system that would analyse visual representations of the same action. The fourth experiment further suggests that the observer viewpoint is important, and that the predictive, internal representation of action aids discrimination of images seen in the first person perspective. It confers advantage neither in discrimination of images that are in the third person view, nor in the discrimination of changes in the fixation cue. The similarity of these null results in Experiment 2 (fixation condition) and Experiment 4 (third person view condition) provides further support against a role of differential attentional load causing the RT savings. If congruence between observed and executed actions allowed greater attention to the discrimination task, we would also expect that to have occurred in the third person condition, as the advantage of imitated actions can be found for both first and third person perspectives (Brass et al., 2001).

These results are therefore consistent with the ideomotor principle (Greenwald, 1970) or common event coding theory (Hommel et al., 2001). According to the event coding theory actions are coded in terms of the perceivable effects they generate. Since perception and anticipated action effects share the same code, this approach predicts that perceived events can prime or induce compatible actions. Indeed, Kilner et al. (2003) have recently shown that observing another human making incongruent movement has a significant interference effect on the execution of movement. But the common coding theory also predicts that action – even if just intended – could modulate perceptual processing. These results are consistent with the report by Aymoz and Viviani (2004). However, our results may go further than predicted by the common event coding theory: that theory would not predict that the reaction time savings should be specific to first person viewpoint, as the third person viewpoint stimuli also share a common code with the planned hand actions.

One unexpected result we found was the significant difference in response speeds between the two image types (*fist* and *flip*). It may simply be that the oddball targets were easier to discriminate in one series than the other, perhaps because some unavoidable differences between the target and main images aided their detection (Fig. 2). In an additional control experiment, eight subjects were instructed to react to the target images, without performing any hand actions. All four combinations of target (*OK* or *peace*) and images series (*fist* or *flip*) were tested. This subject group showed the same pattern of responses as seen in the main experiments—responses were significantly faster for targets shown during the *fist* sequence than the *flip* sequence, regardless of the target image. There was also a significant interaction between target image and the series, further suggesting a difference in matching difficulty. However, these differences are secondary to the main effect seen and the analyses undertaken that included image series as an additional factor showed the effects of interest (congruency versus sequence, or congruency versus viewpoint) remained statistically significant.

Some recent findings have provided a neural mechanism that would allow a direct matching between the visual description of an action and its execution. Recordings of single cell activity in macque monkeys have shown that a subset of neurons in the ventral premotor cortex – area F5 – (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996) and the anterior part of the inferior parietal cortex – PF or Brodmann's area 7b – (Gallese, Fogassi, Fadiga, &

Rizzolatti, 2002; Rizzolatti, Fogassi, & Gallese, 2001) discharge during execution of movement and when it observes another individual performing the same or similar action. Mirror-like neurons have also been reported in the superior temporal sulcus (STS). Neurons in this region respond to face, eye and hand goal directed actions (Perrett et al., 1989). However, they discharge only during observation of an action and not during its execution.

It therefore seems that there are three areas in the primate brain that contain neurons capable of forming a cortical action observation system. Neurons similar to those discovered in monkeys may also exist in humans. For example, transcranial magnetic stimulation of the human motor cortex during observation of hand movements selectively increases motor evoked potentials in muscles normally used to perform the observed hand action (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Gangitano, Mottaghy, & Pascual-Leone, 2004). The motor cortex is therefore selectively potentiated during action observation. Results from other electrophysiological studies (Hari et al., 1998; Strafella & Paus, 2000) support this conclusion, although some spinal changes may also be evoked (Vargas et al., 2004). Furthermore, neuroimaging studies using fMRI (Buccino et al., 2001; Iacoboni et al., 1999; Iacoboni et al., 2001), or PET (Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Rizzolatti et al., 1996) and event-related MEG (Nishitani & Hari, 2000) have localised the neurons responsible for these effects in humans to the ventral premotor and parietal cortices.

So how would action observation and forward modelling interact? Carr, Iacoboni, Dubeau, Mazziotta, and Lenzi (2003) propose that neurons in the STS, PF and F5 form an action representation circuit. During action observation these connections form an inverse model, they claim, converting the visual representation of actions in the STS into a motor plan in F5. For imitation these connections would then act in reverse as a forward model, converting an efferent copy of the motor plan back into a predicted visual representation for comparison with exemplar visual images (a sensory outcome of action). This mechanism could provide the basis for understanding the results of our experiments. When subjects prepare and execute their movements they would activate neurons located in area F5, and through the "forward" connections activate neurons located in areas PF and STS. Thus, preparation of movement may evoke an action plan in motor terms (involving F5 and PF mirror neuron activity) and also in visual terms (invoking PF and STS activity). These motor-evoked visual representations in STS or perhaps other extrastriate areas would then facilitate discrimination of target images, which do not correspond with the predicted sequence (motor–visual priming).

Finally, given that the mirror neuron system has been related to imitating and understanding the actions of others, it was unclear whether one would expect the predictive representations to be congruent with first or third person presentation (Craighero et al., 2002; Vogt et al., 2003). Most of the data on mirror neurons has been collected in allocentric third person conditions. However, it has been proposed that when one perceives one's own actions, there is a closer match between the predicted and actual outcomes of the action than when we perceive the actions of others (Knoblich, Seigerschmidt, Flach, & Prinz, 2002). Therefore,

it could be expected that in motor–visual priming the first person perspective would show stronger effects than the third person perspective, and this was indeed the case. Effects of congruent image presentation were only found during the presentation of stimuli in the first person perspective.

8. Conclusion

We have shown that action execution can assist the discrimination of visual images, when the image sequence is congruent with the executed action, and when the images are shown in the first person perspective. The link between these predictions and visual processing appears to be short lasting. We interpret our results as evidence supporting the hypothesis that forward models in the motor system provide action-specific sensory predictions that are available to cognitive processes independent of motor control.

References

Ahlstrom, V., Blake, R., & Ahlstrom, U. (1997). Perception of biological motion. *Perception*, *26*, 1539–1548.

Aymoz, C., & Viviani, P. (2004). Perceptual asynchronies for biological and non-biological visual events. *Vision Research*, *44*, 1547–1563.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychology (Amsterdam)*, *106*, 3–22.

Brass, M., Bekkering, H., Wohlschlager, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*, 124–143.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 5497–5502.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia*, *40*, 492–502.

Craighero, L., Fadiga, L., Rizzolatti, G., & Umilta, C. (1999). Action for perception: A motor–visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1673–1692.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Science*, *3*, 172–178.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*, 90–92.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.

Frak, V., Paulignan, Y., & Jeannerod, M. (2001). Orientation of the opposition axis in mentally simulated grasping. *Experimental Brain Research*, *136*, 120–127.

Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*, 1771–1788.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(Pt 2), 593–609.

Gallese, V., Fogassi, L., Fadiga, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Attention and performance: vol. XIX* (pp. 247–266). Oxford University Press.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2004). Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. *European Journal of Neuroscience*, *20*, 2193–2202.

Gentilucci, M., Daprati, E., & Gangitano, M. (1998). Right-handers and left-handers have different representations of their own hand. *Brain Research. Cognitive Brain Research*, *6*, 185–192.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research*, *112*, 103–111.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Reviews*, *77*, 73–99.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, *12*, 711–720.

Hamilton, A., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology*, *14*, 493–498.

Han, S., Jiang, Y., Humphreys, G. W., Zhou, T., & Cai, P. (2005). Distinct neural substrates for the perception of real and virtual visual worlds. *Neuroimage*, *24*, 928–935.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 15061–15065.

Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Computation*, *13*, 2201–2220.

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behaviour and Brain Science*, *24*, 849–878.

Iacoboni, M., Koski, L. M., Brass, M., Bekkering, H., Woods, R. P., Dubeau, M. C., et al. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 13995–13999.

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

Jeannerod, M., & Decety, J. (1995). Mental motor imagery: A window into the representational stages of action. *Current Opinion in Neurobiology*, *5*, 727–732.

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

Jordan, M. I. (1995). Computational motor control. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 597–609). Cambridge, Mass: MIT Press.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, *16*, 307–354.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinions in Neurobiology*, *9*, 718–727.

Kawato, M., Furukawa, K., & Suzuki, R. (1987). A hierarchical neural network model for control and learning of voluntary movement. *Biological Cybernetics*, *57*, 169–186.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13*, 522–525.

Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *The Quarterly Journal of Experimental Psychology A*, *55*, 1027–1046.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265–1279.

Musseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology. Human Perception and Performance*, *23*, 861–872.

Musseler, J., Steininger, S., & Wuhr, P. (2001). Can actions affect perceptual processing. *The Quarterly Journal of Experimental Psychology A*, *54*, 137–154.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 913–918.

Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology. Human Perception and Performance*, *20*, 709–730.

Parsons, L. M., Fox, P. T., Hunter Downs, J., Glass, T., Hirsch, T. B., Martin, C. C., et al. (1995). Use of implicit motor imagery for visual shape discrimination as revealed by PET. *Nature*, *375*, 54–58.

Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., et al. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage*, *14*, 749–758.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, *146*, 87–113.

Poulton, E. C. (1974). *Tracking skill and manual control*. London, NY: Academic Press.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET: 1. Obsevation versus execution. *Experimental Brain Research*, *111*, 246–252.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews. Neuroscience*, *2*, 661–670.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, *24*, 113–121.

Smith, D. T., Rorden, C., & Jackson, S. R. (2004). Exogenous orienting of attention depends upon the ability to execute eye movements. *Current Biology*, *14*, 792–795.

Strafella, A. P., & Paus, T. (2000). Modulation of cortical excitability during action observation: A transcranial magnetic stimulation study. *Neuroreport*, *11*, 2289–2292.

Vargas, C. D., Olivier, E., Craighero, L., Fadiga, L., Duhamel, J. R., & Sirigu, A. (2004). The influence of hand posture on corticospinal excitability during motor imagery: A transcranial magnetic stimulation study. *Cerebral Cortex*, *14*, 1200–1206.

Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: Perspective matters. *Neuropsychologia*, *41*, 941–951.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, (Suppl. 3), 1212–1217.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor control. *Science*, *269*, 1880–1882.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature*, *414*, 302–305.