

Motor experience interacts with effector information during action prediction

Lincoln J. Colling (lincoln@colling.net.nz)

Donders Institute for Brain, Behaviour and Cognition, Radboud University Nijmegen
Nijmegen, Netherlands

William F. Thompson (bill.thompson@mq.edu.au)

ARC Centre of Excellence in Cognition and its Disorders, Macquarie University
Sydney, Australia

John Sutton (john.sutton@mq.edu.au)

ARC Centre of Excellence in Cognition and its Disorders, Macquarie University
Sydney, Australia

Abstract

Recent theory suggests that action prediction relies of a motor emulation mechanism that works by mapping observed actions onto the observer action system so that predictions can be generated using that same predictive mechanisms that underlie action control. This suggests that action prediction may be more accurate when there is a more direct mapping between the stimulus and the observer. We tested this hypothesis by comparing prediction accuracy for two stimulus types. A mannequin stimulus which contained information about the effectors used to produce the action and a point stimulus, which contained identical dynamic information but no effector information. Prediction was more accurate for the mannequin stimulus. However, this effect was dependent on the observer having previous experience performing the observed action. This suggests that experienced and naïve observers might generate predictions in qualitatively different ways, which may relate to the presence of an internal representation of the action laid down through action performance.

Keywords: Joint action; embodied cognition; perception-action; action prediction.

Introduction

Many types of joint action require two actors to coordinate their actions. Such coordination is especially demanding for joint actions, such as ensemble music and dance performance, where successful completion of the joint action requires precise temporal synchronisation. In these contexts, it is not possible for individuals to observe and then react to the actions of their co-actors because this would introduce disruptive delays. Rather, individuals must anticipate the actions of their co-actors so that they can plan actions that will align with those actions. Because of these time constraints, researchers have emphasised the role of prediction in recent theoretical accounts of joint action coordination (Csibra, 2008; Wilson & Knoblich, 2005).

Models of predictive mechanisms in motor control, such as forward models and inverse models, can greatly inform our understanding of joint action coordination. Both classes of models are contained within the model of motor control developed by Wolpert and colleagues (e.g., Wolpert, 1997). According to this framework, forward and inverse models are used in tandem to achieve goal-directed behaviour when regular feedback is unreliable because of delays or inaccuracies.

Inverse models act as controllers by transforming a goal state into a series of control commands that are then sent to the controlled system to produce the desired behaviour. In the motor control system, this is implemented by a system that takes the goal state and transforms it into a series of motor commands. Forward models, on the other hand, take the motor commands and transform them into a goal state. The limb also performs a forward mapping from motor commands into a goal state. Therefore, the forward model can be used to predict how the limb is expected to behave. A forward model is particularly useful in motor control where it can be used to bypass delays that occur because feedback must be transmitted from the periphery to centrally located motor control regions. This can compensate for these delays by generating predicted feedback that can be substituted for the delayed feedback.

Motor involvement in action prediction

In addition to their role in intrapersonal action prediction, forward models and inverse models are also implicated in the interpersonal action prediction needed for joint action. Csibra (2008) has suggested that during action observation, an inverse model allows observers to reconstruct the motor codes used to produce the observed action. To support this claim, Csibra cites evidence from electrophysiological studies on monkeys as well as neuroimaging studies from humans that show that neurons in motor regions are active not only when actions are produced but also when the same, or similar, actions are passively observed (for a review, see Rizzolatti & Craighero, 2004). Additionally, Wilson and Knoblich (2005) have proposed that observers are able to construct an internal model of observed actions by mapping the actions onto their own motor systems in a part-by-part, or isomorphic, manner. This internal model acts as a *forward model* by generating a real-time simulation of the observed action that runs in parallel with incoming sensory information. Information from this model can be substituted for incoming sensory information that reaches the observer through observation. By using internally generated information to drive action planning, delays that result from the processing of external actions can be overcome, and this allows co-actors to plan and coordinate

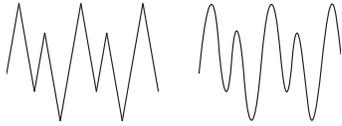


Figure 1: The zigzag (left) and wave (right) patterns used as stimuli during the recording session.

joint actions in time critical situations. Taken together, forward and inverse models provide an effective mechanism that allows actors to use their own action systems in order to emulate¹ the actions of others. We refer to this proposal as the *emulator hypothesis of action prediction*.

A key prediction of the emulator hypothesis is that traces of the observer's motor system should be manifested in the predictions that they generate. The *authorship effect* provides a means of assessing this. The authorship effect refers to the finding that observers are more accurate at generating predictions about recordings of self-generated actions relative to other-generated actions. More generally, the greater the alignment between the motor dynamics of the observer and the motor dynamics of the agent producing the observed action, the more accurate the predictions generated by the observer (Flach, Knoblich, & Prinz, 2003; Keller, Knoblich, & Repp, 2007; Colling, Sutton, & Thompson, 2010, submitted).

Motor involvement in event prediction

In addition to the motor system's role in predicting the actions, evidence from paradigms employing abstract stimuli suggest that the motor system might also be used for sequence prediction in general. For example, findings from fMRI implicate ventral premotor regions in tasks that require participants to generate predictions about abstract sequences (Schubotz & von Cramon, 2004). Similarly, lesions in premotor regions are associated with deficits in sequence prediction (Schubotz, Sakreida, Tittgemeyer, & von Cramon, 2004).

Based on these findings, Schubotz (2007) has suggested that motor simulation is a general mechanism for predicting events. In the case of reproducible events—that is, human actions—these events are simulated using the same means that were initially employed to create the event, by using an internal model of the action. However, in the case of event that can't be mapped onto the body Schubotz argues that predictions are generated using an action model of an effector that best matches the general dynamics of the stimulus. Similarly, impoverished action stimuli lacking detail about which effectors were used to produce the action, or actions that are not in the observers repertoire, might be simulated using this more general mechanism. While this might provide a good general description of the stimulus dynamics it may fail to replicate fine-grained details of the stimuli.

¹Emulate to refers to the process of replicating the functions of a system (e.g., a conspecific's motor system) using different means (e.g., the observer's motor system).

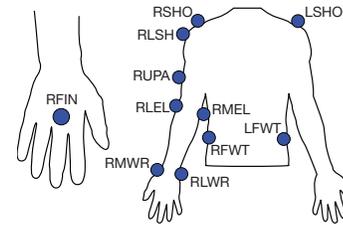


Figure 2: Marker positions for recording session.

Aims of the current study

The primary aim of the present study is to examine the nature of the internal model that observers use during action prediction. In particular, our aim is to examine whether action prediction is achieved via a general purpose predictive system that, although implemented with the motor system, does not rely on a part-by-part simulation of the observed action. The present study measured prediction accuracy by means of an action synchronisation task similar to that reported in Colling et al. (2010, submitted); however, rather than examining differences in synchronisation accuracy for self-produced and other-produced actions, all participants viewed other-produced actions and we instead varied the properties of the stimulus as well as the relevant motor experience of the observers. Both manipulations were designed to modify the information that participants could access to allow them to map the observed actions onto their own action systems. The motor experience manipulation was designed to provide observers with an internal representation of the action onto which they could map the stimulus, while the stimulus manipulation was designed to modify whether the stimulus could be directly mapped onto the observers' bodies.

To modify the information content in the stimulus, we constructed two sets of stimuli so that the stimulus either contained information about what effectors were employed to produce the action (full information) or only contained the motion information required to perform the synchronisation task, but excluded any information about the effectors used to produce the action (point information).

A manipulation was also designed to examine the role of motor experience on action emulation. This was achieved by dividing the participants into two groups and only providing one of the groups with experience with actually performing the action that they would later observe. Schubotz (2007) has suggested that when observers predict actions that are part of their action repertoire, they emulate the actions using an internal model of that action that has been laid down by the experience of producing the action. Naïve observers, on the other hand, might only employ motor regions that match the general dynamics of the movement. If this is the case, then we can predict that the effect of stimulus content would be modulated by motor experience. In particular, we can predict that naïve observers would not incorporate information about the effectors used to produce the movement into their predictive

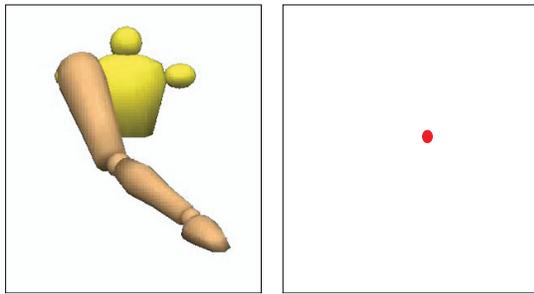


Figure 3: Example stimuli from the full information condition (left) and the point information condition (right).

model; thus, the addition of this information should provide no additional benefit on the synchronisation task.

Methods

Participants

The motor experience group contained of 13 participants (11 females, mean age of 28.1 years). The naïve group contained of 12 participants (8 females, mean age of 20.7 year). All participants were right-handed, and all procedures were approved by the Macquarie University Human Subjects Ethics committee.

Stimuli

In order to create the stimuli for the test session, five right-handed females (mean age of 24.8 years) performed the movement task while their movements were tracked with motion capture.

The movement task involved tracing out wave and zigzag patterns (see Figure 1) as if drawing them on an imaginary blackboard. The patterns were displayed on two large sheets of cardboard measuring 0.594 m × 0.841 m. Both patterns contained five upward and five downward movements alternating between long and short. The two patterns differed in terms of the nature of the direction change at the apex of each upward movement. The direction changed sharply for the zigzag pattern, while there was a smooth, flowing direction change for the wave pattern².

Movements were recorded using an 8-camera 3-D passive optical motion capture system (Vicon MX with 4 Vicon MX-F20 and 4 Vicon MX13+ cameras) at a sampling rate of 200 Hz from markers placed on the subject's shoulders, right arm, right hand, and waist (see Figure 2). Raw motion capture data was resampled to 25 Hz and processed with C-Motion Visual 3D (C-Motion INC, Rockville MD) to create the test stimuli. For the full information condition, the motion

²The difference between the wave and zigzag patterns is of no theoretical interest. Two movement patterns were used only to increase task variety during the stimulus creation phase and ensure participants remained engaged with the task. Statistical analyses confirmed that there were no systematic differences in performance on the test phase as a result of stimulus form (wave, zigzag) and, therefore, this factor was dropped from the analyses reported below.

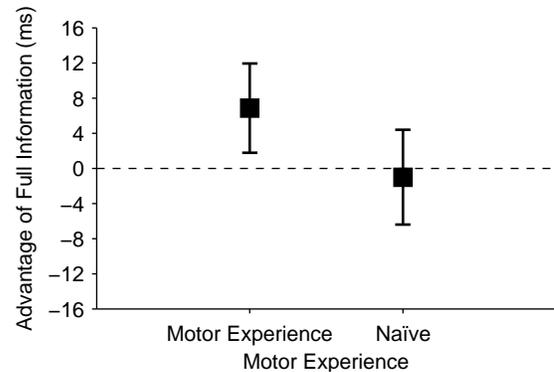


Figure 4: Improvement in synchronisation accuracy for the full information condition relative to the point information condition for the motor experience group and the naïve group. Error bars indicate the 95% confidence interval.

capture data was rendered as an animated character consisting of an upper torso, right arm and hand, while the stimuli for the point information condition consisted of only a single point tracking the hand (see Figure 3).

Procedure

Participants in the motor experience group undertook a movement session that was identical to the task employed during stimulus creation. Participants performed 3 blocks containing 5 repetitions of each pattern (in random order) with their eyes closed to limit visual experience. The movement session and the test session were on average separated by 15.85 days (7 to 27 days).

The task in the test session was to press the response button when the hand of the mannequin, or the marker tracking the hand, reached the apex of each upward movement. Participants were instructed to synchronise the button-press with the display as accurately as possible and were told that this may require them to anticipate when the peak will occur. Each participant performed 4 blocks containing 40 unique stimuli, with equal numbers of full and point stimuli, and equal numbers of wave and zigzag stimuli. Participants in the naïve group were given a brief verbal description of the movement task.

Results

Timing error was calculated as the absolute difference between the timing of the peak in the motion capture trajectory and the timing of the button-press. Absolute timing error was used as a dependent measure because it has been shown to provide a good index of accuracy of hitting the target (Spray, 1986). Absolute timing error was analysed by means of a $2 \times 4 \times 2$ mixed ANOVA with the within-subjects factors Information Content (full information, point information) and Block (1, 2, 3, and 4), and the between-subjects factor of Experience (motor experience, naïve). The Greenhouse-Geisser procedure was used to correct for violations of sphericity.

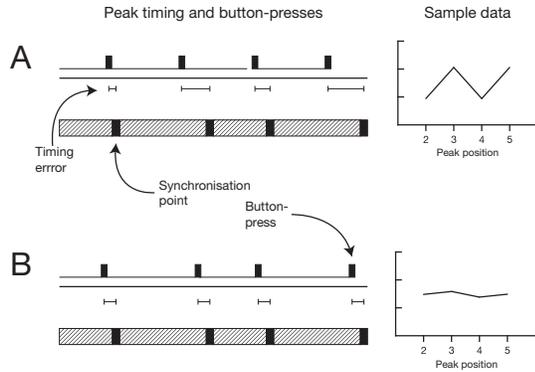


Figure 5: (A) Evenly spaced button-presses results in timing errors that vary as a function of peak number. (B) Timing error that does not vary as a function of peak number is a result of the timing of button-presses varying as a function of peak number.

Where appropriate, we report uncorrected *dfs* along with the corrected *p* value.

There were no systematic differences in synchronisation accuracy related to experimental block, as indicated by the non-significant main effect for Block ($F_{3,69} = 0.250$, $p = .861$, $\epsilon = .518$, $\eta_G^2 = .002$), and the non-significant interactions for Information Content \times Block ($F_{3,69} = 0.368$, $p = .777$, $\eta_G^2 = .001$), Block \times Experience ($F_{3,69} = 1.024$, $p = .352$, $\epsilon = .518$, $\eta_G^2 = .006$) and Information Content \times Block \times Experience ($F_{3,69} = 0.609$, $p = .611$, $\eta_G^2 = .001$). There were also no systematic differences in synchronisation accuracy between the naïve group and the group with motor experience, as indicated by the non-significant main effect of Experience ($F_{1,23} = 0.460$, $p = .504$, $\eta_G^2 = .016$). Furthermore, there were no systematic differences in synchronisation accuracy between the full information displays and point information displays when the data were collapsed across group and block, as indicated by the non-significant main effect for Information Content ($F_{1,23} = 8.573$, $p = .008$, $\eta_G^2 = .003$).

As predicted, the results showed that the effect of information content was modulated by motor experience, as indicated by the significant interaction for Information Content \times Experience ($F_{1,23} = 5.413$, $p = .029$, $\eta_G^2 = .002$). To decompose this interaction, the data were collapsed across block and two paired t-tests were conducted to examine the difference between the two levels of Information Content (full information, point information) for each Experience group. The results of these t-tests showed that the information content effect was found only for the motor experience group ($t_{12} = 2.943$, $p = .012$) but not the naïve group ($t_{11} = -0.411$, $p = .689$). This indicates that timing error was significantly higher for the point stimuli relative to full stimuli for the motor experience group ($M_{\Delta} = 6.855$, 95CI[1.779][11.930]) but not for the naïve group ($M_{\Delta} = -1.007$, 95CI[-6.399][4.385]). Therefore, only the motor experience group was able to take

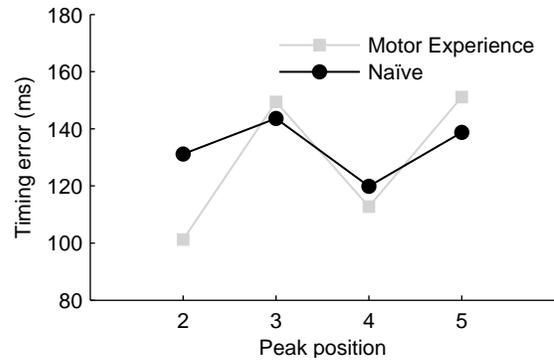


Figure 6: Timing error as a function of peak position for the motor experience group and the naïve group.

advantage of the presence of limb and joint information to enhance synchronisation accuracy (See Figure 4).

A further attempt was made to quantify the difference in task performance between the motor experience group and the naïve group. We examined whether there were differences in task performance that related to whether participants primarily responded to local aspects or global aspects of the stimulus. In the stimuli, the duration of each upward movement alternated from long to short. This irregular pattern leads to local variations in peak timing. Basing responses on global aspects of the stimuli, such as average tempo, would produce a pattern of timing errors that fluctuates from peak to peak. However, by adjusting responses according to the local variations in the stimuli would produce timing errors that are approximately equal for each button-press (see Figure 5). Timing error for each of the final four button-presses (corresponding to each of the final four peaks) was analysed separately for each group by means of a one-way ANOVA with the factor Peak Position (The first peak was dropped from the analysis because the movement leading up to the first peak is neither clearly long nor short).

Analyses showed a significant effect of Peak Position for the naïve group ($F_{3,33} = 5.083$, $p = .031$, $\epsilon = .453$, $\eta_G^2 = .108$), and not for the motor experience group ($F_{3,36} = 1.449$, $p = .254$, $\epsilon = .371$, $\eta_G^2 = .020$). This suggests that for the naïve group timing error changed in a low-high-low-high pattern as the trial progressed, while for the motor experience group peak position did not significantly affect timing error. These results are consistent with the naïve group responding to global aspects of the stimuli and the motor experience group responding to local aspects of the stimuli (see Figure 6).

Discussion

The primary aim of the present study was to investigate the nature of the action emulation employed during action prediction. In particular, we wanted to investigate whether action prediction relies on observers mapping the stimulus onto their body in a part-by-part manner, or whether they just model

the general dynamics of the action without modelling the specifics of the effectors used to create the stimulus. Furthermore, our aim was to investigate the influence of motor experience on action emulation. Schubotz (2007) has suggested that while abstract stimuli and actions that are not part of the observer's repertoire might be simulated using a general purpose mechanism, actions an observer has experience producing are instead simulated using a model that incorporates the specifics of the effectors used to produce the movement. In order to examine these questions, we varied both the information content of the stimuli and the motor experience of the observers.

The effect of stimulus information

In the full information condition, the visual stimulus contained not only the movement information required to perform the task, but also information about the effector used to produce the movement. In the point information condition, the visual stimulus only contained a single moving point. While the point information condition also contained all the movement information required to perform the task, it lacked the additional information about the state of the effectors. As predicted, the results showed enhanced prediction accuracy when observing the full information stimuli.

An alternative explanation for these results is that a difference in the low-level visual features of the stimuli might account for the differences in prediction accuracy. For example, it might be the case that the point information stimulus, which overall contains less visual information, is harder to visually track, and this may manifest as decreased prediction accuracy. However, this could account for the differences in synchronisation accuracy, then this difference should be present in both the experienced and the naïve groups. This was not the case; therefore, the results are not consistent with an explanation based on low-level visual features.

The finding that limb and joint information was able to enhance synchronisation accuracy is also consistent with what is known about the mirror-neuron system, the putative substrate of the action prediction system (Csibra, 2008; Wilson & Knoblich, 2005). Studies by Buccino and colleagues (e.g., Buccino et al., 2004) have shown that during action observation, regions of the motor cortex are activated in a somatotopic fashion. That is, certain regions show specificity for particular effectors in a manner similar to Penfield's (1954) motor homunculus.

The effect of motor experience

A secondary aim of the present study was to examine what influence motor experience would have on prediction accuracy. The results show that motor experience modulated the effect that stimulus type had on prediction performance. In addition, motor experience had an effect on how participants performed the task. In particular, these data show that while the timing error for experienced participants was not affected by the serial position of the peak, the timing error for naïve participants varied according to peak number, and the tim-

ing error for the small peaks was significantly different to the timing error for large peaks.

A pattern of fluctuations would arise if participants kept a relatively steady pace throughout the trial because the spacing of the peaks was not constant throughout the trial, but instead changed according to the height of the peaks. A relatively steady pace for button-presses might occur if participants responded to the global properties of the stimuli, such as the average rhythm (that is, the pace of movement production), or to the general stimulus dynamics. In order to maintain a relatively constant timing error, as seen in the experienced group, participants would need to adjust the timing of each button-press according to the local timing variations in the stimuli that result from the alternating heights of the peaks. This pattern of data, therefore, suggests a global/local bias in stimulus processing that is modulated by motor experience.

The effect of motor experience on processing visual stimuli has recently been noted in several studies. For example, Casile and Giese (2006) have shown that motor training enhances a participant's ability to make a fine-grained visual discrimination of action. In their study, participants were asked to make same/different judgments about gait patterns that they either did or did not have motor experience with. The results showed that participants performed significantly better for trained, or familiar, gait patterns compared with unfamiliar gait patterns, suggesting that in order to make fine-grained visual judgments about the kinematics of an action, observers need to have an internalised model of the action. Similarly, Calvo-Merino, Ehrenberg, Leung, and Haggard (2010) found superior performance on a visual discrimination task of dance moves when those dance moves were part of the observers action repertoire compared with dancers who only had visual experience with the dance moves.

Our findings build on these earlier results, and suggest that observers with motor experience for the observed action are also better able to make fine-grained predictions about the dynamics of a stimulus. This result is consistent with the notion that experienced observers generate predictions about observed actions by employing an internal model of that action that is acquired through motor experience. By mapping the observed action onto their internal model for that action they are better able to capture the fine-grained timing variations in the stimulus because their predictive model more completely captures the constraints specific to the effectors used to produce the action.

Motor experience modulates stimulus effects

We have argued that participants with motor experience are more sensitive to the fine-grained timing differences present in an action because they, unlike naïve observers, employ an internalised model of the observed action in order to generate their predictions. This might also help to explain why the influence of limb and joint information was restricted to the group with motor experience. Several neuroimaging studies have shown experience-related differences in motor system activation when observers view actions performed by other

people. For example, Calvo-Merino, Glaser, Grézes, Passingham, and Haggard (2005) asked expert dancers and non-dancer controls to view videos of dancers performing in one of two styles (ballet or capoeira). The results showed that activation in motor regions was greater when dancers viewed performers of their own style, suggesting that the motor system is preferentially engaged when observing actions that are familiar. A follow-up study (Calvo-Merino, Grézes, Glaser, Passingham, & Haggard, 2006), using male and female ballet dancers extended this finding by showing that motor regions were preferentially activated when viewing gender-specific dance moves. As both male and female dancers presumably have equal visual experience with opposite gender dance moves, but different motor experience, this finding suggests that motor engagement with visually presented stimuli is selective for actions for which observers have specific motor familiarity over and above the effects of visual familiarity.

Studies by Schubotz and colleagues (for a review, see Schubotz, 2007) have also implicated premotor regions in prediction of abstract stimuli and in sequence prediction in general. These stimuli cannot be mapped onto the observer's body and, therefore, they might rather be predicted by using a predictive model that exploits the dynamics of an effector that most closely matches the dynamics of the stimulus. Similarly, differences in motor system activation related to motor experience might suggest that inexperienced observers employ general predictive mechanisms, such as those used for sequence prediction, even when the observed action can, at least in principle, be mapped onto their body in an isomorphic, or part-by-part, manner. The addition of limb and joint information was designed to assist the process of mapping the observed action onto the observer's body; however, if, as the neuroimaging data cited above suggests, naïve observers less readily map the observed action onto their body in an isomorphic manner, then providing information to assist this process should provide no additional benefit. This is indeed what was found in the present study.

References

- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *J Cog Neuro*, *16*(1), 114–126.
- Calvo-Merino, B., Ehrenberg, S., Leung, D., & Haggard, P. (2010). Experts see it all: configural effects in action observation. *Psych Res*, *74*, 400–406.
- Calvo-Merino, B., Glaser, D. E., Grézes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cereb Cortex*, *15*(8), 1243–1249.
- Calvo-Merino, B., Grézes, J., Glaser, D., Passingham, R., & Haggard, P. (2006). Seeing or doing? influence of visual and motor familiarity in action observation. *Curr Bio*, *16*(19), 1905–1910.
- Casile, A., & Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Curr Bio*, *16*(1), 69–74.
- Colling, L. J., Sutton, J., & Thompson, W. F. (2010). Action synchronisation with biological motion. In W. Christensen, E. Schier, & S. J. (Eds.), *ASCS2009: Proceedings of the 9th Conference of the Australasian Society for Cognitive Science* (pp. 49–56). Sydney: Macquarie Centre for Cognitive Science.
- Colling, L. J., Sutton, J., & Thompson, W. F. (submitted). Authorship effects in action synchronisation with biological motion.
- Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition* (pp. 435–459). New York: Oxford University Press.
- Flach, R., Knoblich, G., & Prinz, W. (2003). Off-line authorship effects in action perception. *Brain Cogn*, *53*(3), 503–513.
- Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Conscious Cogn*, *16*(1), 102–111.
- Penfield, W. (1954). Mechanisms of voluntary movement. *Brain*, *77*(1), 1–17.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Ann Rev Neuro*, *27*(1), 169–192.
- Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *TiCS*, *11*(5), 211–218.
- Schubotz, R. I., Sakreida, K., Tittgemeyer, M., & von Cramon, D. Y. (2004). Motor areas beyond motor performance: Deficits in serial prediction following ventrolateral premotor lesions. *Neuropsychology*, *18*(4), 638–645.
- Schubotz, R. I., & von Cramon, D. Y. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *J Neuro*, *24*(24), 5467–5474.
- Spray, J. A. (1986). Absolute error revisited: An accuracy indicator in disguise. *J Motor Behav*, *18*(2), 225–238.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psych Bull*, *131*(3), 460–473.
- Wolpert, D. M. (1997). Computational approaches to motor control. *TiCS*, *1*(6), 209–216.