

This is a repository copy of *Spatio-temporal judgements of observed actions : Contrasts between first- and third-person perspective after motor priming.* .

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/85502/>

Version: Submitted Version

Article:

Brattan, Victoria Caroline, Baker, Daniel Hart orcid.org/0000-0002-0161-443X and Tipper, Steven Paul orcid.org/0000-0002-7066-1117 (2015) Spatio-temporal judgements of observed actions : Contrasts between first- and third-person perspective after motor priming. *Journal of Experimental Psychology: Human Perception and Performance*. pp. 1236-1246. ISSN 1939-1277

<https://doi.org/10.1037/xhp0000079>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Spatio-temporal judgements of observed actions:

Contrasts between first- and third-person perspectives after motor priming.

Victoria C. Brattan, Daniel H. Baker & Steven. P. Tipper

Department of Psychology, The University of York, York, United Kingdom, YO10 5DD

Running head: Action prediction, temporal prediction, time, motor priming & action observation

Author note:

This series of studies forms part of the PhD of Victoria Brattan.

Victoria Brattan is supported by an ESRC PhD Studentship.

Please direct all correspondence about the article or the research materials to Victoria Brattan at victoria.brattan@york.ac.uk.

ABSTRACT

When observing actions, motor simulation processes aid the prediction and understanding of future events. A central issue concerns whether such action simulation serves social functions of interpreting other people, where performance is predicted to be better when third-person perspective (3PP) actions are viewed; or whether it is most beneficial to guide self actions, whereby the first-person perspective (1PP) would be advantageous. We show that in a spatio-temporal judgement task there is an advantage for the prediction of 1PP. However, this is only detected after motor priming whereby participants perform the observed actions prior to making spatio-temporal judgements. The results, firstly, confirm that we draw on our motor experience for the accurate simulation and prediction of action. Secondly, the results suggest that such experience facilitates more accurate state estimation for actions perceived in the 1PP which map more closely onto visual input of self-generated action. More forward prediction error is retained for 3PP viewed actions, which may however have the benefit of compensating for the uncertainty involved in interacting with others.

INTRODUCTION

Anticipation of events in our environment is essential for us to respond with an appropriate movement within an adequate time period. For instance, stepping into an elevator or stepping onto a moving escalator require prediction of the upcoming dynamics of the doors or the stairs to estimate the most appropriate time to step forward. In sport, we see this predictive estimation being performed regularly, as players make judgements about the motions necessary to return a serve in tennis, or prevent a goal in football. Likewise, the ability to generate predictions about other people's actions is crucial to enable us to expedite our own response to them – for our actions to coincide, cooperate, or compete.

However, generating a reactive response in a dynamically changing environment poses many problems. It takes more than 100ms for our movements to be updated as a result of somatosensory feedback (Flanders & Cordo, 1989) and approximately 150-200ms based on visual feedback (Miall, Weir, Wolpert, & Stein, 1993; Saunders & Knill, 2003). Yet we are able to generate responsive movements with great temporal and spatial precision. An explanation for this ability is that our nervous system anticipates the future state of our environment or of other people, thus enabling us to prepare a motor response in a timely manner. It is postulated that we are able to do so based on internal forward (predictive) models developed through our visual and motor experiences, which are updated based on prediction errors, i.e. the discrepancy between the expected outcome and the actual consequence experienced (Wolpert & Flanagan, 2001). Such predictive models allow the nervous system to compensate for the delays in sensory feedback due to neural conduction, receptor transduction, central processing, and noise distortion (Graf et al., 2007; Miall, 2003; Wolpert & Flanagan, 2001), and thus facilitate fine grained visuo-motor control. Such models draw on memory of experience, with accurate estimation (least prediction error) occurring when the motion is predictable and familiar.

As such, visual perception is arguably a predictive activity (Graf et al., 2007; Prinz, 2006; Wilson & Knoblich, 2005) and there is much evidence to date to suggest that motion perception is indeed so. *Representational momentum*, for instance, is the phenomenon whereby static images depicting or implying motion, are perceived as being further forward in their trajectory than is presented; that is, the trajectory of the movement is extrapolated forwards in time and/or space. Representational momentum has been demonstrated in a range of conditions from static images inferring motion (Bertamini, 1993; Freyd, 1983; Reed & Vinson, 1996) to video images of full body biological motion in point light actors (Jarraya, Amorim, & Bardy, 2005). Researchers have suggested that the representational momentum phenomenon reflects the anticipated direction of movement, rather than the perceived actual movement (Hubbard & Bharucha, 1988), and as such reflects predictive processes in perception.

As with representation momentum, the trajectory extrapolation effect has also been detected when observing body movements that match future states. That is, priming effects were only observable for test postures that had been primed with an action that would have resulted in the test pose had the priming action sequence continued further. Such trajectory extrapolation effects have been reported in priming studies involving animated characters (Verfaillie & Daems, 2002), point-light actors (Verfaillie, 1993) and rotating human body movement, but not for actions violating normal human bodily movement (Kourtzi & Shiffrar, 1999).

Oculomotor studies have also provided support for the predictive property of action perception. Flanagan & Johansson (2003) found that when observing another person perform a block stacking task, observers' eye gaze was very similar to the eye gaze-hand coordination produced when performing the same task themselves, and importantly, it maintained a predictive rather than reactive pattern. Similar predictive effects have also been observed in 12 month old infants for goal oriented actions they can perform (Falck-Ytter, Gredebäck, & von Hofsten, 2006). The studies suggest that when observing someone perform an action, eye motor programmes are produced that

are directed by motor representations for action execution, being predictive, not reactive, which would suggest we use eye gaze in a manner appropriate for motor planning and control.

Such a prediction system is said to be underpinned by activity within the action-observation network (AON), a network of brain regions that become active not only when we perform an action ourselves but also when we observe an action being performed by another (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). It is suggested that these shared neural representations employ our own motor system when simply observing an action, allowing us to internally emulate or simulate that action (Gallese, 2005; Grush, 2004; Jeannerod, 2001). Several authors claim that this shared representation system evolved in primates as a means to facilitate social interaction (Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004) – the ‘adaptation hypothesis’ - allowing us to embody others’ expressions and actions, and aiding prediction of other people by drawing on our own motor repertoire (Blakemore & Frith, 2005; Vittorio Gallese & Goldman, 1998; Iacoboni et al., 2005).

A somewhat different perspective postulates that the AON functions not as a consequence of evolution for perception and prediction of others. Instead, it suggests execution-observation common coding has resulted through Hebbian or associative learning, as a consequence of visual input largely co-occurring with motor output (Cook, Press, Dickinson, & Heyes, 2010; Heyes, 2010; Heyes, 2001) – the ‘associative hypothesis’. Advocates of this theory suggest that the system’s principal function is not social, in that it does not have the specific purpose of facilitating action interpretation in others. Rather, the system may have a self-related purpose, to facilitate visuo-motor control.

The etiology and function of the AON are thus debated. Despite this, few studies have attempted to disassociate systems potentially serving a self-oriented predictive function, as opposed to a social, other-oriented function. That is, many studies that have examined action observation and/or prediction have used stimuli that can only be interpreted as others performing an action. Such presentation fails to disentangle differing mechanisms that may underpin action observation and

prediction for actions from different perspectives, which have been associated with self or other (Oberman & Ramachandran, 2008; Chan et al., 2004; Ehrsson et al., 2004; Jarraya et al., 2005; Saxe et al., 2006). If conclusions about the function of shared representations are to be drawn with regards to predictive mechanisms for self or for social purposes, it is essential to attempt to disentangle actions and facilitatory mechanisms relating to self and other.

One method to attempt to differentiate self and other related processes is to utilise stimuli presented in the first person perspective (1PP, egocentric) and third person perspective (3PP, allocentric). This logic was postulated by Oberman and Ramachandran (2008), with the assumption that 1PP presented actions closely resemble the actual visual feedback of performing an action oneself, whereas 3PP stimuli capture the visual input of observing someone else. If a mechanism has emerged for the monitoring of one's own actions, predicting the consequences of an action would be most effectively achieved (that is, have the least prediction error) in the first person perspective. In sharp contrast, if the mechanism has emerged for the social reason of interpreting someone else's actions, action prediction would be most effective in the third-person perspective (Bach, Fenton-Adams, & Tipper, 2014; Oberman & Ramachandran, 2008).

Whilst previous work has largely examined the prediction of consequences of observed actions (e.g. Aglioti, Cesari, Romani, & Urgesi, 2008; Urgesi, Savonitto, Fabbro, & Aglioti, 2012), few studies have explored the temporal dimension of action prediction. Examining temporal aspects of action observation allows us to determine whether the motion perception system is temporally predictive, potentially enabling us to respond to our environment in a timely manner. The current study examines the ability to determine future action states, and specifically the temporal acuity of prediction, via a spatio-temporal judgement task. The technique is adapted from an action occlusion paradigm which has previously been used to explore temporal prediction of actions using pointlight images either through explicit (Parkinson, Springer, & Prinz, 2012; Sparenberg, Springer, & Prinz, 2012; Prinz & Rapinett, 2008) or implicit means (Parkinson, Springer, & Prinz, 2011; Graf et al., 2007).

However, some of these studies have combined transitive object-directed and intransitive non-object related, and familiar and unfamiliar actions, and have found inconsistent results, not always in accordance with the notion that perception is predictive (cf. Sparenberg et al., 2012). However, given that many single cell and neuroimaging studies would indicate that the MNS/AON is particularly activated for goal-directed/object-oriented actions (for monkey mirror neuron firing: Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Iacoboni et al., 2005; Umiltà et al., 2001; for human AON activation: Rizzolatti et al., 1996; Johnson-Frey et al., 2003; Shmuelof and Zohary, 2005) it would be beneficial to examine the temporal accuracy of prediction related specifically to transitive actions. Here we present high quality video images of simple everyday object related actions, such as picking up a wine glass, to examine temporal prediction of action. In Experiment 1, we examine potential differences in a spatio-temporal judgement task when observing actions in 1PP or 3PP conditions. In subsequent Experiments 2 to 4, we determine the effects of visual and motor priming on temporal prediction acuity for 1PP and 3PP actions.

EXPERIMENT 1

METHOD

Participants

Thirty-two participants (13 male) were recruited through the University of York, Department of Psychology participant pool. Participants were excluded if their point of subjective equality (PSE) value was beyond the stimulus range presented in the experiment (between -200 and 200ms) indicating their sensitivity was too low for the present study to determine it accurately. This criterion rejected 8 participants; data from the remaining twenty-four participants (15 female, 9 male) were analysed. Included participants were between 18 and 28 years of age ($M=20.5$, $SD=2.2$) and had normal or corrected-to-normal vision. They received a small payment or course credit in

compensation for their time. Procedures were approved by the Department of Psychology Ethics Committee at The University of York.

Materials

Stimuli were recorded using a Panasonic full HD 1920x1080p (50p) Camcorder, mounted on a tripod. Action sequences were recorded from a bird's eye perspective, looking down onto a table covered with a black cloth. Each action sequence recorded comprised a transitive action involving an actor's forearm and hand reaching towards an object on a table, grasping it, lifting it and removing it from the table in the direction of the actor. Action sequences were recorded with five different objects: a mug, a wine glass, a hammer, a pair of scissors and a jug of water. Each action was performed by a male and by a female actor using their right hand, producing ten template sequences in the first person perspective; that is, the hand appearing from the inferior edge of the screen.

Recordings were edited using Sony Vegas Pro v10, which was used to trim the recordings so that each action sequence began with 300ms prior to the action beginning (i.e. prior to the hand appearing and moving towards the object) and 300ms post-completion of the action. Actions were rendered at 60Hz. Full action sequences (without occlusion) had a duration of 2383ms minimum and 3567ms maximum ($M = 2754.9\text{ms}$).

From these template sequences in the 1PP, 3PP action sequences were created by flipping the image about the horizontal axis (see Figure 1). 3PP actions were therefore a mirror-match, rather than an anatomical-match, of the 1PP actions. This manipulation was based on studies demonstrating that we imitate mirror-matched actions better than anatomically matched actions (Chiavarino, Apperly, & Humphreys, 2007) and that whilst perspective and handedness are important in sense of ownership, peri-dominant-hand space is the preferred reference frame to distinguish other from self (Conson, Aromino, & Trojano, 2010; Conson, Alaerts, Heremans, Swinnen, & Wenderoth, 2009; Mazzarella, & Trojano, 2009). Therefore, in experimental studies examining action observation mechanisms,

mirror-matched limb actions in the allocentric perspective (3PP) would be preferable over the use of anatomically-matched actions, which would have an additional processing disadvantage¹.

Presentation of visual stimuli and response recording was programmed using Python v2.7 (www.python.org) and Pygame modules (www.pygame.org). Visual stimuli were presented on a 1280x1024 monitor, with 1024x576 image dimensions on a black surround, at a viewing distance of 64cm, subtending a visual angle of 28x15 degrees, at a 60Hz screen refresh rate.

An occlusion image was created using a grey-scale average of a frame from a video prior to the start of the action (eg. a cup on the table top without the hand in shot was averaged to grey-scale).

Insert figure 1 around here

Design and Procedure

Experimental trials began with 1000ms of an action sequence video. The action video was then replaced by the grey-scale occlusion image, presented for a fixed duration of 500ms. The occluder was then followed by the continuation of the action video until the end of the sequence. The point from which the action sequence continued post-occlusion was manipulated in 13 time steps, with the time between the last frame pre-occlusion and the first frame post-occlusion ranging from 300ms to 700ms in steps of 33ms (2 frames). This produced 13 offset periods between -200 and 200ms. Thus, an offset period of zero corresponded to a coherent trial, whereby the post-occlusion - pre-occlusion frame difference temporally corresponded exactly to the duration of the occlusion (500ms).

¹ Bach, Fenton-Adams, & Tipper (2014) investigated observation of anatomical- and mirror-matched action observation and found no evidence for differences between these 3PP conditions.

Participants were instructed to imagine the continuation of the action during the occlusion period. Once the action continued after occlusion, participants were requested to respond with a key press to indicate whether the point from which the action continued was temporally earlier or later than the actual action would be expected to be given the occlusion period and their expectation. Participants were asked to respond whilst the action was still continuing and before completion of the sequence when the screen went to black. Task responses were given with the right hand. A self-paced inter-trial interval was employed, requiring the participant to press the spacebar to progress to the next trial. Figure 2 depicts an experimental trial.

An 'earlier' trial refers to a negative temporal offset (<0); that is, the time between the last frame pre-occlusion and the first frame post-occlusion is less than the occlusion period ($<500\text{ms}$). A 'later' trial refers to a positive temporal offset (>0); the time between the last frame pre-occlusion and the first frame post-occlusion is greater than the occlusion period ($>500\text{ms}$).

The experiment began with a *familiarisation phase*, and then a *practice phase*, before proceeding to the main *experimental phase*. In the familiarisation phase, participants viewed each template video (i.e. without occlusion) twice (once in each perspective), and were asked to simply observe. Twenty practice trials (10 in each perspective) were then presented, allowing participants to adjust their judgement and response speed. Trials presented in this practice phase were the temporal offset extremes (offsets of -200ms and 200ms) for each object in each perspective (gender of the hand was randomised). 'Correct/Incorrect' feedback was provided to participants at the end of each practice trial.

The main experimental phase consisted of 390 trials in each of the 1PP and 3PP conditions (780 trials in total) (5 actions \times 13 offset periods \times 6 repetitions), divided into ten equal blocks. The 1PP and 3PP conditions were presented in alternate blocks and were counterbalanced between participants. Trials that were not responded to within the given response time were repeated once at the end of the block. Each block began with 10 familiarisation trials in the perspective (1PP/3PP) corresponding to

that block. Each block ended with a 'Break' screen allowing participants to rest their eyes and to proceed to the next block when ready to do so.

The temporal offset was varied using the method-of-constant-stimuli. Trials were presented in a pseudo-random order, ensuring that the same object or temporal offset was never presented three times consecutively. From the proportion of 'late' responses per stimulus offset condition, logistic psychometric functions were fitted to each individual's data to estimate their point of subjective equality (PSE).

Insert Figure 2 around here.

RESULTS & DISCUSSION

Data were collapsed across object and actor gender and analysed in terms of temporal offset (test motion) and perspective (see Figure 1).

Firstly, participants' proportion of 'late' responses across the 13 temporal offset conditions were analysed in a one-way ANOVA. Data for both the 1PP and 3PP conditions showed a statistically significant effect of temporal offset ($F(12,276) = 66.67, p < .001, \eta_p^2 = .74$ for the 1PP; $F(12,276) = 85.83, p < .001, \eta_p^2 = .79$ for the 3PP). Importantly, the linear trend was significant for both 1PP and 3PP conditions ($F(1,23) = 225.33, p < .001, \eta_p^2 = .91$; $F(1,23) = 316.81, p < .001, \eta_p^2 = .93$ respectively). This indicates that participants could perform the task, as data show a significant linear increase in the proportion of late responses with increasing temporal offset.

For each individual participant's data, a logistic function was fitted to the proportion of 'late' responses for each temporal offset condition by minimising the mean squared error of the fit. The number of free parameters was determined per individual by nested modelled F-tests as some

participants had a high lapse rate. This determined the main parameter of interest: the point of subjective equality (PSE). See Figure 3 for an example individual fitted psychometric function.

Insert Figure 3 about here.

Temporal offsets were determined relative to the occluder duration. The PSE determines the temporal offset that would elicit an 'early' / 'late' response probability of .5, and is a determinant of the subjective experience of a 'real-time' outcome. A negative PSE value thus indicates that a condition in which less video had been removed than the occluder time was perceived as being coherent with the occluder, and would in turn indicate that mental simulation was progressing slower than the real action itself. A positive PSE indicates that a 'later' condition was perceived as being coherent with the occluder, and that the mental simulation was progressing faster than the real action.

An average PSE was determined across participants for each perspective. For the 1PP condition, an average PSE of 19.29ms (SE=5.05) was observed, which was not significantly different from zero ($p = .095$). For the 3PP condition, an average PSE of 37.29ms (SE=5.05) was observed, which differed significantly from zero ($t(23) = 3.069$; $p = .005$). A repeated measures t-test demonstrated that there was no significant difference between the PSEs for the two perspective conditions ($t(23) = -1.78$; $p = .09$) (see Figure 4, Experiment 1).

Insert Figure 4 about here.

Before interpreting the data it is worth discussing what PSE value would constitute optimal performance; that is, whether a PSE close to zero or a significantly positive PSE is optimal for engagement with the environment and/or motor control. It is important that we are able to anticipate an action outcome and that our internal simulation of that action continue such that predictions of future states can guide action. However, it is critical for visuomotor performance that

this prediction be as accurate as possible, where the forward model and reality match. Representational momentum studies demonstrate that there is an automatic simulation process of future states. However, such tasks require participants to recall from memory the last stimulus seen and in doing so participants make errors in reporting stimuli in loci they would have moved to. In sharp contrast, our task presents stimuli that have continued along their action trajectory and the task is to identify their accurate action state. In this situation, performance closer to zero is closer to reality (i.e. there is less prediction error between the simulated action and the real action dynamics). Hence, subsequent motor processes directed towards the stimulus would be more accurate. We therefore posit that a PSE closer to zero is indicative of optimal performance.

Given the above premise, there are two aspects of the data to be discussed. First, action simulation processes appear to be more rapid than the actual observed actions. This is an interesting contrast to previous similar studies that have shown that action simulation processes run more slowly than the actual action (e.g., Prinz & Rapinett, 2008; Sparenberg et al., 2012). At this stage we cannot be certain why this contrast emerged, but note a number of interesting methodological differences that might mediate the difference in results. Firstly, previous studies have focussed on whole-body complex actions, combining transitive and intransitive actions. Given that many single cell and neuroimaging studies would indicate that the MNS/AON is particularly activated for goal-directed/object-oriented actions (e.g. Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti et al., 1996) it may be the case that lack of goal/intention in intransitive conditions influences the simulation processes on which temporal judgements are made (Gallese & Goldman, 1998; Jeannerod, 2001). Secondly, previous studies have used action stimuli with varying degrees of familiarity to participants. Activities such as throwing a basketball or leapfrogging, as used in these previous studies, may arguably be recently *unfamiliar* to participants. The stimuli in the Sparenberg et al. (2012) study therefore do not differentiate between actions for which participants can and cannot easily draw upon their own motor repertoire to enable simulation/prediction mechanisms. Thirdly, the use of point light actions (Parkinson, Springer, & Prinz, 2012; Sparenberg et al., 2012;

Parkinson, Springer, & Prinz, 2011; Graf et al., 2007;) may also influence the nature of timing judgements. Finally, researchers have observed that showing less information post-occlusion degrades performance in occlusion tasks (Parkinson et al., 2012). As each of the occlusion studies described present a simple static image post-occlusion on which to base a decision, it might be argued that performance accuracy is not as would be expected if actions were allowed to continue until their completion, as in our present study.

As noted, our stimuli contrast on all these dimensions. They are all highly familiar and simple transitive reach-to-grasp actions, which are presented in high quality full colour videos, and the display for judgement is a continuation of the action rather than a static image. Which of these stimulus properties mediates the change in temporal judgement decisions is not clear at this time. Certainly, however, the stimulus properties that shift temporal simulations from slower to faster than the actual observed action are certainly worthy of further study.

The central focus of this study was to examine contrasts between 1PP and 3PP views of action. As a reminder, and in accordance with the logic proposed by Oberman & Ramachandran (2008), an account arguing that simulation of observed actions evolved as a process to facilitate the understanding of other people might predict better spatio-temporal judgements when viewing 3PP actions. In contrast, an account arguing that simulation processes emerged to predict one's own action states might predict more accurate performance when viewing 1PP actions. The results of this first study are equivocal. At most, based on this first study, it would appear that the prediction system is not viewpoint dependent, and that there is an anticipation bias whereby we simulate actions faster than real world dynamics and thus overestimate the action state in spatio-temporal terms.

Although there is a trend for the 1PP viewing condition to be somewhat more accurate than 3PP, this is not significant in this study. The lack of consistent contrast between 1PP and 3PP has certainly been observed in other situations, for example, in behavioural studies (Bach et al., 2014) and fMRI

studies, where common neural population codes for producing and observing action are observed in parietal cortex in both 1PP and 3PP views (e.g., Oosterhof, Tipper, & Downing, 2012). However, in both of these studies other situations did detect contrasts between 1PP and 3PP. For example, when more complex aspects of action had to be integrated, Bach et al., (2014) observed faster responses when observing 1PP than 3PP actions, and in fMRI ventral premotor cortex (PMv) encoded joint production and perception of actions only in the 1PP (Oosterhof et al, 2012).

Therefore, in the following experiments we investigate other experimental situations that might reveal contrasts between 1PP and 3PP actions, as well as improve performance. In Experiment 2, via visual priming manipulations, we attempt to facilitate the processing of 3PP actions.

EXPERIMENT 2

Many of the occlusion studies discussed above found that prediction performance was significantly impaired when the whole body actions were inverted (Graf et al., 2007; Parkinson et al., 2011; Sparenberg et al., 2012). Under the assumption that humans have no visual experience of inverted movements and cannot draw on their motor repertoire to simulate such actions (Loula, Prasad, Harber, & Shiffrar, 2005), the authors of these studies concluded that simulation of biological motion requires more than trajectory extrapolation alone. Several studies have examined how visual experience is sufficient to influence how we perceive or predict actions. A body of research has demonstrated how individuals can acquire motor skills or modify behaviour based on outcomes through observational learning (Buchanan & Wright, 2011; Chamley, 2003; Carrol & Bandura, 1982). Furthermore, perceptual experience alone has been found to influence how we perceive other's actions. Cross, Stadler, Parkinson, Schütz-Bosbach, & Prinz (2011), for instance, found that temporal prediction of actions improved consistently after observational training of gymnastic moves. In neuroimaging, Higuchi, Holle, Roberts, Eickhoff, & Vogt (2012) demonstrated that activation

intensities in prefrontal areas of the AON during observational practice of guitar chord playing predicted the behavioural effects on chord playing. These studies suggest involvement of motor areas, and emergence of motor resonance processes, through observational experience alone.

METHOD

Participants

Thirty-three participants (9 male) were recruited through the University of York, Department of Psychology participant pool. As in Experiment 1, participants were excluded if their point of subjective equality (PSE) value was beyond the stimulus range presented in the experiment. Nine participants were excluded on this basis and subsequent analyses were conducted on the remaining 24 participants (16 female, 8 male). All were aged between 18 and 33 years of age ($M=22.3$, $SD=3.7$) and had normal or corrected-to-normal vision. They received a small payment or course credit in compensation for their time. Procedures were approved by the Department of Psychology Ethics Committee at The University of York.

Materials & Procedure

The same stimuli and paradigm were used as in Experiment 1. In Experiment 2, after the familiarisation phase and before the practice phase of the main computer based task, participants observed the experimenter perform the transitive actions on the same objects from across the table using their left hand (thus replicating the image of the action flipped about the horizontal axis as in the 3PP condition). Each transitive action was performed 4 times on each object. The same main experimental computer task as in Experiment 1 was then performed by participants.

RESULTS & DISCUSSION

As in Experiment 1, all participants produced monotonically increasing proportions of 'late' responses with increasing temporal offset for both perspective conditions, indicating that they could perform the task well.

From the logistic functions fitted to each individual participant's data, an average PSE was determined across participants for each perspective. For the 1PP condition, an average PSE of 23.55ms (SE=4.39) was observed, which was not significantly different than zero ($p = .091$). For the 3PP condition, an average PSE of 29.43ms (SE=4.39) was observed, which differed significantly from zero ($t(23) = 2.240$; $p = .035$). However, a repeated measures t-test demonstrated that there was no significant difference between the PSEs for the two perspective conditions ($t(23) = -.671$; $p = .509$) (see Figure 4, Experiment 2). A combined analysis of Experiments 1 and 2 also failed to detect any differences between the baseline Experiment 1 and the visual 3PP action priming condition of Experiment 2 on the PSE for 1PP or 3PP actions.

The results of this experiment are therefore very clear. Although participants received real world visual experience of another person reaching out and grasping the objects in the same manner as subsequent 3PP visual displays, there was no effect on spatio-temporal judgements. As noted, it was ensured that the viewed priming actions were a close match to the 3PP actions in the computer task, and hence it was predicted 3PP performance would improve. The results are unable to provide support for an account claiming the simulation process is primarily involved in the prediction of another person's actions.

However, a counter argument may be that our priming events were limited, certainly much reduced compared with prior studies (e.g., Cross et al, 2011). Nevertheless, in these latter studies, participants were learning much more complex and unfamiliar actions such as gymnastic movements. For typical participants there would be no prior motor representations of these complex and difficult actions. In contrast, we would argue that the very simple everyday reach-to-grasp actions in our study would benefit from a few exposures, if visual exposure is effective in temporal

judgement tasks. The following two experiments manipulate motor experience with a similar level of four exposures to each stimulus. As will be seen, with this similar level of exposure, contrasts between 1PP and 3PP now emerge. Hence the studies to be reported suggest large amounts of exposure are not necessary when everyday actions are simple and very familiar, and furthermore, motor representations are more influential than prior visual exposure on temporal judgements of viewed actions.

EXPERIMENT 3

The importance of prior motor experience in the simulation of action has previously been demonstrated. For example, Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard (2005) conducted fMRI in expert ballet dancers, expert capoeira dancers and inexpert controls whilst they watched ballet or capoeira dance sequences. Greater activation was observed in areas of the AON when participants watched dance in their expertise. In a later study, Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard (2006) utilised the fact that male and female ballet dancers have different motoric dance experiences but have equal levels of visual familiarity of ballet moves. The male and female ballet dancers observed gender-specific ballet moves whilst being scanned in fMRI. They found increased activity in the AON when participants observed action within their motor repertoire compared to the actions they frequently observed but of which they had no physical experience. The authors concluded that the AON responds with motor experience over and above visual experience and that this system integrates observed actions with movements in one's own motor repertoire (see also, Cross, Hamilton, & Grafton (2006); Knoblich & Flach (2001) for similar observations). Hence these studies demonstrate that action simulation mechanisms are enhanced by the capability of performing an observed action, and that we integrate information from motor experience with current visual input, which can facilitate predictive mechanisms to determine action outcomes.

Therefore, in the current study we provide motor experience prior to the spatio-temporal judgement task. That is, participants reach to grasp and pick up each of the 5 objects in the study 4 times. We predicted two possible outcomes. First, that there might be a general improvement in temporal judgement accuracy for all forms of viewed action after motor priming. That is, both 1PP and 3PP spatio-temporal judgements are more accurate. However, second, because the motor priming is by necessity an egocentric 1PP activity, temporal judgements of 1PP actions may differ from those of 3PP actions.

METHOD

Participants

Thirty-two participants (12 male) were recruited through the University of York, Department of Psychology participant pool. As in Experiment 1 and 2, participants were excluded if their point of subjective equality (PSE) value was beyond the stimulus range presented in the experiment. Eight participants were excluded on this basis and subsequent analyses were conducted on the remaining 24 participants (13 female, 11 male). All were aged between 18 and 31 years of age ($M=22.5$, $SD=3.4$) and had normal or corrected-to-normal vision. They received a small payment or course credit in compensation for their time. Procedures were approved by the Department of Psychology Ethics Committee at The University of York.

Materials & Procedure

The same stimuli and paradigm were used as in Experiments 1 and 2. In Experiment 3, however, after the familiarisation phase and before the practice phase, participants performed the transitive actions themselves, with eyes open. Participants were asked to observe the experimenter place the object on a table covered with a black cloth, on a marker set at 34cm from the participant's edge of the table. Participants were asked to begin each action with their hands off the table, and then reach

towards, grasp and remove each object from the table as if they were about to use the object and in a manner that replicated as precisely as possible the action just observed (eg. grasp the wine glass by placing the fingers on the stem of the glass to lift the object), such that participants received motor experience that was specific to the actions they would judge. On completion of each action (when the object was removed from the table completely towards the participant), the experimenter took the object from the participant and placed it back on the marker ready for the next action. Each action was performed with the participant's right hand, four times on each of the five objects. The same main experimental computer task as in Experiment 1 was then performed by participants.

RESULTS & DISCUSSION

As in the previous experiments, all participants produced monotonically increasing proportions of 'late' responses with increasing stimulus interval offsets, demonstrating that they could perform the task well.

From the logistic functions fitted to each individual participant's data, an average PSE was determined across participants for each perspective. For the 1PP condition, an average PSE of -3.26ms (SE=4.87) was observed, which did not differ significantly from zero ($p = .790$). For the 3PP condition, an average PSE of 26.84ms (SE=4.87) was observed, which was marginally greater than zero, but did not reach significance ($p = .083$). A repeated measures t-test demonstrated that there was a significant difference between the PSEs for the two perspective conditions ($t(23) = -3.093$; $p = .005$, $d = .63$) (see Figure 4, Experiment 3).

This is the first of this series of studies to detect a significant contrast between 1PP and 3PP viewed action when judging temporal aspects of action. It appears that prior motor experience is important for facilitating performance when viewing 1PP actions. This contrasts with the findings of Experiment 2 where prior visual experience of 3PP actions had no significant effect on performance.

The results of Experiments 2 and 3 are in line with previous observations concerning the role of vision and motor experience on action simulation processes. For example, Aglioti et al. (2008) and Urgesi et al. (2012) examined action anticipation in basketball and volleyball players respectively and compared this with supporters who had similar visual experience. Only those with prior motor representations could accurately judge the future states of throws and serves based on viewed players' body kinematics. Furthermore, participants given physical practice training improved their abilities to predict actions by reading body kinematics. Those assigned to observational training alone improved only in understanding the ball trajectory (Urgesi et al., 2012). Our findings are aligned with this prior work showing motor experience is also required to temporally predict bodily actions. However, our results go further by demonstrating that motor experience can selectively facilitate temporal judgements of actions viewed from a first person perspective.

EXPERIMENT 4

A further issue concerns whether the effects produced in Experiment 3 are purely due to action priming. Note that when practising the reach and grasp actions participants were able to observe their actions, as would be expected in such a naturalistic situation. Therefore, it is unclear whether the difference between 1PP and 3PP detected in Experiment 3 is due to motor priming or whether it is caused by a combination of motor priming and concomitant visual input from the 1PP.

Experiment 4 therefore adopts an approach similar to that of Casile & Giese (2005). They assessed the visual recognition of gait before and after non-visual motor training, in which participants were blindfolded whilst learning a novel upper-body coordinated movement. The study showed a selective improvement in recognition of subsequently observed actions that were physically trained compared to novel actions, despite participants having no visual input during the training phase. The Casile and Giese (2005) study therefore demonstrated that motor experience, without visual

experience, has a direct influence on action recognition. We now examine whether pure motor experience when participants are unable to observe their own action also has a facilitation effect when making temporal judgements and whether this is selective for 1PP actions.

METHOD

Participants

Thirty-six participants (16 male) were recruited through the University of York, Department of Psychology participant pool. As in all other experiments, participants were excluded if their point of subjective equality (PSE) value was beyond the stimulus range presented in the experiment. Twelve participants were excluded on this basis and subsequent analyses were conducted on the remaining 24 participants (12 female, 12 male). All were aged between 18 and 29 years of age ($M=22.9$, $SD=3.5$) and had normal or corrected-to-normal vision. They received a small payment or course credit in compensation for their time. Procedures were approved by the Department of Psychology Ethics Committee at The University of York.

Materials & Procedure

The same stimuli and paradigm were used as in Experiments 1, 2 and 3. In Experiment 4, after the familiarisation phase and before the practice phase, participants performed the transitive actions themselves. Participants were instructed to reach and grasp each object as in Experiment 3, such that they received motor experience specific to the actions to be judged, but *with eyes closed*. On completion of each action (when the object was removed from the table completely towards the participant), they were instructed to open their eyes when the object was again placed in its usual location. Participants were then able to look at the object until they were ready to grasp it, at which point they closed their eyes and completed the reach-to-grasp action. These self-paced actions were

performed with the participant's right hand, four times on each of the five objects. The same main experimental computer task as in Experiment 1 was then performed by participants.

RESULTS & DISCUSSION

As in all previous experiments, all included participants produced monotonically increasing proportions of 'late' responses for increasing temporal offsets, indicating they could perform the task well in both perspective conditions.

From the logistic functions fitted to each individual participant's data, an average PSE was determined across participants for each perspective. For the 1PP condition, an average PSE of -14.01ms (SE=6.26) was observed, which did not differ significantly from zero ($p = .397$). For the 3PP condition, an average PSE of 18.46ms (SE=6.26) was observed, which also did not differ significantly from zero ($p = .134$). Importantly, a repeated measures t-test demonstrated that there was a significant difference between the PSEs for the two perspective conditions ($t(23) = -2.594$; $p = .016$, $d = .53$) (see Figure 4, Experiment 4).

Experiment 4 examined the effect of motor priming when there was no concurrent visual input. It has essentially confirmed the findings of Experiment 3. Indeed, a combined analysis of Experiments 3 and 4 showed a highly significant main effect of 1PP-3PP viewpoint ($F(1,46) = 15.57$, $p < .001$, $\eta_p^2 = .25$), but no main effect of experiment ($F(1,46) = .28$, $p = .60$) or interaction ($F(1,46) = .02$, $p = .88$). Thus, when participants have some motor experience of the actions to be subsequently judged there is a significant contrast between 1PP and 3PP viewed actions. That is, 3PP actions are simulated at a faster rate than 1PP actions and the latter tend to be somewhat more accurate.

Combined analysis across experiments:

Figure 4 summaries the overall data pattern in the spatio-temporal judgement tasks. We have noted, in line with previous research, the critical importance of motor experience in the simulation of actions. To further extract the pattern of performance across experiments we performed a

combined analysis. This examined the effects of perspective (1PP vs 3PP) and the effects of motor priming (no motor priming vs motor priming). The latter involved the combination of Experiments 1 and 2 in the no motor priming condition and Experiments 3 and 4 in the motor priming condition. This 2 (perspective) x 2 (yes/no motor priming) ANOVA demonstrated no significant effect of motor priming. However, of central importance, there was a significant effect of perspective ($F(1,94) = 17.61, p.001, \eta_p^2=.16$). That is, in all studies the spatio-temporal prediction performance in the 1PP condition tended to be more accurate than the 3PP condition, with the latter 3PP revealing simulation processes tending to be faster than the observed actions. However, this main effect of perspective was modified by a marginally significant interaction between perspective and yes/no motor priming ($F(1,94) = 3.53, p=.06, \eta_p^2=.04$). Planned comparisons revealed that the 3PP was not significantly influenced by whether participants experienced motor priming prior to the task ($t(94) = .79, p=.43$). However, in sharp contrast, there was a significant change in the accuracy of the 1PP temporal judgements, with significantly slower motor simulations after motor priming ($t(94)=2.13, p=.034, d= .44$). Figure 5 depicts these differences in the temporal prediction of action in 1PP and 3PP with and without motor priming.

Insert Figure 5 around here.

DISCUSSION

There are a number of conclusions that can be drawn concerning the function and mechanisms of the action simulation processes, although our conclusions concerning the etiology of the action observation network remain limited. The adaptation theory of the AON suggests that we recruit motor regions to create motor simulations of observed actions as a means to anticipate and interpret the actions of other people. That is, the account suggests the AON is a mechanism underlying action prediction that has evolved to facilitate social interaction (Rizzolatti & Arbib, 1998;

Rizzolatti & Craighero, 2004). By this account we reasoned that when viewing 3PP actions that more closely match what would be perceived when observing another person, action timing judgements might be more accurate than when viewing 1PP actions. In all experiments we have found no evidence for this pattern. Indeed, even when 3PP actions were visually primed with real world reach and grasp actions that equated to the actions subsequently observed, 3PP actions gained no advantage (Experiment 2). In contrast, 1PP action timing judgments tended to be more accurate.

Although 3PP action prediction remained less accurate we do not rule out a role of the AON in the prediction of other people's actions. Indeed the forward prediction error observed for 3PP actions, indicating the simulation is further along the motion path, would enable compensation for any unpredictability in the actions of other people. In essence, as when intercepting a moving object over which we have no control, the motor system has to predict future states to place effectors appropriately for the object's arrival; then similarly, as we never have control over the final properties of another person's actions, forward prediction error would enable adjustments to be made. Thus, for an unfamiliar agent or action it would be more advantageous to slightly over anticipate, and we note the forward errors are small ($< 40\text{ms}$) enabling efficient interpersonal interactions.

However, it is worth noting that our observation that the simulation process for 3PP actions appeared to be somewhat faster than the actual observed action contrasted with previous similar studies (Sparenberg et al., 2012). That is, in the Sparenberg et al. (2012) study, motor simulation appeared to be slower than the actual actions. They additionally manipulated the occlusion duration and found the same negative bias (slower simulation than real action) for the different occluder durations, suggesting a stable time error. As a means to reconcile their findings with previous studies that suggest perception is predictive, the authors suggested this error may be due to a switching cost from a perception to a simulation mechanism. However, we suggest that it is worth considering the role of stimulus type, whether point-light images, transitive vs intransitive, whole body vs single hand

grasps, familiar vs unfamiliar, and whether the stimulus to be judged is static or a continuation of the action. As we have revealed with contrasts in viewpoint (1PP vs 3PP), further study of stimulus type may provide new insights into when and why action simulation processes are faster or slower than reality, and in turn when perception or simulation is necessarily predictive.

In contrast to the 3PP data, 1PP action prediction tended to be more accurate (closer to zero) in all conditions. Following the logic of Oberman and Ramachandran (2008) and that the 1PP more closely matches the visual input when performing an action oneself, the results support the notion that action observation systems emerged to guide one's own actions. At the least, the results would suggest that the system is less attuned to serving a social function as it is to providing an accurate state estimation for 1PP actions.

Of note, and in support of the idea that the system emerged for self action prediction, is that the contrast between 1PP and 3PP was most evident after motor priming, where participants received brief experience of grasping and lifting the objects in a manner similar to what would subsequently be observed (Experiments 3 & 4). It is striking that only minimal motor priming (4 reaches to each object) was sufficient to improve timing judgements. Our findings are in line with the literature clearly showing that it is prior motor experience/skill that influences motor simulation processes (e.g., Calvo-Merino et al., 2005) and that action prediction based on body kinematics is improved following motor practice (Urgesi et al. 2012). In particular, motor priming without any visual input (Experiment 4) is able to facilitate 1PP action judgements, and this result supports that of Casile & Giese (2005). The lack of contrast between Experiment 3, in which there was both visual and motor priming, and Experiment 4 in which there was only motor priming, might suggest that our effects were purely motor and vision played little role. However, we are cautious concerning this conclusion, as the condition examining purely visual observation of 1PP action without motor output was not possible in our current design. Thus it may yet be demonstrated that visual and physical experience may provide complementary and mutually reinforcing benefits to action prediction.

In sum, this series of experiments has demonstrated more than one property of the AON. The experiments examined the ability to make spatio-temporal judgements about future states of observed action. The ability to develop internal forward models of action and run an action simulation process is important to enable accurate visuomotor processes. It appears we have an anticipatory bias when observing actions of which we have little or no first-hand experience, but which reduces as we gain physical experience of these actions. Moreover, prediction differences can be observed between viewing action that closely matches our visual experience when performing an action (1PP) and that experienced when observing another person (3PP). The data confirm that motor priming has a selective effect on spatio-temporal prediction of 1PP rather than 3PP actions. This effect of motor priming, firstly, supports the theory that internal forward models of action and simulation processes draw upon one's own motor representations, and secondly, supports the notion that the major role of such action simulation and prediction processes, and their underlying mechanisms (the AON), is to facilitate one's own actions. This said, such action simulations can be extended to enable similar computations when observing another person's actions, although they retain greater prediction error which may arguably have the social benefit of dealing with uncertainty and unpredictability of others.

REFERENCES

- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, 11(9), 1109–1116. doi:10.1038/nn.2182
- Alaerts, K., Heremans, E., Swinnen, S. P., & Wenderoth, N. (2009). How are observed actions mapped to the observer's motor system? Influence of posture and perspective. *Neuropsychologia*, 47(2), 415–22. doi:10.1016/j.neuropsychologia.2008.09.012
- Bach, P., Fenton-Adams, W., & Tipper, S. P. (2014). Can't touch this: the first-person perspective provides privileged access to predictions of sensory action outcomes. *Journal of Experimental Psychology. Human Perception and Performance*, 40(2), 457–64. doi:10.1037/a0035348
- Bertamini, M. (1993). Memory for position and dynamic representations. *Memory & Cognition*, 21(4), 449–57.
- Blakemore, S.-J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, 43(2), 260–7. doi:10.1016/j.neuropsychologia.2004.11.012
- Buchanan, J. J., & Wright, D. L. (2011). Generalization of action knowledge following observational learning. *Acta Psychologica*, 136(1), 167–78. doi:10.1016/j.actpsy.2010.11.006
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cerebral Cortex (New York, N.Y. : 1991)*, 15(8), 1243–9. doi:10.1093/cercor/bhi007
- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology : CB*, 16(19), 1905–10. doi:10.1016/j.cub.2006.07.065
- Casile, A., & Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology*, 16(1), 69–74.
- Chamley, C. P. (2003). *Rational Herds. Economic models of social learning*. New York: Cambridge University Press.
- Chan, A. W.-Y., Peelen, M. V., & Downing, P. E. (2004). The effect of viewpoint on body representation in the extrastriate body area. *Neuroreport*, 15(15), 2407–10.
- Chiavarino, C., Apperly, I. a., & Humphreys, G. W. (2007). Exploring the functional and anatomical bases of mirror-image and anatomical imitation: the role of the frontal lobes. *Neuropsychologia*, 45(4), 784–95. doi:10.1016/j.neuropsychologia.2006.08.007
- Conson, M., Aromino, A. R., & Trojano, L. (2010). Whose hand is this? Handedness and visual perspective modulate self/other discrimination. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, 206(4), 449–53. doi:10.1007/s00221-010-2418-9
- Conson, M., Mazzarella, E., & Trojano, L. (2009). Numbers are represented in egocentric space: effects of numerical cues and spatial reference frames on hand laterality judgements. *Neuroscience Letters*, 452(2), 176–80. doi:10.1016/j.neulet.2009.01.043

- Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology: Human Perception and Performance*, 36(4).
- Cross, E. S., Stadler, W., Parkinson, J., Schütz-Bosbach, S., & Prinz, W. (2011). The influence of visual training on predicting complex action sequences. *Human Brain Mapping*, 34(2), 467–86. doi:10.1002/hbm.21450
- David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., ... Vogeley, K. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *Journal of Cognitive Neuroscience*, 18(6), 898–910. doi:10.1162/jocn.2006.18.6.898
- Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science (New York, N.Y.)*, 305(5685), 875–7. doi:10.1126/science.1097011
- Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, 9(7), 878–9. doi:10.1038/nn1729
- Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, 424(6950), 769–71. doi:10.1038/nature01861
- Flanders, M., & Cordo, P. J. (1989). Kinesthetic and visual control of a bimanual task: Specification of Direction and Amplitude. *The Journal of Neuroscience*, 9(2), 447–453.
- Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, 33(6), 575–81.
- Gallese, V. (2005). Embodied simulation : From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, 4, 23–48.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain : A Journal of Neurology*, 119 (Pt 2), 593–609.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the mind-reading, 2(12), 493–501.
- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *NeuroImage*, 36 Suppl 2, T22–32. doi:10.1016/j.neuroimage.2007.03.017
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *The Behavioral and Brain Sciences*, 27(3), 377–96; discussion 396–442.
- Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, 5(6), 253–261.
- Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, 34(4), 575–83. doi:10.1016/j.neubiorev.2009.11.007
- Higuchi, S., Holle, H., Roberts, N., Eickhoff, S. B., & Vogt, S. (2012). Imitation and observational learning of hand actions: prefrontal involvement and connectivity. *NeuroImage*, 59(2), 1668–83. doi:10.1016/j.neuroimage.2011.09.021

- Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics*, 44(3), 211–21
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3), e79. doi:10.1371/journal.pbio.0030079
- Jarraya, M., Amorim, M.-A., & Bardy, B. G. (2005). Optical flow and viewpoint change modulate the perception and memorization of complex motion. *Perception & Psychophysics*, 67(6), 951–61.
- Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage*, 14(1 Pt 2), S103–9. doi:10.1006/nimg.2001.0832
- Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: interactions of perception and action. *Psychological Science*, 12(6), 467–72.
- Kourtzi, Z., & Shiffrar, M. (1999). Dynamic representations of human body movement. *Perception*, 28(1), 49–62. doi:10.1068/p2870
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology. Human Perception and Performance*, 31(1), 210–20. doi:10.1037/0096-1523.31.1.210
- Miall, R. C. (2003). Connecting mirror neurons and forward models. *Neuroreport*, 14(17), 2135–7. doi:10.1097/01.wnr.0000098751.87269.77
- Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a smith predictor? *Journal of Motor Behavior*, 25(3), 203–16. doi:10.1080/00222895.1993.9942050
- Oberman, L. M., & Ramachandran, V. S. (2008). How do shared circuits develop? *Behavioural and Brain Sciences*, 31(1), 34–35.
- Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Viewpoint (in)dependence of action representations: an MVPA study. *Journal of Cognitive Neuroscience*, 24(4), 975–89. doi:10.1162/jocn_a_00195
- Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2013). Crossmodal and action-specific: neuroimaging the human mirror neuron system. *Trends in Cognitive Sciences*, 17(7), 311–8. doi:10.1016/j.tics.2013.04.012
- Parkinson, J., Springer, A., & Prinz, W. (2011). Can you see me in the snow? Action simulation aids the detection of visually degraded human motion. *Quarterly Journal of Experimental Psychology* (2006), 64(8), 1463–72. doi:10.1080/17470218.2011.594895
- Parkinson, J., Springer, A., & Prinz, W. (2012). Before, during and after you disappear: aspects of timing and dynamic updating of the real-time action simulation of human motions. *Psychological Research*, 76(4), 421–33. doi:10.1007/s00426-012-0422-3
- Prinz, W. (2006). What re-enactment earns us. *Cortex*, 42, 515–517.

- Prinz, W., & Rapinett, G. (2008). Filling the Gap : Dynamic Representation of Occluded Action. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions* (pp. 223–236). Amsterdam: IOS Press.
- Reed, C. L., & Vinson, N. G. (1996). Conceptual effects on representational momentum. *Journal of Experimental Psychology. Human Perception and Performance*, 22(4), 839–50.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–92. doi:10.1146/annurev.neuro.27.070203.144230
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, 12(2), 149–54
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews. Neuroscience*, 11(4), 264–74. doi:10.1038/nrn2805
- Saunders, J. a, & Knill, D. C. (2003). Humans use continuous visual feedback from the hand to control fast reaching movements. *Experimental Brain Research*, 152(3), 341–52. doi:10.1007/s00221-003-1525-2
- Saxe, R., Jamal, N., & Powell, L. (2006). My body or yours? The effect of visual perspective on cortical body representations. *Cerebral Cortex (New York, N.Y. : 1991)*, 16(2), 178–82. doi:10.1093/cercor/bhi095
- Sparenberg, P., Springer, A., & Prinz, W. (2012). Predicting others' actions: evidence for a constant time delay in action simulation. *Psychological Research*, 76(1), 41–9. doi:10.1007/s00426-011-0321-z
- Umiltà, M. a, Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing. a neurophysiological study. *Neuron*, 31(1), 155–65.
- Urgesi, C., Savonitto, M. M., Fabbro, F., & Aglioti, S. M. (2012). Long- and short-term plastic modeling of action prediction abilities in volleyball. *Psychological Research*, 76(4), 542–60. doi:10.1007/s00426-011-0383-y
- Verfaillie, K. (1993). Orientation-dependent priming effects in the perception of biological motion. *Journal of Experimental Psychology*, 19(5), 992–1013.
- Verfaillie, K., & Daems, A. (2002). Representing and anticipating human actions in vision. *Visual Cognition*, 9(1/2), 217–232.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3), 460–73. doi:10.1037/0033-2909.131.3.460
- Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology : CB*, 11(18), R729–32.

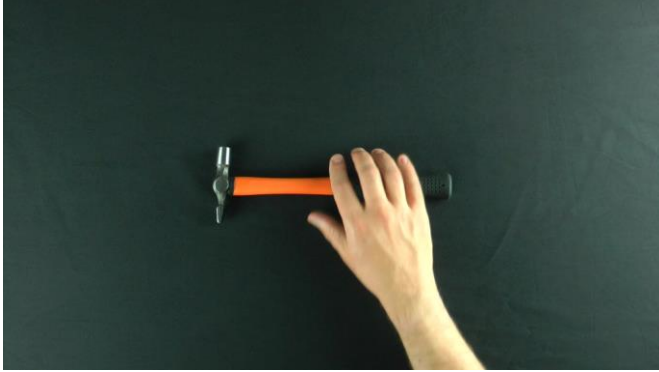
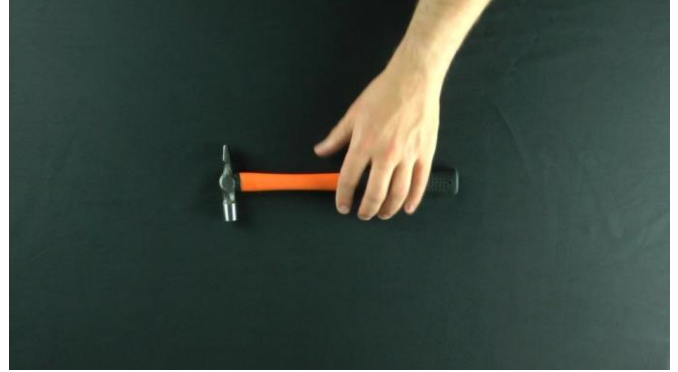
FIGURES**1PP****3PP**

Figure 1: Static images taken from example video stimuli in the first person perspective (1PP) and the third person perspective (3PP) conditions. 3PP action sequences were created from the 1PP video sequences by flipping the images about the horizontal axis.

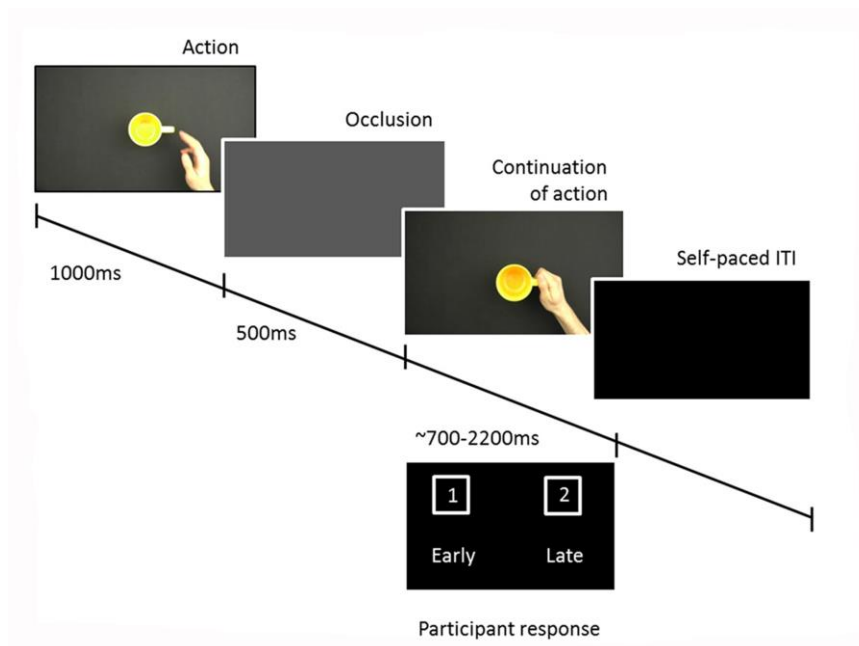


Figure 2. Schema of a trial. Each trial began with 1000ms of the start of an action sequence, followed by an occlusion screen for a fixed duration of 500ms. The continuation of the action then followed until its completion, with the re-start being incoherent-early, incoherent-late, or coherent with the length of the occluder. Participants had to respond during the continuation of the action as to whether they thought the action re-started from a point that was too early or too late.

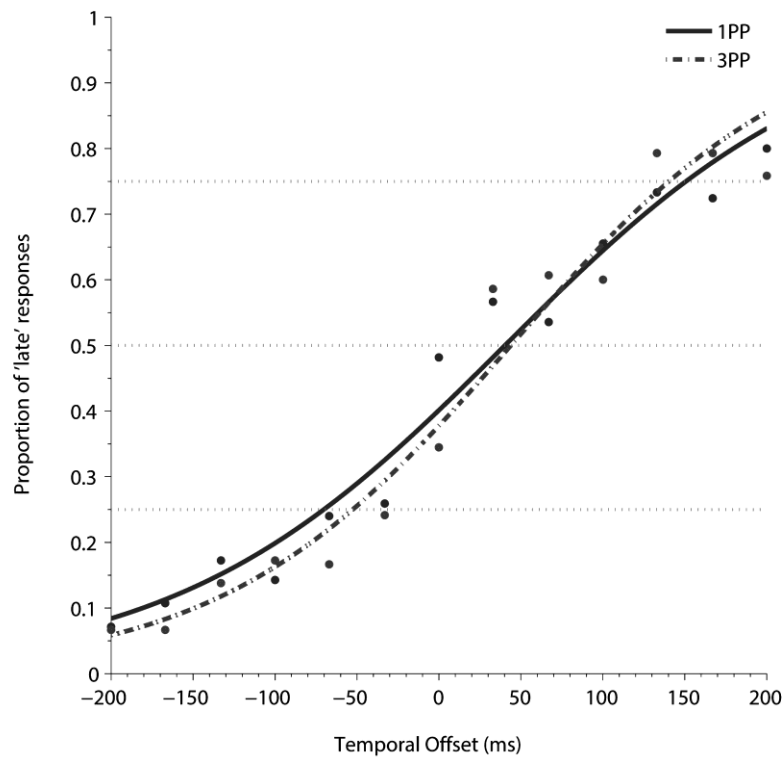


Figure 3. Example logistic function fitted to an individual participant's proportion of 'late' responses (pLate) per temporal offset condition. The solid line depicts the function for first person perspective actions (1PP) and the dashed line for third person perspective actions (3PP).

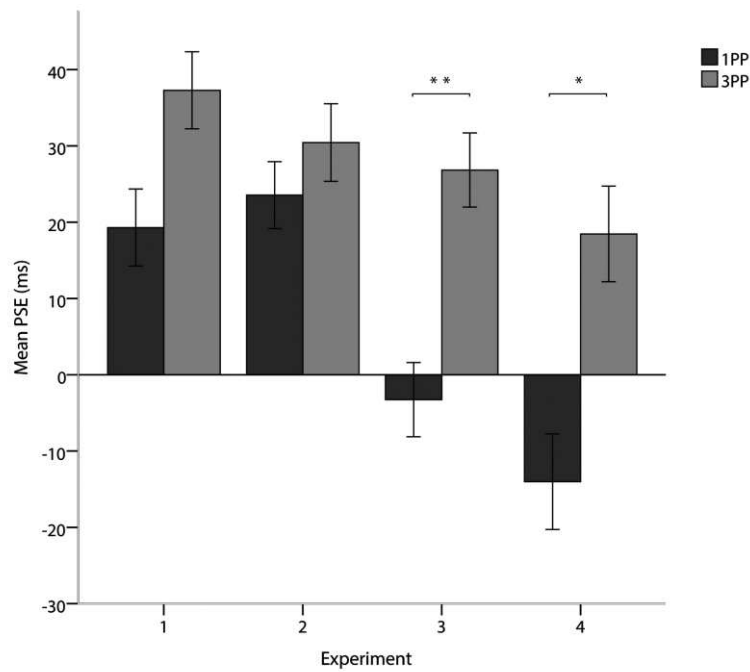


Figure 4. Bar graph depicting the mean point of subjective equality (PSE) for each perspective in each experiment. Blue bars show mean PSEs for first person perspective (1PP) actions, green bars show mean PSEs for third person perspective (3PP) actions. Significance levels are represented as ** for $p < .005$, * for $p < .05$. Error bars depict ± 1 standard error.

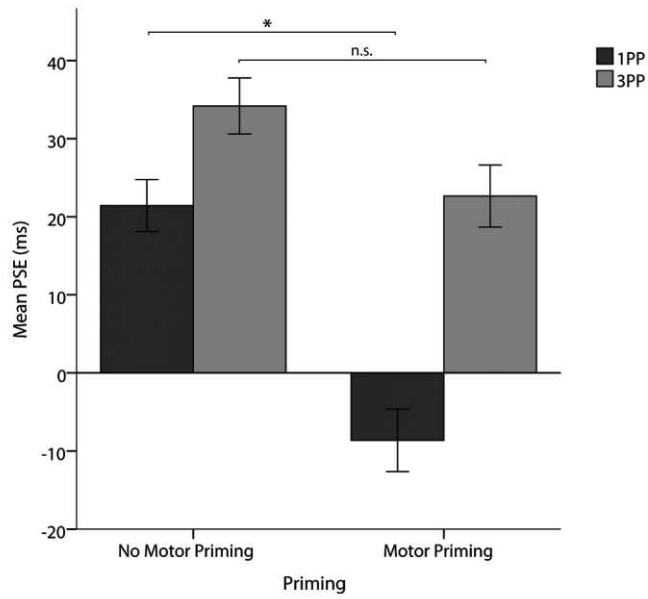


Figure 5. Bar graph depicting the mean point of subjective equality (PSE) for each perspective with and without motor priming. Blue bars show mean PSEs for first person perspective (1PP) actions, green bars show mean PSEs for third person perspective (3PP) actions. Significance levels are represented as * for $p < .05$. Error bars depict ± 1 standard error.