Original Articles

# Predicting actions from subtle preparatory movements

Maryam Vaziri-Pashkam *, Sarah Cormiea, Ken Nakayama

*Vision Sciences Laboratory, Department of Psychology, Harvard University, Cambridge, MA 02138, United States*

ABSTRACT

To study how people anticipate others' actions, we designed a competitive reaching task. Subjects faced each other separated by a Plexiglas screen and their finger movements in 3D space were recorded with sensors. The first subject (Attacker) was instructed to touch one of two horizontally arranged targets on the screen. The other subject (Blocker) touched the same target as quickly as possible. Average finger reaction times (fRTs) were fast, much faster than reactions to a dot moving on the screen in the same manner as the Attacker's finger. This suggests the presence of subtle preparatory cues in other parts of the Attacker's body. We also recorded videos of Attackers' movements and had Blockers play against unedited videos as well as videos that had all preparatory cues removed by editing out frames before Attacker finger movements started. Blockers' fRTs in response to the edited videos were significantly slower (∼90 ms). Also, reversing the preparatory movements in the videos tricked the Blockers into choosing the incorrect target at the beginning of their movement. Next, we occluded various body parts of the Attacker and showed that fRTs slow down only when most of the body of the Attacker is occluded. These results indicate that informative cues are widely distributed over the body and Blockers can use any piece from a set of redundant cues for action prediction. Reaction times in each condition remained constant over the duration of the testing sessions indicating a lack of learning during the experiment. These results suggest that during a dynamic two-person interaction, human subjects possess a remarkable and built-in action reading capacity allowing them to predict others' goals and respond efficiently in this competitive setting.

Published by Elsevier B.V.

## 1. Introduction

To navigate the social environment, we often need to predict the goals of other agents based on their movements. Coordinated group dances, competitive sports, or even a simple handshake require accurate predictions of others' movements. What makes these predictions possible? Human body movements follow distinct patterns due to biomechanical constraints (Johansson, 1973). Moving a hand towards a target on a table, for example, may require lifting the elbow and abducting the arm. Other more distributed adjustments may also be necessary to position the center of gravity of the body appropriately. Does the human visual system have access to information about these biomechanical constraints? Can they be used to predict the goals of others in a simple interaction? Does this predictive ability require training?

Humans are able to extract a diversity of information from viewing the actions of others. They can determine types of actions

(Johansson, 1973) as well as gender (Kozlowski & Cutting, 1977; Troje, Sadr, Geyer, & Nakayama, 2006), identity (Loula, Prasad, Harber, & Shiffrar, 2005; Troje, Westhoff, & Lavrov, 2005), emotion (Atkinson, Dittrich, Gemmell, & Young, 2004; Chouchourelou, Matsuka, Harber, & Shiffrar, 2006; Dittrich, Troscianko, Lea, & Morgan, 1996) and size of the actors (Jokisch & Troje, 2003) and the properties of the manipulated objects (Runeson & Frykholm, 1981) from spatiotemporal patterns of the movements of the body parts (Cutting, Moore, & Morrison, 1988). Besides reading ongoing bodily cues, we can predict future events from observing actions. Infants as young as 11 months can anticipate the goal of an adult's action based on prior familiarization with that goal (Cannon & Woodward, 2012). This goal prediction ability reaches its full potential in adults (Frith & Frith, 2006; Sebanz & Knoblich, 2009) and will also include predictions of social intents (Ansuini, Cavallo, Bertone, & Becchio, 2015; Lewkowicz, Quesque, Coello, & Delevoye-Turrell, 2015; Manera, Becchio, Cavallo, Sartori, & Castiello, 2011; Quesque & Coello, 2015; Quesque, Delevoye-Turrell, & Coello, 2016; Sartori, Becchio, & Castiello, 2011), as well as the consequences of action in the physical world. For example, Diaz, Fajen, and Phillips (2012) showed that subjects could predict

the direction of a ball after viewing the movements of a kicker. Similar results have been found in tennis (Farrow & Abernethy, 2003), cricket (Müller, Abernethy, & Farrow, 2006), badminton (Abernethy & Zawi, 2007), squash (Abernethy, Gill, Parks, & Packer, 2001), baseball (Ranganathan & Carlton, 2007), volleyball (Starkes, Edwards, Dissanayake, & Dunn, 1995), basketball (Aglioti, Cesari, Romani, & Urgesi, 2008), and darts (Knoblich & Flach, 2001). Other than explicit reports, studies of human eye movement behavior during action observation demonstrate that gaze positions follow the predicted action goals (Ambrosini, Pezzulo, & Costantini, 2015; Flanagan & Johansson, 2003; Flanagan, Rotman, Reichelt, & Johansson, 2013; Rotman, Troje, Johansson, & Flanagan, 2006).

Another remarkable feature of human action reading ability is that even partial information is sufficient for predicting the future course of an action. For example Louis-Dam, Orliaguet, and Coello (1999) asked actors to reach for an object to move it to a target zone and showed that observers who had only viewed the beginning of the reach were able to predict whether the target zone was close or far. Similar predictive abilities have been found in reaching (Martel, Bidet-Ildei, & Coello, 2011; Pesquita, Chapman, & Enns, 2016), weightlifting (Runeson & Frykholm, 1983), speech (Abry, Cathiard, Robertribes, & Schwartz, 1994), writing (Kandel, Boë, & Orliaguet, 1993; Orliaguet, Kandel, & Boë, 1997) and sign language (Pennel, Coello, & Orliaguet, 1999). Also, Graf et al. (2007) used point-light displays to show that even when part of the timecourse of an action is obscured, humans can judge whether the part after occlusion belonged to the same action sequence. These results along with other similar evidence (Parkinson, Springer, & Prinz, 2011, 2012) demonstrate that humans are sensitive to the spatiotemporal specifics of a movement; viewing the beginning of a movement triggers predictions about the future course of that movement. Most of these studies have relied on off-line reports by the subjects in response to videos of moving actors. In a typical experiment, a video or a movement sequence is played and cut at various time points, and subjects are asked to decide the outcome based on partial information (Runeson & Frykholm, 1983). These psychophysical studies are limited because they allow the subjects time to reflect on what they have seen. Overcoming this limitation requires studying action prediction in real time, in a naturalistic interactive setting.

In everyday life reading actions occurs extemporaneously. Humans continuously predict the goals of others to inform their own actions. Only by measuring movements in real time we can hope to characterize such naturally occurring processes. In fact, studies of competitive sports have shown that real-time responses in naturalistic settings might provide different results from off-line responses (Farrow & Abernethy, 2003; Ranganathan & Carlton, 2007).

Here, we would like to examine subjects' ability to predict the immediate goals of others' actions in the context of a realistic interaction. Using a motion-tracking device, we measured the movements of one subject in response to another. This design allows for moment-to-moment analysis of the subjects' movements to determine if they anticipate their opponent's goals. The task was a competitive reaching task in which one subject (Attacker) had to choose a target and tap it with their finger, and another (Blocker) had to block the same target by tapping it soon after the Attacker. We found that subjects were surprisingly fast in responding to their opponent, much faster than when they responded to a dot projected on the screen that moved to the targets. Reaction times were fast from the beginning of the experiment with no need for training. In subsequent experiments, we demonstrated that the Blocker could use predictive cues present well ahead of the finger movement of the Attacker to reduce their reaction time. We showed that removing the predictive cues slo-

wed down the Blockers, and inaccurate cues tricked the Blockers into reaching for the wrong target. In the next experiment, we explored the location of the predictive cues and showed that they are distributed over various body parts of the Attacker. Together these results demonstrate that humans can efficiently read out cues from multiple body parts of their opponent for movement anticipation and can readily use these cues to guide their own actions.

### 1.1. General methods

#### 1.1.1. Apparatus
Stimulus generation and data analysis were done on a Windows computer with MATLAB Psychtoolbox software. Hand movements were tracked with Polhemus Liberty, an electromagnetic position and orientation measuring system with an update rate of 240 Hz. A small position-tracking sensor ($1.27 \times 2.22 \times 1.9$ cm) was attached to the tip of the right index finger to record the 3D position of the fingertip.

#### 1.1.2. Subjects
All subjects were aged 18–35, were right-handed and had normal or corrected-to-normal vision. Subjects gave their informed consent prior to the experiments and received compensation for their participation. All experiments were approved by the Committee on the Use of Human Subjects at Harvard University. The number of subjects is detailed under the Methods section of each experiment.

## 2. Experiment 1: fast finger reactions

### 2.1. Methods

#### 2.1.1. Subjects
11 pairs of subjects participated in this experiment.

#### 2.1.2. Stimuli and procedure
Two subjects sat across from each other (~1.2 m apart) separated by a large (1.2 m × 1.5 m) Plexiglas screen (each subject was ~63 cm from the screen). Two small pieces of foam (5 cm × 5 cm) were affixed to the screen to serve as targets. At the beginning of each session, the position of the two targets and the starting points were calibrated separately for each sensor to account for minor variations across experimental sessions. Subjects were randomly assigned one of two roles: Attacker or Blocker. A beep sound, audible to both subjects, prompted the start of each trial, at which point the Attacker chose and reached for one of the two targets and the Blocker responded by reaching for the same target as fast as possible, attempting to beat the Attacker. The Blocker was announced to be the winner if they hit the same target as the Attacker within a time window after the Attacker. The size of the time window was adjusted for each pair so that the Blocker won in approximately half the trials. To do that, in each trial (except for the first five trials), the time window was set to be equal to the median hit time difference between the Attacker and Blocker in the prior trials. If the Blocker hit the target after this window or hit the wrong target, the Attacker was announced to be the winner of the trial. The Attacker was instructed to go directly to the target without any attempt to trick the Blocker. The Attacker sat behind an opaque panel that covered their body from the waist down. Inter-trial interval was randomly set to be between 1 and 4 s. Both the Attacker and Blocker started their movements from a flat resting spot placed ~28 cm from the screen (Fig. 1, also see Video 1 for a full video of two subjects engaged in the game). Each pair of subjects completed two blocks of 30 trials.
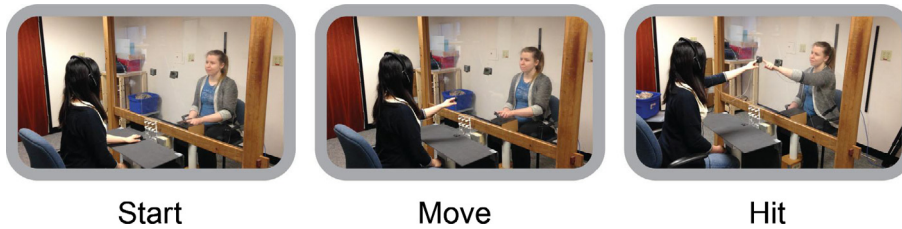
**Fig. 1.** Set up of Experiment 1. Subjects completed a competitive reaching task while seated facing each other separated by a Plexiglas screen. Two targets were affixed to the screen equidistant from both Subjects. The Attacker is pictured on the left, the Blocker on the right.

*2.1.3. Analysis*

In this experiment and all the subsequent experiments, the finger reaction time (fRT) for each trial was calculated as:

$$fRT = T_b - T_a$$

where $T_a$ is the first time point when the absolute speed of the Attacker in the 3D space exceeded a speed threshold (25 cm/s for 15 samples in a row) and $T_b$ is the first time point when the horizontal speed of the Blocker's finger towards their final target exceeded the same speed threshold. For the Blocker's start point, we used the horizontal speed towards the target instead of the absolute 3D speed in order to account for changes of direction. Fig. 2 shows two sample trials one without (Fig. 2a) and one with (Fig. 2b) a change of direction. Note that calculating the finger reaction time based on the horizontal speed towards the target allowed us to accurately measure the true reaction time of the Blocker for their final choice in both types of trials.

To calculate movement times for the Blockers, we subtracted the time in which they hit the target from their finger start time ($T_b$). The peak velocity was calculated as the highest Euclidean velocity during the movement time. The standard deviation of the final hit points was calculated for and then averaged across x and y directions and the two targets to obtain average endpoint variability.

*2.2. Results*

Fig. 2c shows 30 sample trajectories of one pair of subjects. Both Attackers and Blockers followed relatively consistent movement paths. The Blockers had a peak movement velocity of 656.7 ± 124.9 cm/s, movement time of 219.9 ± 57.1 ms, and endpoint variability of 1.49 ± 0.5 cm. The Blockers' final accuracy was high (92.7% ± 6.29). The analysis of finger reaction times (fRTs) showed very fast fRTs for the Blocker in response to the Attacker (155.7 ms, See Fig. 2d), especially when compared to similar hand movement reaction times previously reported in the cognitive psychology literature (Song & Nakayama, 2006). To our knowledge, this is the first time that such fast finger reactions have been reported in naïve subjects in a competitive interaction. What is the source of these fast reactions? We will address this question in the next two experiments.

## 3. Experiment 2: Human vs. Dot

The previous experiment demonstrated fast finger reaction times in response to the Attacker's finger movements. Did Blockers only focus on the finger movements of Attackers to achieve fast reactions, or did they gather cues from viewing other parts of the Attackers' body? If the finger movement is the only source of information for the Blockers, they should be equally fast reacting to any object that moves similarly to the Attackers' finger. To investigate this possibility, in Experiment 2, we measured the fRTs of a new set of Blockers in response to a dot moving on a screen in an essentially identical manner to an Attacker's finger and compared it to fRTs in response to human Attackers.
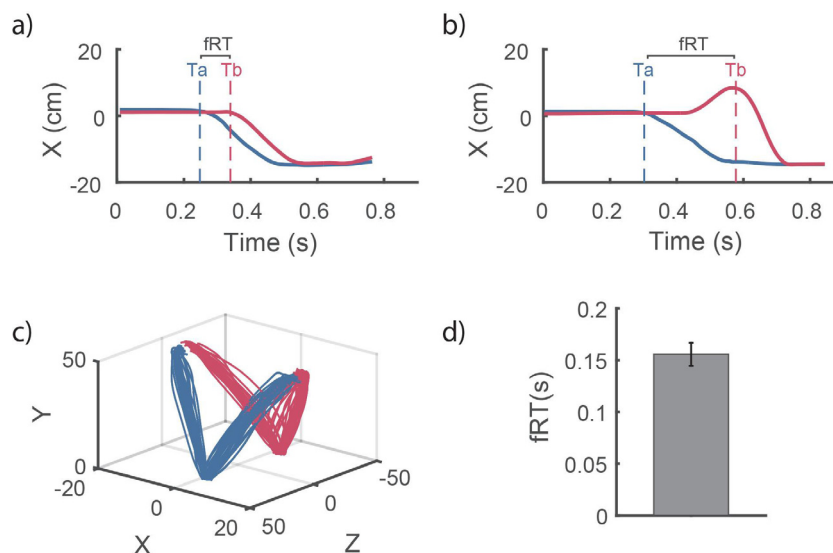


**Fig. 2.** (a) A typical trial, horizontal position (X) of a subject pair plotted against time. The blue line depicts the Attacker's motion and the red line depicts the Blocker's motion. The dashed lines represent the start of the finger movements of the Attacker ($T_a$) and the Blocker ($T_b$). (b) A less common trial, the Blocker changes direction mid-movement. (c) 3D plot of all trials from a representative pair of subjects. (d) Average fRT in Experiment 1. The error bar shows ±1 SEM.

### 3.1. Methods

#### 3.1.1. Subjects

31 subjects participated in this experiment, 10 pairs in the Human condition, and 11 single subjects in the Dot condition. One subject from the Dot condition was removed from the analysis due to very low accuracy (67%). Including this subject did not qualitatively change any of the results.

#### 3.1.2. Stimuli and procedure

This experiment consisted of two conditions: The Human condition and the Dot condition. The procedure of the Human condition was similar to Experiment 1 except that Attackers wore headphones and were told which target to choose at the start of each trial, thus eliminating any Attacker direction bias. Auditory instructions to indicate one of the two targets were the words "left" or "right" generated by a text-to-speech program. Also, the beep that was audible to the Blocker in the first experiment was eliminated in this experiment to remove other sources of predictability besides the Attacker's movement. The inter-trial interval in this experiment was set to 1 s. In the Dot condition stimuli were back-projected on a semi-transparent sheet affixed to the Plexiglas screen using a ViewSonic projector (1024 × 768, 60 Hz) at 70 × 50 cm (Fig. 3a). Twenty random motion paths were selected from Attacker data collected in Experiment 1 and used to create the dot stimulus for this experiment. The horizontal and vertical positions of the Attacker's finger were represented by the location of the dot on the screen. The trial started with a dot presented at the bottom of the screen equidistant from the two targets. Because the angular size of the finger increases as it approaches the targets, the dot's diameter started at 0.67 cm and increased to 1.34 cm at the endpoint (Fig. 3b). The Blocker was instructed to try to beat the moving dot and tap on the same target as fast as possible. The Blocker success in both the Human and Dot conditions was determined in the same manner as Experiment 1. In the Dot condition, the hit point was determined from the original 3D movements. Subjects in the Human condition completed two blocks of 30 trials, and those in the Dot condition completed three blocks of 40 trials. There was no noteworthy reason for this difference in the number of trials, and including only the first 30 trials in the Dot condition did not qualitatively change the results.

#### 3.1.3. Analysis

The analysis in this experiment was similar to Experiment 1. In addition to calculating all the measures in Experiment 1, in each condition, we also calculated the fRTs over the course of the experiment. To do this, for each subject we measured the fRT in each trial for each experimental condition. We then ran a linear regression analysis with the trial number as the independent variable and fRT as the dependent variable. We then performed a one-sample $t$-test on the regression slopes to determine whether they are greater than zero. For this analysis to have equal trial numbers across the two conditions, we only analyzed the first 60 trials of the Dot condition.

### 3.2. Results

Subjects were highly accurate in both experimental conditions (Dot: 99.7%, Human: 97.6%) with a slightly higher performance in the Dot condition ($t(18) = 3.03$, $p < 0.01$). The two conditions were similar to each other in many kinematic measures including movement time (M = 221.6, 224.6, SD = 37.8, 49.3 ms for the Human and Dot conditions respectively, $t(18) = 0.15$, $p = 0.88$), end point variability (M = 1.19, 1.33, SD = 0.39, 0.23 cm for the Human and Dot conditions respectively, $t(18) = 0.97$, $p = 0.34$), and peak velocity (M = 639.2, 627.4, SD = 96.4, 130.4 cm/s for the Human and Dot conditions respectively, $t(18) = 0.42$, $p = 0.82$). Also, these measures in the Human condition of this experiment were not significantly different from the results of Experiment 1 ($t(19) = 0.08$, 0.36, 1.51, ps = 0.93, 0.73, 0.15 for movement time, peak velocity and endpoint variability). Given these null results, and for brevity, we have refrained from reports of these measures for subsequent experiments.

The fRTs in the Human condition were slower (∼28 ms) than those in the previous experiment ($t(19) = 0.88$, $p = 0.15$). This could have been caused by the presence of the beep, and possible predictability of Attackers' target choice in the first experiment, or the fact that Attackers were directed to the target in this experiment instead of choosing the target themselves. In line with the latter possibility, a recent study (Pesquita et al., 2016) has demonstrated that reaction times in response to actors who voluntarily choose a target are faster than responses to those who are directed to the target. Despite the slower reactions compared to Experiment 1, as Fig. 4a shows, fRTs were still ∼116 ms slower in the Dot condition compared to the Human condition (Dot: 301.2 ms compared to Human: 184.4 ms; $t(18) = 8.15$, $p < 0.0001$). Since we have a small sample size for which it is hard to ascertain an underlying normal distribution, we also performed the analyses in this experiment as well as all the other analyses reported in this manuscript using bootstrapping technique free from any assumptions of normality. Results stayed qualitatively similar in all cases.

These results replicate those of the first experiment demonstrating fast fRTs in response to human Attackers. Furthermore, they show that the Blockers could not have been focused solely
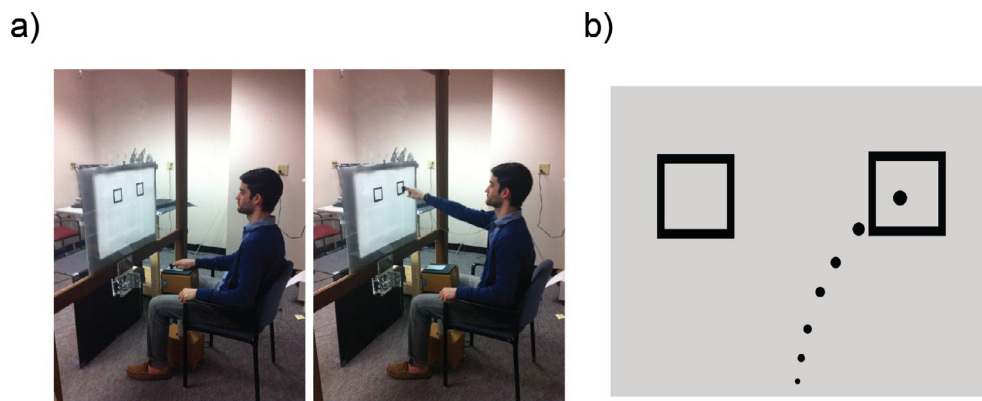
## a)

## b)



**Fig. 3.** Setup of the Dot condition in Experiments 2 and 5. (a) Stimuli were back-projected on the Plexiglas screen. (b) Blockers responded to a dot moving from the bottom of the screen to one of the two targets.
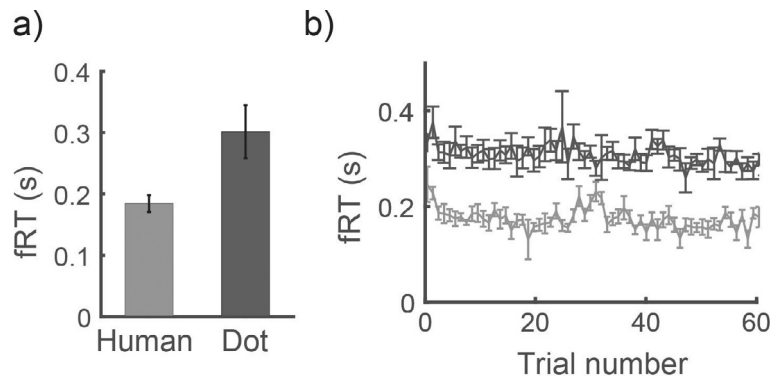
**Fig. 4.** Results of Experiment 2. (a) Average fRTs in the Human and Dot conditions. Error bars show ±1 SEM. (b) Average fRTs over the course of the experiment. The y-axis represents fRTs and the x-axis represents trial number. Subjects showed no sign of learning over the course of the experiment. Error bars show ±1 SEM.

on the finger position of the Attacker as the Dot condition in this experiment preserveed the finger movement information and the Blockers were slower in response to the dot.

Are subjects' fRTs fast from the beginning of the experiment or do they learn through the course of the experiment? To answer this question, we explored the change in reaction times across trials (Fig. 4b). A linear regression analysis (see Methods) showed no significant change of fRT across trials for the Human ($t(9) = 1.17$, $p = 0.27$) condition. This analysis showed similar results on the data from Experiment 1 ($t(10) = 1.31$, p = 0.22). The Dot condition however showed a marginally significant decrease in fRT across trials ($t(9) = 2.11$, p = 0.064). Subjects were fast in the Human condition from the beginning with no evidence of learning. What is the cause of these uncharacteristically fast fRTs? One possibility is that Blockers are not responding to the movement of the finger but to preparatory movements that happen prior to the movement of the finger. A reach movement is not merely an isolated thrust of the arm. Due to the connection of the arm to the rest of the body and the biomechanical constraints of the human musculoskeletal system, it is impossible to move the arm without engaging other body parts. These preparatory movements might cue the Blockers to the ultimate goal of the Attackers. In other words, it is possible that the putative fast reactions are in fact regular reactions to events that happen earlier than the finger start time.

## 4. Experiment 3: removing the predictive cues from videos

If the fast reactions are driven by body movements present before the Attackers start to move their finger, removing this information should slow down the responses. To test this, we videotaped an Attacker, and modified the videos to remove the predictive cues and measured the finger reaction times of the Blockers in response to the modified videos. This condition eliminates all possible predictive cues before the Attacker's finger starts to move. For comparison, and to ensure that playing against the videos was not qualitatively different from playing against a human Attacker, we also measured finger reaction times in response to unmodified videos as well as a real human Attacker.

### 4.1. Methods

#### 4.1.1. Subjects
10 subjects participated in this experiment.

#### 4.1.2. Stimuli and procedure
A lab member was videotaped as an Attacker playing against six Blockers in a set up similar to the Human condition of Experiment 2. We positioned the camera lens (GoPro Hero3+ Black Edition,

recording with $1280 \times 720$ spatial resolution at 60 Hz) slightly above and in front of the Blocker's eyes ~at 60 cm away from the Plexiglass screen to capture the visual scene from the point of view of the Blockers. To simulate the experience of playing against a real Attacker, we projected these videos with size $35.5 \times 70$ cm on the screen (same screen as in Experiment 2) so that the size of the Attacker in the video in visual angles roughly matched that of the real Attacker. A total of 747 videos of single trials (trial clips) were extracted from the videos and for each trial clip, the frames that matched the start and end of the finger movement towards the target were identified. The experiment consisted of three conditions: Real, FullVid, CutVid. In the Real condition, subjects played against the same lab member shown in videos. In the FullVid condition, subjects played against unedited trial clips. In the CutVid condition, we removed all frames before Attacker start time ($T_a$) and replaced them with the first frame of the trial clip in which no movement had yet occurred (Fig. 5a). The first frame of each trial clip was the frame immediately after the instruction was provided to the Attacker. On average, ~30.29 frames (~504.89 ms) were removed from the videos in this condition. A random set of trial clips was selected for each condition. No video was repeated for a given subject. Each subject completed two blocks of 30 trials for each experimental condition. Inter-trial interval was randomly set to be between 1 and 4 s in all conditions. The order of conditions was counterbalanced across subjects.

### 4.2. Results

Subjects were highly accurate in all three conditions (Real: 99.5%, FullVid: 98.5%, CutVid: 100%) with no significant difference between the conditions (ps > 0.05). A repeated-measures ANOVA showed a significant effect of condition on fRT (F(2,9) = 73.63, p < 0.001). The reaction times in the Real and FullVid conditions were both fast and were not significantly different from one another ($t(9) = 1.072$, p = 0.31). As such, these results replicate the results of the two previous experiments. In addition, they show that all the cues responsible for fast reaction times are preserved in the video clips. The reaction times in the CutVid condition, however, were significantly slower than both the Real (fRT difference 116 ms, $t(9) = 15.02$, p < 0.0001) and the FullVid conditions (fRT difference 105 ms, $t(9) = 8.12$, p < 0.0001). Removing the video frames with preparatory cues slowed down the Blockers. Note that the difference between the full and cut videos is subtle. In fact, when the two videos are compared with one another it is hard to tell the two apart (see Video 2 that shows an example CutVid and FullVid trial clip side by side. The left panel shows a CutVid clip and the right panel shows a FullVid clip). We debriefed the subjects after the experiment and 7 out of 10 subjects did not notice that
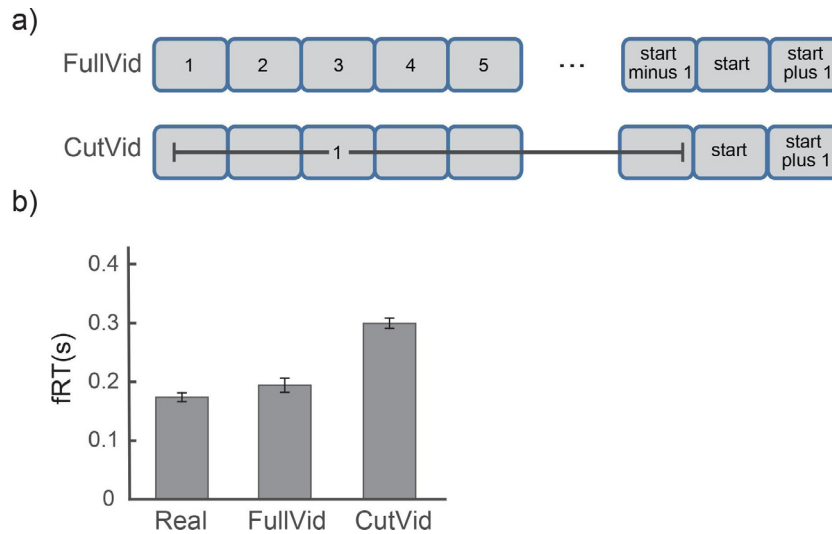
**Fig. 5.** Results of Experiment 3. (a) Schematics of the FullVid and CutVid conditions in Experiment 3. In the CutVid condition, all the frames before the start of the Attacker were removed and replaced with the first frame of the trial clip. (b) Results of Experiment 3, Average reaction times for the three experimental conditions. Error bars show ±1 SEM.

the videos were cut in the CutVid blocks. Despite the subtlety of this difference, it still provided the subjects with sufficient information for a speeded accurate response. These results suggest that Blockers indeed use predictive information before the finger movement to determine the goal of the Attacker and to reduce their reaction times.

## 5. Experiment 4: incongruent predictive cues in videos

If the preparatory information is predictive of Attackers' goals, then incongruent preparatory cues should cause Blockers to react slower or choose the wrong target. To test this possibility, we measured the fRT of subjects in response to modified videos with misleading preparatory cues and compared that to videos with correct preparatory cues.

### 5.1. Methods

#### 5.1.1. Subjects

20 subjects participated in this experiment, 10 in the Congruent condition and 10 in the Incongruent condition. One subject was removed from the incongruent condition due to very low accuracy (54%). Including this subject did not qualitatively change any of the results.

#### 5.1.2. Stimuli and procedure

We used the videos from Experiment 3. We divided each video into two epochs. Epoch1 included frames from when the Attacker first received instructions until their finger began to move (on average ~30 frames equal to ~500 ms), and epoch2 consisted of the frame in which the Attacker started to move (Ta) and all frames after that until a target was hit (on average ~13 frames equal to ~222 ms). The experiment consisted of Congruent and Incongruent conditions. In the Congruent condition, epoch1 and epoch2 from two trials with the same direction were combined. In the Incongruent condition, epoch1 and epoch2 from two trials with different directions were combined. Note that for both congruent and incongruent conditions, we drew frames for epoch1 and epoch2 from two different videos to keep the two conditions as comparable as possible so as to not unfairly advantage the congruent case because of continuity. No video was repeated for a given subject. Each subject completed one block of 80 trials. Inter-trial interval was randomly set to be between 1 and 4 s in both conditions. Half of the subjects completed the Congruent condition and the other half the Incongruent condition.

#### 5.1.3. Analysis

As before, we measured the average fRT and accuracy in this experiment. In addition, we performed an analysis to compare moment-by-moment accuracy across conditions. On each trial and at each time point, we determined if the subjects were moving towards the correct target or not. To do this, we used the speed in the horizontal direction, and if the speed vector was towards the correct direction, the accuracy at that time point was set to 1, otherwise it was set to 0. We then averaged this instantaneous accuracy across trials and subjects to determine the time-course of the increase in accuracy during the trial in each condition. From this, we can determine if subjects are moving towards the correct target at each time point and if this behavior varies between conditions.

### 5.2. Results

We first measured the average accuracy and reaction time of the subjects in the Congruent and Incongruent conditions. At first glance, the results showed no difference between conditions for either accuracy (Ms = 0.98, 0.92 and SDs = 0.04, 0.92 for the congruent and incongruent conditions respectively, $t(18) = 0.82$, $p = 0.41$) or reaction time (Ms = 189.19, 210.27 and SDs = 55.18, 58.97 ms for the congruent and incongruent conditions respectively, $t(18) = 1.6$, $p = 0.12$). However, a closer look at the data from individual subjects revealed that in the Incongruent condition subjects had different patterns of results. Some had slow reaction times and were highly accurate, while others responded early and were inaccurate (Fig. 6a). The average accuracy and reaction time thus does not fully capture the intricacies of movement in the Incongruent condition. To get around this problem and to have a closer look at the Blockers' behavior during the trial, we ran a separate analysis. We used the instantaneous horizontal direction of the finger to define the average accuracy across all Blockers at each time-point during the trial (see Methods). Accuracy rose smoothly in the Congruent condition. In the Incongruent condition, however, the initial accuracy was lower than chance level indicat-
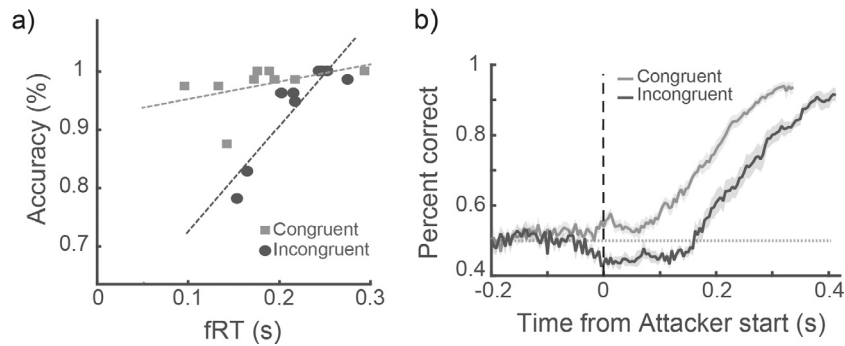
**Fig. 6.** Results of Experiment 4. (a) Individual subjects' results in Experiment 4. Each dot represents one subject with their average accuracy plotted on the y-axis and fRT plotted on the x-axis. In the Congruent condition, subjects were accurate with a small positive relationship between the accuracy and reaction time. In the Incongruent condition, however, this relationship was much more pronounced. The dashed lines represent least-square linear regression fit. (b) The average accuracy for each time point determined as the proportion of trials in which the Blocker was moving towards the correct target. The light gray and the dark gray lines depict the results for the Congruent and the Incongruent conditions respectively. The dashed line represents the Attacker start time and the dotted line represents chance level (0.5). The shaded regions represent ±1 SEM.

ing that on average subjects chose the incorrect target initially and then corrected their choice later in the trial (Fig. 6b). These results demonstrate that it is possible to trick the Blockers into choosing the wrong target through systematic manipulation of the preparatory information.

## 6. Experiment 5: where in the body?

Where are these predictive cues? Are they concentrated on a single body part? If so covering that part should slow down reaction times. To test this, we covered the screen with occluders, allowing only selected parts to be visible in each condition. We chose two non-overlapping regions, one in which only the head and shoulders were visible (Top) and another in which only the torso and arms were visible (Torso). To determine the contribution of the eyes, we added another Top condition in which subjects wore sunglasses (Top-SG). And to minimize all possible visible cues we had a head-only condition in which subjects also wore sunglasses (Head-SG). We then measured the fRT of Blockers in response to occluded Attackers (Fig. 7a) and compared that to responses to dots back-projected on the screen.

### 6.1. Subjects

50 subjects participated in this experiment, 10 individual subjects participated in the Dot-full and Dot-end conditions, 10 pairs participated in the All, Top and Torso conditions and 10 separate pairs participated in the Top-SG and Head-SG conditions.

### 6.2. Stimuli and procedure

This experiment consisted of 7 conditions: All, Top, Torso, Top-SG, Head-SG, Dot-full and Dot-end. In each condition parts of the Attacker's body was obscured from the view of the Blocker. The All condition and the Dot-full conditions were similar to the Human and Dot conditions in Experiment 2 respectively with the exception that the inter-trial interval varied randomly between 1 and 4 s in this experiment. This was to ensure that the predictability of the trial timing were not the source of the fast reaction times and that the results could be generalized to a broader range of conditions with more unpredictability. In the Top condition, only the Attacker's shoulders and head were visible. In the Torso condition, only the waist to shoulders was visible. In the Top-SG condition, sunglasses were added to the Top condition. The Head-SG condition was the same as the Top-SG condition, but the shoulders and neck were also obscured. To have a better control for the

Head-SG condition we ran an extra Dot-end condition in this experiment in which the dot appeared only in the final target position instead of moving from the start point to the target. Subjects completed two blocks of 30 trials for each condition with the Human Attacker and three blocks of 40 trials for each of the Dot-full and Dot-end conditions. In the cases that the subjects performed more than one condition, the order of the blocks was counterbalanced across subjects.

### 6.3. Analysis

For this experiment, in addition to calculating the average fRTs, we ran an analysis to determine the amount of learning by measuring the decrease in reaction time over the course of the experiment in each condition (see Methods of Experiment 2).

### 6.4. Results

Subjects were highly accurate in all conditions (Ms = 0.99, 0.97, 0.97, 0.97, 0.96, 0.95, 0.99, 0.998, and SDs = 0.008, 0.03, 0.02, 0.05, 0.05, 0.006, 0.003 for the All, Torso, Top, Top-SG, Head-SG, Dot-full, and Dot-end conditions, respectively). Given the near ceiling accuracy in all conditions, there was not much room for variability across conditions; we thus focused our analysis on the fRTs in this experiment. We replicated the results of Experiment 2. In the All condition, the fRTs were 105 ms faster than those in the Dot-full condition. These results once again show that Blockers use preparatory cues before the start of the Attacker's finger movement for speeded reactions in the All condition. Note that the reaction times in this experiment were overall slower than those in Experiment 2 as the inter-trial interval was unpredictable (Klemmer, 1956). Nevertheless, the difference between the two conditions was preserved in this Experiment ($t(18) = 4.62$, $p < 0.001$).

Focusing on the first three conditions (All, Torso, and Top), a repeated-measures ANOVA showed a significant effect of condition on fRT ($F(2,9) = 6.54$, $p < 0.01$). The reaction times in the Torso condition were not significantly different from the All condition ($t(9) = 1.21$, $p = 0.26$) and were ~94 ms faster than the Dot-full condition ($t(18) = 4.47$, $p < 0.001$). In other words, preparatory cues in just the torso and arms were enough to help the Blockers achieve fast reaction times. The reaction times in the Top condition were slightly slower (~35 ms) than the All condition ($t(9) = 3.72$, $p < 0.01$), but still significantly faster (70 ms) than the Dot-full condition ($t(18) > 3.06$, $p < 0.01$). These results indicate that there are also preparatory cues in the head and shoulders, but they are avail-
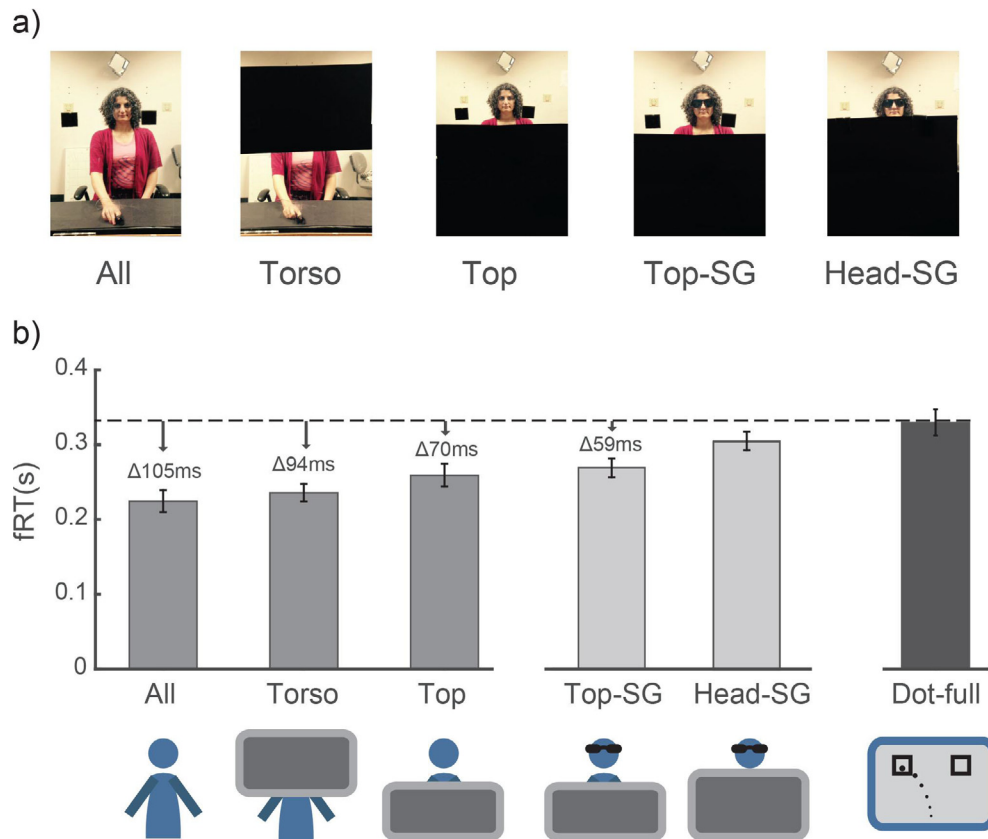
**Fig. 7.** Results of Experiment 5. (a) Setup of all the conditions in Experiment 5 that have a human Attacker. In each condition, part of the Attacker's body was occluded. (b) Average fRTs in the 5 conditions with the human Attacker compared to the fRTs in the Dot-full condition.

able slightly after those in the torso and arms. Note that Blockers view completely non-overlapping regions of the Attacker's body in the Torso and Top conditions. Moreover, in the Top condition, the hands and arms are not visible until the very end of the Attacker's movement. Thus, the fact that both of these conditions have faster reaction times than the Dot-full condition demonstrates that the informative cues are distributed over large swaths of the body that are indicative of the goal of the movement.

To further restrict the Blocker's view, in a new set of subjects we measured the reaction times in two additional conditions. The Top-SG condition was similar to the Top condition, but Attackers wore dark sunglasses to cover their eyes. In the Head-SG condition, the Attackers wore sunglasses, and only their head from the chin up was visible. The fRTs in the Top-SG condition were ∼59 ms faster than the Dot-full condition ($t(17) = 2.70$, $p < 0.05$) and ∼46 ms slower than the All condition ($t(17) = 2.36$, $p < 0.05$). In other words, covering the eyes did not slow fRTs as much as the Dot condition, and subjects could use cues from the shoulders and the head to achieve fast reactions. In natural reaching, it is often the case that the eyes move to the target before the manual reach (Prablanc, Echallier, Komilis, & Jeannerod, 1979). Therefore, it is possible that Blockers would be able to use Attackers' eye gaze for goal prediction. The two conditions Top-SG and Top allow us to directly test whether Blockers use Attackers' gaze for fast reactions. The reaction times of the two conditions were not significantly different ($t(17) = 0.56$, $p = 0.58$), suggesting that the visibility of the eyes was not enough to explain the fast reactions in our competitive reaching setting.

The Head-SG condition was the only condition in this experiment with slow fRTs comparable to the Dot-full condition ($t(17) = 1.06$, $p = 0.30$). This condition was ∼82 ms slower than the All

condition ($t(17) = 4.21$, $p < 0.001$). But even in this condition, there was some predictive information. Note that subjects saw the finger only towards the end of the movement in the Head-SG condition, while in the Dot condition movement of the dot started from the bottom of the screen. Therefore, the fact that the fRTs in the Head-SG condition were not slower than the Dot condition indicates that even the head of the Attacker contains some predictive information. In fact, to have a better control for the Head-SG condition, we ran another block in which we only presented the dot at the end of the movement at the final target location (see Methods). The fRTs in this condition were significantly slower than the Head-SG condition ($t(17) = 2.95$, $p < 0.01$) supporting the presence of predictive cues in the head of the Attacker. Overall the results from all the conditions with the human Attacker determine that to perform this task subjects can gather information from any available body part and that the preparatory cues are distributed over the body. Further analysis of the reaction times over the course of the experiment in each condition (See Methods) showed no signs of learning. A linear regression analysis showed no change in the reaction times throughout the block for all conditions with the human Attacker ($ts < 1.64$, $ps > 0.14$). Similar results were found in the Dot condition ($t(9) = 0.72$, $p = 0.49$). The Blockers were able to efficiently collect the predictive cues from any available body part of the Attacker to predict the movement goals without any need for training.

## 7. Discussion

In this study, we measured the hand movements of subject pairs during a dynamic interaction. One subject (Attacker) was asked to reach for one of two targets, and another (Blocker) was

asked to reach for the same target as quickly as possible. Reaction times of the Blocker were surprisingly short, approximately 100 ms faster than expected, suggesting that preparatory cues before the start of the Attacker's movement are predictive of their goals and Blockers are able to use those cues for such speeded responses. Experiments using video recordings of an Attacker confirmed this, showing that (1) with the putative preparatory cues removed, reaction times slowed correspondingly and (2) when these cues were reversed, Blockers moved towards the incorrect target if they relied on the early cues. Furthermore, we found that the preparatory cues were widely distributed over the body of the Attacker.

A critical difference between our study and most previous studies of action prediction (e.g. Abry et al., 1994; Graf et al., 2007; Kandel et al., 1993; Martel et al., 2011; Orliaguet et al., 1997; Parkinson et al., 2011, 2012; Pennel et al., 1999; Pesquita et al., 2016; Runeson & Frykholm, 1983) is that the other studies lacked the essentials of a real interaction. Subjects viewed point-light walkers or videos of actors on a computer screen. The videos were cut to show only part of the movements and subjects make non-speeded predictions about the goal of the movement. However, in real-time interactions, people do not explicitly report their predictions of others' actions. The process is automatic and implicit, part of the observer's sensory-motor loop (Neisser, 1976). Investigating the process of action reading outside a natural context is limited. For example, subjects might have time to focus on specific body parts when they are not forced to respond immediately. Also viewing a few frames of a video might feel unnatural, force subjects to use strategies that are not generally used in a naturalistic setting. In our experiment, using continuous recordings of hand movements, we were able to show that subjects are not only able to read others' actions, but they can also efficiently incorporate their predictions into their own actions for speeded responses.

In Experiment 2, we showed that participants were faster in response to a human Attacker compared to a dot moving on the screen in the same manner as the Attacker's finger. One possible source of fast reaction times in this experiment could be the social and competitive nature of the task (Georgiou, Becchio, Glover, & Castiello, 2007) or the presence of more biological features in the case of a human Attacker compared to the Dot. However, the results of experiment 3 using videos suggest that fast reaction times do not result from these factors and show the direct role of preparatory information in speeded responses. In experiment 3 we used video recordings of an Attacker to be able to systematically manipulate the predictive cues. The reaction times in the cut videos in which the predictive information was removed were slower than those in response to full videos. Note that the social and competitive nature of the task and the biological nature of the visual signal is essentially the same between the intact and the manipulated videos. Thus, even with the sense of urgency supplied by the video, subjects were slow in response to the videos with only a fraction of a second removed.

Previous experiments have demonstrated the relevance of the beginning of a reaching movement in predicting the future course of a reach (Martel et al., 2011; Pesquita et al., 2016). These studies have used videos and have cut the movement at various time points and asked the subjects to predict the future course of an action. Our Experiments 3 and 4 complement these findings and in realistic settings demonstrates the causal role of this early information in fast reaction times. When the beginning of the movement was removed, Blockers were not able to respond in a speeded manner to the Attackers.

In Experiment 4, we manipulated the videos to trick the subjects with incongruent preparatory movements and showed that this incongruent movement affects the initial response of the Blockers. Similar paradigms have been used in the study of soccer kicks (Tomeo, Cesari, Aglioti, & Urgesi, 2012). However, Tomeo

et al. did not show the effect of incongruent information on realistic interactions and used video recordings and explicit psychophysical reports from observers. They also only demonstrated the effect of incongruent movements in a sports activity that requires more expertise compared to an everyday action such as reaching. In addition, the length of incongruent information in Tomeo et al. (2012) was much longer (more than 1 s) compared to our case in which the incongruent information was much more subtle. Our results thus extend the findings of Tomeo et al. (2012) to an everyday action in a realistic setting and show the causal role of subtle preparatory movements in determining the future course of an action.

The results of the last experiment demonstrate that our ability to predict action goals is remarkably resilient to occlusion of body parts and suggest that redundant cues are present in large swaths of the body. Similar manipulations have been used in studies of social intention of actions (Sartori et al., 2011), consequences of actions on a manipulated object (e.g., a soccer ball, Diaz et al., 2012), and attention control for an action (Pesquita et al., 2016). Our experiment 5 uses a similar occlusion paradigm and extends these findings to predictions of the future course of a reaching goal from the subtle preparatory movements. In our experiment, reaction times slowed down only when most of the body of the Attacker was occluded. In the Top and Torso conditions, for example, the visible body parts were fully non-overlapping. Yet, the reaction times in both conditions were fast indicating that the informative cues were present in the head/shoulders as well as the torso/arms. Note that the informative cues in the Top condition were not dependent on reading the eyes, as wearing sunglasses in the Top-SG condition did not change the speed of response. Even the Head-SG condition was not fully stripped of predictive information. In this condition head movement alone without shoulders or eyes had residual predictive power. In sum, distributed cues were present in various body parts that inform about the goal of the movement. These results are in line with previous literature in sports showing that distributed bodily cues are informative about future actions (Diaz et al., 2012; Huys, Smeeton, Hodges, Beek, & Wiliams, 2008). Our results extend these findings and show that these distributed cues can be used in a simple reaching task for fast responses. We further show that even a subset of these distributed cues is enough for making highly accurate predictions.

These results suggest that humans have implicit knowledge of the biomechanical constraints that govern bodily movements. In the human body, due to specific joint properties and musculoskeletal connections, movement of one limb often requires moving other body parts. In a reaching movement, the amount of torque on the elbow is directly affected by the movement in the shoulder joint (Hollerbach & Flash, 1982). Moreover, moving a limb often requires postural adjustments in order to stabilize the body in the new form. The first muscles to contract in a reaching task are those in the torso and lower limb used to adjust the center of gravity (Belen'kii et al., 1967). The knowledge of these movement principles is essential for predicting the actions of others. For example in our task reaching for a target will engage the shoulders and might require leaning to one side to adjust the center of gravity. The shoulder and trunk movements, in this case, are integral parts of the reaching movement and in a sense cannot be isolated from the final finger movement. It is possible that subjects maintain two models for the left and right reaches in their mind and compare each model to incoming evidence at each moment to make predictions about the action goal.

A notable finding in our study was that subjects were fast from the beginning of the experiment. In other words, no learning was required to achieve fast reactions. This stands in contrast with seemingly similar studies where significant improvements in performance over the course of the experiment were observed (Diaz

et al., 2012). There are several differences between our study and Diaz et al. (2012). First, they rely on explicit judgments that might require more training for action prediction as opposed to real-time interactions. Second, in their study, they have focused on action prediction in a particular sport that may also require more training compared to a simple everyday task such as reaching in our case. Third, the predictions in their study concern the effect of a kicking action on an object (the soccer ball) that again might need more training than the prediction of immediate goal of an action in our case. Our study shows that adult subjects do not need training for predicting immediate action goals in everyday interactions.

The subtlety of the preparatory cues used to predict the Attacker's movements is particularly significant. This is evident in the videos of Experiment 3 (see Video 2) where the difference between the full videos (that include preparatory cues) and the cut videos (with no preparatory cues) is barely noticeable. These results are consistent with previous literature studying goal prediction in the context of action observation. It has been shown that after only 100–170 ms of hand movement, observers' gaze predict the eventual goal of movement (Ambrosini et al., 2015; Rotman et al., 2006). Our results extend these findings from action observation to realistic interactions and show that even subtle preparatory movements before the start of an action are taken into account for action prediction in real time. The ability to read these subtle cues further attests that humans are expert body readers. On reflection, this is perhaps not surprising considering the fact that humans are a social species, and the demands of a social environment often require reading subtle cues from the body of other individuals to predict their intentions, emotions and other attributes.

What are the underlying mechanisms of this remarkable action prediction ability? Two main theories have been suggested for action prediction. The "simulation account" states that action prediction requires a simulation in the action production system (Flanagan & Johansson, 2003; Rizzolatti, Fogassi, & Gallese, 2001). These studies suggest the mirror neuron system as the source of action prediction (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Umiltà et al., 2001). The "inference account", on the other hand, states that action prediction does not require the involvement of the motor system, but is a cognitive process that infers the goal of the movement from the knowledge of the intentions of the agent as well as the context and visual information (Brass, Schmitt, Spengler, & Gergely, 2007; Csibra, 2008). Note that the two accounts may not be mutually exclusive (Ambrosini et al., 2015; Falck-Ytter, 2012). Our current results cannot provide evidence for or against either of these accounts, but they suggest that a simulation or an inference of an action should rely on kinematic cues gathered from any available body part and should happen with inputs as short as a fraction of a second. Future experiments using similar methods could directly test the two accounts in the context of realistic interactions.

Taken together these results show that humans rely on subtle preparatory cues distributed over the bodies of others to anticipate their goals. More broadly, our study could open the door to future studies of human interactions in real-time settings, to have a deeper understanding of human body movements and what humans are sensitive to during such interactions. These studies could, in turn, inform machine learning and modeling techniques to improve human-computer interface.

## Conflict of interest

The authors declare no competing financial interests.

## Acknowledgements

## Appendix A. Supplementary material

The raw data for this manuscript is saved in a public repository on GitHub (Vaziri-Pashkam, Cormiea, & Nakayama, 2017) located at: https://github.com/mvaziri/PredictingActions_RawData. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2017.06.014.

## References

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception-London, 30*(2), 233–252.

Abernethy, B., & Zawi, K. (2007). Pickup of essential kinematics underpins expert perception of movement patterns. *Journal of Motor Behavior, 39*(5), 353–367.

Abry, C., Cathiard, M. A., Robertribes, J., & Schwartz, J. L. (1994). The coherence of speech in audiovisual integration. *Cahiers de Psychologie Cognitive – Current Psychology of Cognition, 13*(1), 52–59.

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.

Ambrosini, E., Pezzulo, G., & Costantini, M. (2015). The eye in hand: Predicting others' behavior by integrating multiple sources of information. *Journal of Neurophysiology, 113*(7), 2271–2279.

Ansuini, C., Cavallo, A., Bertone, C., & Becchio, C. (2015). Intentions in the brain the unveiling of Mister Hyde. *The Neuroscientist, 21*(2), 126–135.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717–746.

Belen'kii, V. E., Gurfinkel, V. S., & Pal'tsev, E. I. (1967). On the control elements of voluntary movements. *Biofizika*.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology, 17*(24), 2117–2121.

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science, 15*(2), 292–298.

Chouchourelou, A., Matsuka, T., Harber, K., & Shiffrar, M. (2006). The visual analysis of emotional actions. *Social Neuroscience, 1*(1), 63–74.

Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In *Sensorimotor foundations of higher cognition. Attention and performance XXII* (pp. 435–459).

Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Attention, Perception, & Psychophysics, 44*(4), 339–347.

Diaz, G. J., Fajen, B. R., & Phillips, F. (2012). Anticipation from biological motion: The goalkeeper problem. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 848.

Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*(6), 727–738.

Falck-Ytter, T. (2012). Predicting other people's action goals with low-level motor information. *Journal of Neurophysiology, 107*(11), 2923–2925.

Farrow, D., & Abernethy, B. (2003). Do expertise and the degree of perception-action coupling affect natural anticipatory performance? *Perception-London, 32*(9), 1127–1140.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature, 424*(6950), 769–771.

Flanagan, J. R., Rotman, G., Reichelt, A. F., & Johansson, R. S. (2013). The role of observers' gaze behaviour when watching object manipulation tasks: Predicting and evaluating the consequences of action. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1628), 20130063.

Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Research, 1079*(1), 36–46.

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

Georgiou, I., Becchio, C., Glover, S., & Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition, 102*(3), 415–433.

Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *Neuroimage, 36*, T22–T32.

Hollerbach, J. M., & Flash, T. (1982). Dynamic interactions between limb segments during planar arm movement. *Biological Cybernetics, 44*(1), 67–77.

Huys, R., Smeeton, N. J., Hodges, N. J., Beek, P. J., & Wiliams, A. M. (2008). On the dynamic information underlying visual anticipation skill. *Perception & Psychophysics, 70*(7), 1217–1234.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Attention, Perception, & Psychophysics, 14*(2), 201–211.

Jokisch, D., & Troje, N. F. (2003). Biological motion as a cue for the perception of size. *Journal of Vision, 3*(4), 252–264.

Kandel, S., Boë, L. J., & Orliaguet, J. P. (1993, September). Visual detection of coarticulatory anticipation or . . . guessing what has not yet been written. In *Virtual reality annual international symposium, 1993, 1993 IEEE* (pp. 148–154). IEEE.

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *Journal of Experimental Psychology, 51*(3), 179.

Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science, 12*(6), 467–472.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics, 21*(6), 575–580.

Lewkowicz, D., Quesque, F., Coello, Y., & Delevoye-Turrell, Y. N. (2015). Individual differences in reading social intentions from motor deviants. *Frontiers in Psychology, 6*.

Louis-Dam, A., Orliaguet, J. P., & Coello, Y. (1999). Perceptual anticipation in grasping movement: When does it become possible. *Studies in Perception and Action, V*, 135–139.

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.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research, 211*(3–4), 547–556.

Martel, L., Bidet-Ildei, C., & Coello, Y. (2011). Anticipating the terminal position of an observed action: Effect of kinematic, structural, and identity information. *Visual Cognition, 19*(6), 785–798.

Müller, S., Abernethy, B., & Farrow, D. (2006). How do world-class cricket batsmen anticipate a bowler's intention? *The Quarterly Journal of Experimental Psychology, 59*(12), 2162–2186.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. WH Freeman/Times Books/Henry Holt & Co.

Orliaguet, J. P., Kandel, S., & Boë, L. J. (1997). Visual perception of motor anticipation in cursive handwriting: Influence of spatial and movement information on the prediction of forthcoming letters. *Perception-London, 26*, 905–912.

Parkinson, J., Springer, A., & Prinz, W. (2011). Can you see me in the snow? Action simulation aids the detection of visually degraded human motion. *The Quarterly Journal of Experimental Psychology, 64*(8), 1463–1472.

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), 421–433.

Pennel, I., Coello, Y., & Orliaguet, J. P. (1999). Visual perception of motor anticipation in sign language. *Studies in Perception and Action*, 144–148.

Pesquita, A., Chapman, C. S., & Enns, J. T. (2016). Humans are sensitive to attention control when predicting others' actions. *Proceedings of the National Academy of Sciences, 113*(31), 8669–8674.

Prablanc, C., Echallier, J. F., Komilis, E., & Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. *Biological Cybernetics, 35*(2), 113–124.

Quesque, F., & Coello, Y. (2015). Perceiving what you intend to do from what you do: Evidence for embodiment in social interactions. *Socioaffective Neuroscience & Psychology, 5*.

Quesque, F., Delevoye-Turrell, Y., & Coello, Y. (2016). Facilitation effect of observed motor deviants in a cooperative motor task: Evidence for direct perception of social intention in action. *The Quarterly Journal of Experimental Psychology, 69*(8), 1451–1463.

Ranganathan, R., & Carlton, L. G. (2007). Perception-action coupling and anticipatory performance in baseball batting. *Journal of Motor Behavior, 39*(5), 369–380.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*(2), 131–141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*(9), 661–670.

Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of Neurophysiology, 96*(3), 1358–1369.

Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance, 7*(4), 733.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General, 112*(4), 585.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition, 119*(2), 242–252.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science, 1*(2), 353–367.

Song, J. H., & Nakayama, K. (2006). Role of focal attention on latencies and trajectories of visually guided manual pointing. *Journal of Vision, 6*(9), 11.

Starkes, J. L., Edwards, P., Dissanayake, P., & Dunn, T. (1995). A new technology and field test of advance cue usage in volleyball. *Research Quarterly for Exercise and Sport, 66*(2), 162–167.

Tomeo, E., Cesari, P., Aglioti, S. M., & Urgesi, C. (2012). Fooling the kickers but not the goalkeepers: Behavioral and neurophysiological correlates of fake action detection in soccer. *Cerebral Cortex, bhs279*.

Troje, N. F., Sadr, J., Geyer, H., & Nakayama, K. (2006). Adaptation aftereffects in the perception of gender from biological motion. *Journal of Vision, 6*(8), 7.

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Attention, Perception, & Psychophysics, 67*(4), 667–675.

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), 155–165.

Vaziri-Pashkam, M., Cormiea, S., & Nakayama, K. (2017). *PredictingActions_RawData, Github repository*. <https://github.com/mvaziri/PredictingActions_RawData>.