Humans are sensitive to attention control when predicting others' actions

Ana Pesquita[a,1], Craig S. Chapman[b], and James T. Enns[a,c]

[a]Department of Psychology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4; [b]Faculty of Physical Education and Recreation, University of Alberta, Edmonton, AB, Canada T6G 2H9; and [c]Department of Psychology, University of Western Australia, Crawley, WA 6009, Australia

Studies of social perception report acute human sensitivity to where another's attention is aimed. Here we ask whether humans are also sensitive to how the other's attention is deployed. Observers viewed videos of actors reaching to targets without knowing that those actors were sometimes choosing to reach to one of the targets (endogenous control) and sometimes being directed to reach to one of the targets (exogenous control). Experiments 1 and 2 showed that observers could respond more rapidly when actors chose where to reach, yet were at chance when guessing whether the reach was chosen or directed. This implicit sensitivity to attention control held when either actor's faces or limbs were masked (experiment 3) and when only the earliest actor's movements were visible (experiment 4). Individual differences in sensitivity to choice correlated with an independent measure of social aptitude. We conclude that humans are sensitive to attention control through an implicit kinematic process linked to empathy. The findings support the hypothesis that social cognition involves the predictive modeling of others' attentional states.

social perception | attention | action prediction | autism spectrum | action observation

Imagine giving a sales pitch when you notice a potential buyer reaching for her phone. Whether she has shifted her attention to the phone voluntarily or whether she did so because it blinked unexpectedly is critical information about her mental state and possibly about the success of your sales pitch. Here we report that humans can implicitly distinguish between these two kinds of attention control in the observed actions of others.

We already know that humans are remarkably sensitive to where someone is attending (1, 2). This ability not only holds important clues to dangers and opportunities in the environment, but it contributes to a representation of the other's mental state [i.e., a theory of mind (3, 4)]. However, do these representations only fill out the content of the other's mind, or do they also hold information on the control of that content? A recent theory proposes that social awareness involves the predictive (forward) kinematic modeling of other people's attention (5–7). These models include the nature of control, so that the spatial and temporal consequences of an attentional state can be predicted in the actions of others before they occur.

The control of attention is among the most widely studied topics in all of cognitive science (8–10). Attention is endogenous when we voluntarily decide to act on an event in the environment; it is exogenous when the action is governed by environmental factors, such as a sudden local change in appearance or sound. Here we studied observers' sensitivity to attention control. We presented observers with videos of actors reaching to one of two possible targets while either choosing (endogenous control) or being directed (exogenous control) to one target.

Fig. 1A illustrates the reaching task from the actor's perspective. Because our goal was to test for sensitivity to how the reaches were controlled—not sensitivity to overt differences in the onsets or movement times of the reaches—we eliminated temporal cues that might distinguish chosen from directed actions, and we randomized the trials shown to observers to eliminate any trial-to-trial contingencies. How we accomplished these design goals is described in

detail in *SI Methods*, which describes how the stimulus set of videos was generated, along with the temporal and kinematic characteristics of the reaches shown in the videos (Tables S1 and S2). The analyses of the actors' video clips also compared kinematic features of the reaches, to confirm that they varied as we expected between chosen (endogenous) and direct (exogenous) reaches. In short, chosen reaches tended to take longer to achieve peak acceleration, and the limb trajectories toward the targets were more curved than direct reaches, reflecting their greater decisional uncertainty (11).

We presented chosen and directed videos to observers and ask them to predict the target of the actor's actions. Two alternative hypotheses were considered. If observers based their predictions solely on the kinematic cues of the reaching actions, they should fare better on directed trials, because those reaches take less time to reach peak acceleration and move more directly through space to the target location. We call this the "physical signal hypothesis" and contrast it with what we call the "social prediction hypothesis." In the social prediction hypothesis, choice actions follow more naturally and predictably from the prechoice mental and postural states of an actor than directed actions (12). Actions that are directed by an unpredictable external signal are less likely to be congruent with the actor's recent mental and postural history. Thus, if observers can capitalize on bodily cues reflecting the actors' internal biases toward one target, they should fare better on chosen trials, because these early cues predict the actors' ultimate target choice (13).

In the following experiments, we report empirical support for the social prediction hypothesis by pursuing five specific questions: (*i*) Are humans sensitive to endogenous vs. exogenous attention control in others? (*ii*) Is sensitivity to attention control consciously accessible to observers? (*iii*) Where on the actor's body can the attention control signal be seen? (*iv*) How early in the time course of observed actions is the attention control

Fig. 1. Illustrations of the method from the actors' and observers' perspectives. (*A*) Illustration of the method from the actors' perspective. Actors were filmed through Plexiglas reaching to two possible targets. On chosen trials, both locations were lit ,and actors had to choose (not shown); on directed trials only one location was lit, and actors were directed to reach to that location (as shown). (*B*) Illustration of the method from the observers' perspective. Observers responded to each video by pressing a spatially mapped key as rapidly as possible to indicate where the actor was reaching.

signal available? (*v*) Is sensitivity to attention control linked to social aptitude?

Experiment 1: Are Humans Sensitive to Attention Control in Others?

Fig. 1*B* illustrates the person perception task from the observer's perspective. The first experiment tested the sensitivity of observers to the actors' attention control states by asking observers to indicate rapidly whether the target of the actor's reach was left or right. The fundamental question is whether observers would be faster to predict the end-target of an actor's movements when they were chosen vs. directed.

The results showed that observers were faster to discriminate the location of an actor's reach when it was chosen than when it was directed. This difference in reaction times is surprising when one considers that the reaching kinematics favored the direct reaches. However, it is consistent with the claim that social awareness involves a predictive model of the attentional state of others (5–7).

This model not only includes information about where the other is attending, but it includes ongoing predictions about the decision being undertaken by the other.

Methods.

Observers. Thirty participants (18 female, 4 left-handed) with a mean age of 21.9 (SD = 4.6) were recruited from the University of British Columbia (UBC) Human Subject Pool to serve as observers. The only exclusion criterion was failing to report normal or corrected to normal vision. Observers received partial course credit in exchange for 1 h of time, as approved by the UBC Behavioral Research Ethics Board. All participants read and signed a written informed consent document before testing. The document described the procedures, informed participants they would receive partial credit in a qualifying psychology course, and that they could withdraw from participation at any point without penalty.

Procedure. Fig. 1*B* illustrates the experiment from the observer's perspective. Observers sat at a desktop computer, with their task being to press one of two keys, spatially mapped to the target locations, as rapidly as possible. Observers were instructed to treat this task as a competitive game in which they could "beat the actor" by making the correct response before the actor's finger reached the target location. However, they were also told to minimize their errors by making no more than 10–20% errors overall. Each trial began with the observer's two index fingers resting on these keys and their eyes on a fixation cross for 1–1.5 s. The presentation of the fixation cross was followed by a video clip showing an actor reaching for a target and the observer's response (Movies S1–S16).

The session began with eight practice trials involving an actor who was not used in the main test. Observers were told that actors would reach left and right an equal number of times and at random. The 100 trials for each actor were shown in a single block, in counterbalanced order across observers, and observers were given a short break between each of the four blocks of trials. At the conclusion of the session, observers completed the 50-item Autism-Spectrum Quotient (AQ) (14).
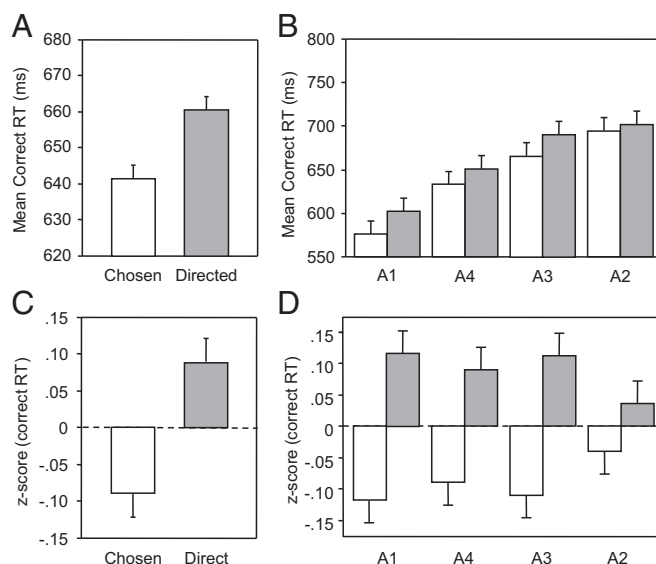
Fig. 2. Observers' mean responses in experiment 1 (*n* = 30). (*A*) Mean correct response time (RT). Error bars are ±1 SEM, following the Loftus & Masson (27) procedure for within-subjects designs. (*B*) Mean correct RT for each of the four actors. (*C* and *D*) The data in *A* and *B* after each observer's correct RT has been converted to z-scores to standardize the distributions for individual differences in mean speed and variance.

Results. Fig. 2 shows the mean correct response time (RT) in the chosen and directed conditions overall (Fig. 2*A*) and for each of the four actors ranked by the speed with which observers could discriminate whether they were pointing left or right (Fig. 2*B*). Fig. 2 *C* and *D* show the data after each observer's correct RT had been converted to z-scores to control for the larger differences in the mean speed and variance of the four actors' reaches (Fig. 2*B*). Both of these analyses make it clear that RT was faster in the chosen than in the directed condition for each of the four actors (A1–A4). This conclusion was supported by the following analyses.

Incorrect trials and responses >3 SDs from the mean were excluded. Response accuracy, correct RT, and z-scores of correct RT were each subjected to repeated-measures ANOVA examining the effect of condition (chosen or directed) and actor (A1–A4). Z-scores were computed on the correct RT values by subtracting each observer RTs from the mean RTs of that observer to the corresponding actor, and dividing this difference by the SD of the observer's RTs for this actor.

Observers responded correctly on 81% of trials (SEM = 0.7%), with significant differences in accuracy between actor videos [$F(3, 87) = 15.31$, $P < 0.001$, $\eta^2 = 0.346$ (in rank order A3 = 85%, A2 = 83%, A4 = 81%, and A1 = 75%)], but no differences between condition ($P > 0.25$) or an interaction ($P > 0.09$). Analysis of correct RT indicated significant main effects of condition, with responses to chosen reaches made significantly faster than responses to directed reaches [$F(1, 29) = 70.39$, $P < 0.001$, $\eta^2 = 0.708$; actor $F(3, 87) = 31.48$, $P < 0.001$, $\eta^2 = 0.521$; and interaction, $F(3, 87) = 3.21$, $P < 0.03$, $\eta^2 = 0.100$].

To test whether the choice advantage was influenced by observer accuracy, we included overall accuracy as a between-subjects factor, after dividing the participants into more accurate (mean accuracy = 93% correct) and less accurate (mean accuracy = 69% correct) halves. This analysis indicated no interaction of condition × accuracy [$F(1, 28) = 1.38$, $P < 0.25$, $\eta^2 = 0.012$], with both groups showing a 19-ms advantage in the chosen condition. This result indicates that the difference between responses to chosen and directed actions is not due to a speed–accuracy trade-off.

Z-scores of correct reaction times of observers to each actor were computed to consider the effect of condition after controlling for the large variability in reaching behavior between actors. In these analyses, the main effect of actor was no longer significant, but there was a main effect of condition [$F(1, 29) = 80.51$, $P < 0.001$, $\eta^2 = 0.735$]. In the experiments that follow, we undertake similar analysis of accuracy, correct RT, and z-scores, but for simplicity, we will only present graphs showing the mean z-scores and their SEs. None of the conclusions differed depending on whether an analysis was based on raw RT or on z-scores.

Experiment 2: Is Sensitivity to Attention Control Consciously Accessible to Observers?

Experiment 1 showed that observers' speeded responses to actors' reaches are faster when the target of the reach is chosen rather than directed. However, it is one thing for a social prediction model to influence kinematic behavior (i.e., the observer's spatially mapped response); it is another to have this information accessible at a conscious level. In this experiment, we replicated the conditions of the previous one, but, in addition, we asked whether observers could correctly report the attention control state of the actors.

We found no evidence, either in the observers as a group or among individual observers, that their explicit attempts to discriminate chosen from directed actions exceeded the chance level of guessing. However, these same observers were able to distinguish these two types of reaches in their speeded kinematic responses. This pattern of findings implies that sensitivity to attention control influences an observer's action, but that it is not accessible to the observer's conscious awareness.

Methods. The method in experiment 2 was identical to experiment 1 with the following exceptions: (*i*) Thirty different observers (10 female, 2 left-handed) with mean age of 23.1 (SD = 4.3) served as observers. (*ii*) In the instructions, observers were shown the actor's perspective in the video and informed that on a random half of the trials the actor had chosen which target to reach and on the other half of trials they were directed. (*iii*) After the observer's speeded response to indicate the direction of the actor's reach, observers were asked to make a second response, indicating whether they believed the reach had been chosen or directed. They responded to this question ("Did the actor choose where to point?") by pressing one of two specially marked keys at the top of the keyboard marked as "yes" and "no."

Results. Fig. 3 shows the mean z-scores of correct RT in the chosen vs. directed conditions (Fig. 3*A*) and shows the proportion of hits and false alarms observers made in response to the question of whether the video they had just responded to represented a chosen or directed trial (Fig. 3*B*), after rank-ordering observers in terms of their response biases from conservative (reluctant to respond "chosen") to liberal (reluctant to respond "direct"). See Dataset S1 for data coded for hit and false alarm responses. These data show that the main finding of experiment 1 replicated under these conditions (i.e., correct responses were faster on chosen than directed trials) but that observers were unable to report whether the actors they were responding to were chosen or not. These conclusions were supported by the following analyses.

Observers responded correctly on 78% of trials (SEM = 0.8%), with significant differences in accuracy between actor videos [$F(3, 87) = 5.84$, $P < 0.001$, $\eta^2 = 0.169$ (in rank order A3 = 79%, A2 = 79%, A4 = 78%, and A1 = 73%)], but no differences between condition or any interaction ($P > 0.50$). Analysis of correct RT indicated significant main effects of condition [$F(1, 29) = 23.42$, $P < 0.001$, $\eta^2 = 0.447$] and actor [$F(3, 87) = 34.67$, $P < 0.001$, $\eta^2 = 0.545$]. Examination of the relation between the choice advantage and accuracy indicated that the mean choice advantage was 21 ms for the 15 participants who were most accurate (mean accuracy = 92% correct) and only 7 ms for the 15 participants who were least accurate (mean accuracy = 64% correct) [$F(1, 28) = 7.24$, $P < 0.01$, $\eta^2 = 0.136$]. Analysis of z-scores also indicated a main effect of condition [$F(1, 29) = 14.74$, $P < 0.001$, $\eta^2 = 0.337$].

Analyses of the proportion of hits and false alarms in response to the question of whether a video represented a chosen or directed trial revealed no significant differences, either when the data were aggregated as a group or for any observer individually (all $P > 0.25$).

We also replicated this insensitivity in explicit reports in a new sample of 30 observers, who (*i*) were not asked to predict the target locations and (*ii*) were given trial-by-trial accuracy feedback on
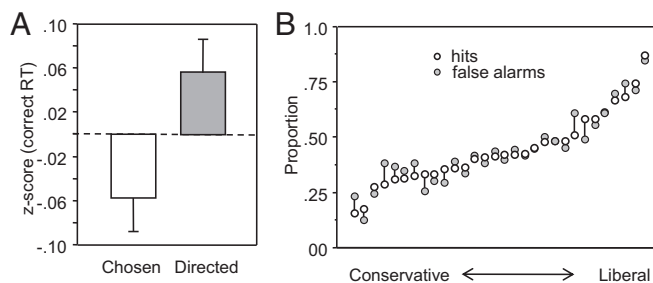
Fig. 3. Observers' mean responses in experiment 2 (*n* = 30). (*A*) Mean z-scores of correct RT in experiment 2. Error bars are ±1 SEM, following the Loftus & Masson (27) procedure for within-subjects designs. (*B*) The proportion of hits and false alarms of observers trying to discriminate chosen from directed trials, after rank-ordering observer's response biases from conservative (reluctant to respond "chosen") to liberal (reluctant to respond "direct").

their guesses about whether the observer was choosing or reacting on each trial, so that they could devote their full attention to the task. The results were the same. Not a single one of the observers had a hit rate that differed significantly from their false-alarm rate.

Experiment 3: Where on the Actors' Body Can the Attention Control Signal Be Seen?

Extant theories of social cognition have focused on the eyes as the primary source of information about social attention (3, 15). More recent evidence suggests that head and body position also play a role (6, 16). In experiment 3, we investigated where the control signal is coming from in the video clips of the actors. To do so, we selectively masked either the head (leaving the torso and limbs visible) or the body (leaving only the head visible) of the actors, while again asking observers to make a speeded response to the target of the actor's reach.

The results showed that observers' sensitivity to attention control cues was robustly resistant to the occlusion of actors' body parts. The signal was available in both the head and body conditions, suggesting that the cues to the decision are distributed throughout the body. This result is in line with other research on the bodily clues regarding people's intentions (17). For example, how one reaches for a Lego piece allows a partner to predict the intention to cooperate or compete (18). The kinematics of running reveals the intention to deceive a sports opponent (19). Observers are able to perceive the value of the poker hand in the arm kinematics of players (20). The present adds to this previous work by showing that observers are sensitive to the bodily cues of attention control. It will be important in future studies to record the kinematics of actors beyond their end effector (i.e., finger movement), perhaps by using point-light displays to isolate features of bodily movements that carry the signal of attention control.

Methods. The method in experiment 3 was identical to experiment 1 with the following exceptions. (*i*) There were 30 different observers (24 female, all right-handed) with mean age of 21.1 y old (SD = 2.17). (*ii*) The 400 videos were each shown twice, once showing only the actor's head (including face, neck, and eyes) and once showing only the actor's body (torso and arms). Head and body videos were randomly interspersed in each block of trials. (*iii*) The display monitor was an active touch surface and was much larger (83 × 67 cm), such that the actor videos were approximately life size. (*iv*) Observers began each trial with the index finger of their right hand at a center home position marked on the table. They responded to each video by reaching as rapidly as possible to the target location they thought the actor was reaching toward. The instructions were to beat the actor if at all possible, without making >10–20% errors. We recorded the observer's reach initiation time and movement time on each trial using Optotrack to sample the 3D position of the right index finger at 200 Hz.

Results. Fig. 4 shows the mean z-scores of correct RT in the chosen vs. directed conditions, separately for trials in which only the body and limbs were visible vs. when only the head was visible. These data show that observers were more sensitive to the difference between chosen and directed trials when the body and limbs were visible than when the head was visible. These conclusions were supported by the following analyses.

Observers responded correctly on 78% of trials (SEM = 0.6%), with significant differences in accuracy between actor videos [$F_{(3, 87)} = 21.23$, $P < 0.001$, $\eta^2 = 0.423$ (in rank order A4 = 81%, A3 = 81%, A2 = 74%, and A1 = 74%)], but not between conditions ($P > 0.09$). Response accuracy was also significantly greater when the body was visible (mean = 82%) than when the head was visible (mean = 73%) [$F_{(1, 29)} = 37.06$, $P < 0.001$, $\eta^2 = 0.561$].

Analysis of correct RT indicated that chosen trials were faster by 11 ms than direct trials [$F_{(1, 29)} = 12.76$, $P < 0.02$, $\eta^2 = 0.306$]; there were actor differences [$F_{(3, 87)} = 43.35$, $P < 0.001$, $\eta^2 = 0.599$];
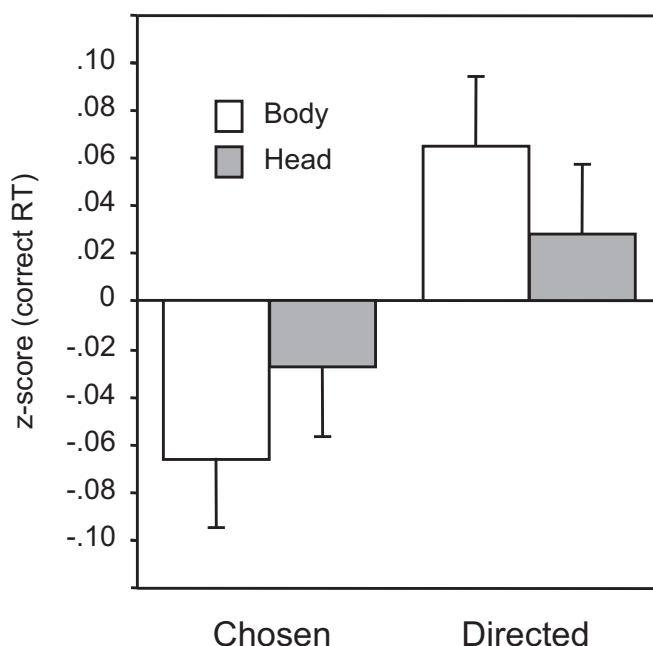

Fig. 4. Mean z-scores of correct RT in experiment 3 (*n* = 30), separately for trials in which the body and limbs were visible vs. when only the head was visible. Error bars are ±1 SEM, following the Loftus & Masson (27) procedure for within-subjects designs.

and responses when the body was visible were faster by 134 ms than when only the head was visible [$F_{(1, 29)} = 76.48$, $P < 0.001$, $\eta^2 = 0.725$]. Responses to choice movements were faster than responses to direct movements by 14 ms when the body was visible and 8 ms when the head was visible [$F_{(1, 29)} = 1.44$, $P < 0.25$, $\eta^2 = 0.047$], but the responses on head trials were also slower (134 ms) and more variable (SE of 12 ms vs. only 6 ms for body trials). Analysis of z-scores, which controlled for these differences, indicated a significant advantage on chosen over direct trials [$F_{(1, 29)} = 18.89$, $P < 0.001$, $\eta^2 = 0.394$], with this effect being significantly larger when the body was visible than when only the head was visible [$F_{(1, 29)} = 5.84$, $P < 0.02$, $\eta^2 = 0.168$]. Examination of the relation between the choice advantage and accuracy indicated that the choice advantage was larger for the 15 participants who were most accurate (mean accuracy = 86%, mean z-score difference = 0.134) than for the 15 participants who were least accurate (mean accuracy = 69%, mean z-score difference = 0.050) [$F_{(1, 28)} = 4.50$, $P < 0.04$, $\eta^2 = 0.084$].

Experiment 4: How Early Is the Attention Control Signal Available?

In experiment 4, we examined the time course of sensitivity to attention control by presenting the actor's videos cut at different lengths and asking observers to indicate the likely end target of the actor's actions from these brief segments. This question was guided by the theoretical idea that modeling of another's attentional state requires the observer to predict the actor's behavior even before it begins (6, 7) and by previous results showing that early prediction is essential in social coordination tasks (21).

The results showed that the advantage for responding on choice trials was evident in the first 100 ms of processing. This finding means that observers were able to use the preparatory movements that preceded the actor's reach to make a target location prediction. This finding is consistent with theories emphasizing the predictive nature of modeling social attention (5–7), which means that the sooner one can predict another's action, the more time they will have to consider and execute appropriate reactions (7, 22).
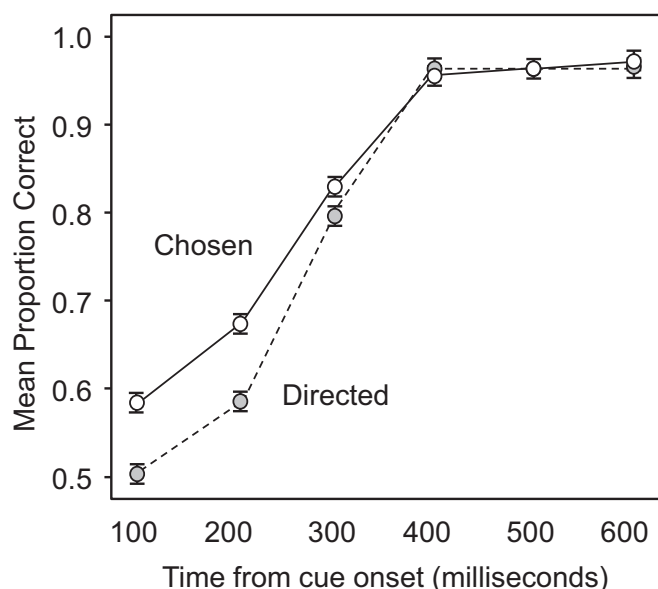
Fig. 5. Mean proportion correct response in experiment 4 (*n* = 30). Error bars are ±1 SEM, following the Loftus & Masson (27) procedure for within-subjects designs.

The musculoskeletal constraints of the body require that moving one limb often requires the activation of other body parts. For example, initiating an arm-reaching movement requires the engagement of the shoulders, torso, and even the lower limbs to make the necessary postural adjustments to stabilize the body (23). Humans appear to have implicit knowledge of these biomechanical principles and use this knowledge to predict others' actions. For example, basketball experts are able to predict the end result of a shot before the ball leaves the athlete's hand (22). Observers of a soccer player are able to predict the kick direction before the foot-to-ball contact (24). Deception in rugby runners is detected above chance before the runner changes direction (19). More closely related to the present task, a competitive reaching study showed that preparatory cues (i.e., movements and postural configurations preceding the lift-off of the finger) give opponents an advantage (13). The present findings extend this evidence by showing that observers are sensitive to the attention-control states engaging preparatory movements.

Method. The method in experiment 4 was identical to experiment 1 with the following exceptions. (*i*) There were a total of 30 different observers (17 female, 3 left-handed) with mean age of 22.71 y old (SD= 3.43). (*ii*) Using the same pool of videos as in previous experiments, we cut each video at six different lengths from the onset of the cue (0–100 ms to 0–600 ms, in 100-ms steps). (*iii*) Videos were randomly sampled from this pool on each trial. Observers reported the likely end target of the actor's reach, so percentage correct became the dependent measure. Because this method involved guessing on many trials when the segments were short, speed of responding was not emphasized. (*iv*) Observers completed two blocks of 600 trials, separated by a short break. Each block consisted on the presentation of 100 videos from a single actor, and the two actors selected for each observer were counterbalanced across observers.

Results. Fig. 5 shows the mean proportion of correct responses in the chosen and directed conditions as a function of the time from the onset of the actor's cue. These data show that observers can predict the target location more accurately for the chosen than the directed condition at the shortest two video lengths. This conclusion

was supported by an ANOVA indicating significant main effects of condition [$F(1, 29) = 23.90$, $P < 0.001$, $\eta^2 = 0.452$] and time [$F(5, 145) = 1149.99$, $P < 0.001$, $\eta^2 = 0.975$] and an interaction [$F(5, 145) = 27.54$, $P < 0.001$, $\eta^2 = 0.487$]. Simple effects testing indicated that the chosen advantage in accuracy was significant at 100 and 200 ms (both $P < 0.01$), but not at the longer time bins (all $P > 0.15$).

Question 5. Is Sensitivity to Attention Control Linked to Social Aptitude?

If the sensitivity of observers' responses to the attentional state of actors reflects the mental modeling of social attention, then individual differences in the strength of this sensitivity may be related to social aptitude on a broad scale. To test this hypothesis, we correlated individual differences in social sensitivity to attention control with self-reported social aptitude, as measured by the AQ (14).

The analyses indicated that observers with higher social aptitude also exhibit stronger sensitivity to attention-control states in their kinematic responses. This finding bolsters the hypothesis that sensitivity to attention control arises from the involuntarily tendency for humans to model the attentional states of others (3, 4, 6), because the sensitivity is observed in people with generally greater social empathy and communication skills.

Method. We asked observers in all four experiments to fill out the 50-item AQ (14), which captures variation in the tendency toward autistic traits in the general population. Individuals with higher level of autistic-like traits show a nonclinical propensity to empathize less strongly with others and to engage in systemized thinking (e.g., great attention to detail, rigid interests), whereas individuals with lower levels of autistic traits display the opposite cognitive profile. One observer in experiment 3 did not complete the AQ questionnaire.

Results. To examine possible relationships between observers' social aptitude and their sensitivity to the attentional state of actors, we assigned each observer a sensitivity score based on their mean difference in z-scores between the directed and chosen conditions. In experiments 1 and 2, this overall score consisted of the mean difference in z-scores between chosen and directed trials across all four actors. In experiment 3, we used the mean difference score only for the body condition, which provided a stronger and more reliable signal than the head condition, and in experiment 4, we used the mean difference score in the 100- and 200-ms time bins, where the signal was strongest.
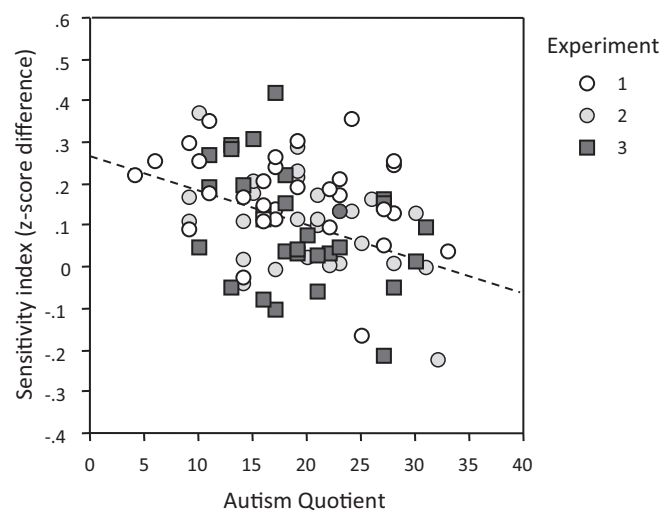
Fig. 6. Scatterplot of the relation between observer's speeded sensitivity scores in experiments 1–3 and their AQ scores (*n* = 89).

Fig. 6 shows a scatterplot of observers' speeded sensitivity score in experiments 1–3 and their AQ scores. Experiments 1–3 portrayed a negative relationship between the measure of speeded response sensitivity and the AQ [$r(28) = -0.284$, $P = 0.13$, $r(28) = -0.478$, $P = 0.01$; and $r(27) = -0.387$, $P = 0.04$, respectively], but there was almost no correlation in experiment 4, where response sensitivity was measured in accuracy rather than speed [$r(28) = -0.004$]. This finding is consistent with observers with greater social aptitude being able to respond more rapidly to an actor who is selecting their reach with intention rather than reacting. The correlation over all observers in experiments 1–3 was $r(87) = -0.366$ ($P < 0.001$).

Conclusion

This study offers evidence on the perceptual mechanisms underlying social cognition. When observers were given the opportunity to predict the location of a videotaped actor's reach, they were faster to do so when the actor was deciding where to reach than when the actor was being directed by an external cue. This result was observed despite our care in removing all temporal cues from the sampling of the actor's reaches and in randomizing the two types of reaches shown to observers. This finding implies that the decision undertaken by the actor is visible to the observer before being executed by the actor. However, tests of whether the observer's sensitivity to the actor's choice was consciously accessible were negative. Tests of where the signals about the actor's choices were coming from indicated that it was widely distributed over the body, though stronger in the torso and limbs than in the head. Tests of when the signal was available indicated that it was influential even before the actor's limb started moving. Finally, sensitivity in the speeded decisions of observers was correlated with a paper-and-pencil measure of social aptitude.

These findings are consistent with recent theoretical proposals claiming that social awareness involves the predictive (forward) kinematic modeling of the action consequences of others' attentional states (5–7). With regard to the actors in the present study, the results show that early kinematic cues in the execution of chosen actions carry predictive information about the actor's ultimate choice. This observation is consistent with evidence indicating that action components are not independent of one another; at any moment in time, internal mental biases and existing bodily states unconsciously influence the unfolding of the subsequent movements in a sequence (25). The results of this study therefore support the hypothesis that chosen actions follow more naturally and predictably from the prechoice mental and postural states of the actor than directed actions. Observing the stream of consistent kinematic cues in an actor's chosen behavior is therefore what we believe underlies the ability to predict the outcome of the reach earlier in time.

With regard to observers in the present study, the findings support the general idea that the brain is a prediction machine (25, 26), which, in the realm of social observation, means that we are continuously gathering information to update models of others' internal states, so that we can predict what they will do next as soon as possible (5–7). The main finding of this study is that is easier to do for most observers when actors are choosing to act rather than being directed externally. The secondary findings (*i*) that this sensitivity to choice in the kinematics of others is not consciously accessible to observers, but (*ii*) that it is correlated with an independent measure of social aptitude in everyday life, bolsters the view that social action observation is a fast and implicit kinematic process linked to empathy.

1. Friesen CK, Kingstone A (1998) The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychon Bull Rev* 5(3):490–495.
2. Langton SR, Bruce V (2000) You must see the point: Automatic processing of cues to the direction of social attention. *J Exp Psychol Hum Percept Perform* 26(2):747–757.
3. Baron-Cohen S (1995) The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. *Joint Attention: Its Origins and Role in Development*, eds Moore C, Dunham PJ (Erlbaum, New York), pp 41–59.
4. Calder AJ, et al. (2002) Reading the mind from eye gaze. *Neuropsychologia* 40(8):1129–1138.
5. Webb TW, Graziano MSA (2015) The attention schema theory: A mechanistic account of subjective awareness. *Front Psychol* 6:500.
6. Graziano MSA (2013) *Consciousness and the Social Brain* (Oxford Univ Press, New York).
7. Graziano MSA, Kastner S (2011) Human consciousness and its relationship to social neuroscience: A novel hypothesis. *Cogn Neurosci* 2(2):98–113, 10.1080/17588928.2011.565121.
8. Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32(1):3–25.
9. Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3(3):201–215.
10. Posner MI, Rothbart MK (2007) Research on attention networks as a model for the integration of psychological science. *Annu Rev Psychol* 58:1–23.
11. Gallivan JP, Chapman CS (2014) Three-dimensional reach trajectories as a probe of real-time decision-making between multiple competing targets. *Front Neurosci* 8(215):215.
12. Rosenbaum DA, Herbort O, van der Wel R, Weiss DJ (2014) What's in a grasp? *Am Sci* 102(5):366.
13. Cormiea S, Vaziri-Pashkam MKN (2015) Unconscious reading of an opponent's goal. *J Vis* 15(12):43.
14. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E (2001) The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord* 31(1):5–17.
15. Perrett DI, Emery NJ (1994) Understanding the intentions of others from visual signals: Neurophysiological evidence. *Curr Psychol Cogn* 13(5):683–694.
16. Langton SR, Watt RJ, Bruce I (2000) Do the eyes have it? Cues to the direction of social attention. *Trends Cogn Sci* 4(2):50–59.
17. Becchio C, Manera V, Sartori L, Cavallo A, Castiello U (2012) Grasping intentions: From thought experiments to empirical evidence. *Front Hum Neurosci* 6(117):170–175.
18. Manera V, Becchio C, Cavallo A, Sartori L, Castiello U (2011) Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Exp Brain Res* 211(3-4):547–556.
19. Mori S, Shimada T (2013) Expert anticipation from deceptive action. *Atten Percept Psychophys* 75(4):751–770.
20. Slepian ML, Young SG, Rutchick AM, Ambady N (2013) Quality of professional players' poker hands is perceived accurately from arm motions. *Psychol Sci* 24(11):2335–2338.
21. Sebanz N, Knoblich G (2009) Prediction in joint action: What, when, and where. *Top Cogn Sci* 1(2):353–367.
22. Agliotti SM, Cesari P, Romani M, Urgesi C (2008) Action anticipation and motor resonance in elite basketball players. *Nat Neurosci* 11(9):1109–1116.
23. Hollerbach MJ, Flash T (1982) Dynamic interactions between limb segments during planar arm movement. *Biol Cybern* 44(1):67–77.
24. Diaz GJ, Fajen BR, Phillips F (2012) Anticipation from biological motion: The goalkeeper problem. *J Exp Psychol Hum Percept Perform* 38(4):848–864.
25. Hawkins J, Blakeslee S (2007) *On Intelligence* (Macmillan, London).
26. Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 36(3):181–253.
27. Loftus GR, Masson MEJ (1994) Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1(4):476–490.

Supporting Information

Pesquita et al. 10.1073/pnas.1601872113

SI Methods

Stimulus Construction. We filmed actors reaching to one of two possible targets while either choosing (endogenous control) or being directed to (exogenous control) one target (Fig. 1*A*). Actors were recruited from the same population as observers in the following experiments. A total of 11 potential actors were filmed. Five actors were excluded due to technical failures in the recordings. From the remaining six actors, we selected four (two females, ages 19–21) who followed instructions in all respects and consented to having their reaches recorded for presentation to other participants as stimulus materials. Movies S5, S6, S7, and S8 are available upon request.

Actors were seated at a table facing a Plexiglas panel positioned 56 cm from the table edge. Actors were filmed at 50 fps, 800×800 pixels, by using a Flea3 camera placed on the opposite of the Plexiglas. Two LED lights facing the actor served as cues. The LED lights were positioned 20 cm to the right and left side of a central fixation point located at the average actor's eye level. The videos started at cue presentation and ended after the reach was completed. Actors were instructed to begin each trial by fixating the central point. The fixation point was followed by the simultaneous onset of an auditory beep and the visual cue(s). On directed trials, one of the two LEDs was illuminated randomly, and actors were instructed to reach and touch it as rapidly as possible; on chosen trials, both LEDs were lit, and the instructions were to rapidly choose one LED to touch. Actors were instructed to make each choice in the moment and to try to select the left and right LEDs approximately equally often, which they did (50.87% right overall). The intertrial interval was kept deliberately short (1,000–1,500 ms after each response) to prevent strategic choosing in advance of the cue. Each actor completed a total of 100 trials in both the chosen and directed conditions. Critically, observers could not see the cues for chosen vs. directed actions that were visible to the actors (see Movies S1–S16 for examples of these recordings).

Equating Initiation and Movement Time. Because our goal was to test for sensitivity to how the reaches were controlled—not sensitivity to overt differences in the onsets or movement times of the reaches—we first eliminated temporal cues that might distinguish chosen from directed actions. From a pool of 800 video clips (4 actors × 200 trials), we first selected 100 videos at random from each actor and ranked them according to their initiation and movement times. The t tests evaluated whether there were significant differences in either initiation or movement times. If a test was positive, the videos in the tails of the distribution were replaced by randomly selected from the remaining videos until no differences remained.

One hundred test clips were selected for each actor (400 total), with an equal number of chosen and directed reaches that were not significantly different in initiation time [$t(49) = -0.81; 1.49; 0.29; -0.95$] or movement time [$t(49) = 0.06; -0.87; -0.78; 0.10$] for actors 1–4, respectively. However, there were still naturally occurring differences between actors in their overall initiation time [$F(3, 392) = 75.09, P < 0.001, \eta^2 = 0.363$ (means in rank order A3 = 302 ms, A2 = 299 ms, A4 = 282 ms, and A1 = 205 ms)] and movement time [$F(3, 392) = 771.23, P < 0.001, \eta^2 = 0.855$ (means in rank order A2 = 757 ms, A4 = 619 ms, A3 = 589 ms, and A1 = 387 ms)].

Manipulation Check. To determine whether actors' reaches were influenced by attention control, we tested for subtle kinematic differences between conditions. Our hypothesis was that chosen reaches would express the decision required on those trials, with longer times to peak acceleration and curved trajectories reflecting the uncertainty of choosing vs. reacting (12). To test this hypothesis, we compared chosen and directed reaches on eight kinematic measures as shown in Table S1.

Four of the eight measures were consistent with the hypothesis, and none trended in the opposite direction. In comparison with directed reaches, chosen reaches had a marginally longer mean time to peak velocity, a higher mean vertical area under the curve, a larger mean ascending angle, and a longer mean ascending distance. These findings support the hypothesis that choosing to reach to a target location results in greater kinematic uncertainty than when being directed to the same target location.

In addition to these kinematic differences between conditions, each of the eight measures differed significantly between actors, as one might expect, given each actor's individual style of responding. However, with only one exception, these differences in individual actor style did not interact significantly with the reported main effects for chosen vs. direct reaches. The exception was that peak velocity was significantly higher for chosen than directed reaches for actor 1 [$t(49) = 3.15, P = 0.01$], whereas the other actors did not differ on this measure.

To help us understand whether the kinematic measures pointed to a common underlying factor, we submitted the eight kinematic measures, along with the temporal features for each reach in the stimuli set, to a principal component analysis (PCA). To further focus this analysis on only those kinematic effects that distinguished chosen from direct reaches, we performed the PCA after first computing z-scores for each measure. These z-scores were computed by dividing the difference between the measurement value and the mean of that measurement for the corresponding actor per the SD of that measurement for that actor. This computation meant that there were no longer any differences between actors in these measures or interactions between actor and condition.

Visual inspection of a scree plot, showing the total variance accounted for by the PCA as a function of an increasing number of potential components, revealed a plateau after the first component (21.98%). Measurement loadings on this component were generally positive for chosen reaches and negative for directed reaches, leading to a significant difference overall [$F(1,392) = 5.20, P = 0.02, \eta^2 = 0.01$]. Table S2 shows the weights associated with each measure. This pattern supports the hypothesis that chosen reaches reflect endogenous orienting, portraying a reaching pattern in which slower movements take longer to achieve peak velocity, have marked transitions from center to end-side, and display arched vertical trajectories. Conversely, exogenous orienting has a reactive nature, which is reflected by a relationship between faster reaches that tend to quickly achieve peak velocity and have straighter trajectories from home to target.

Together, these efforts to equate the temporal parameters of chosen and direct reaches, along with the kinematic analyses indicating greater decisional uncertainty for chosen reaches, meant that we had a stimulus set allowing for testing whether observers, blind to the condition under which the actors were reaching, were nonetheless sensitive to actors' attention control states. To make the test of this hypothesis even more stringent, we randomized the order of video clips, thereby eliminating trial-to-trial contingencies that might give observers additional clues to the actor's state.

Table S1. Means of eight kinematic measures taken on the distribution of reaches used as stimulus materials in the experiments

| Kinematic measurements | Chosen | ANOVA test | | |
		Direct	F-test	P	η^2
PV	2,256	2,200	3.471	0.063	0.001
TimePV	191.3	187.1	3.065	0.081	0.003
h-AUC	18,390	18,703	0.808	0.369	0.002
SCA	60.26	59.55	0.927	0.336	0.002
SCD	121.5	120.8	0.047	0.829	0.000
v-AUC	42,138	41,035	9.564	0.002	0.015
AA	62.57	62.34	3.396	0.066	0.007
AD	412.8	411.0	5.451	0.020	0.010

PV, peak velocity, the maximum velocity achieved during the reach; TimePV, time to peak velocity, time elapsed from movement initiation until peak velocity. h-AUC, horizontal area under the curve, area to the outer side of the horizontal trajectory. Thus, higher values of h-AUC correspond to larger inward curvatures in the trajectories, indexing more accentuated transitions from side-neutrality (center) to side-selection (left vs. right). SCA, side-commitment angle, the angle between the most inwards point (point of highest side-neutrality) and the most outwards point (point of highest side-commitment) in the reaching trajectory. Larger angles correspond to more exaggerated transitions from side-neutrality (center) to side-selection (left vs. right). SCD, side-commitment distance, the distance between the points of lowest and highest side-commitment. Thus, larger side-commitment distance values represent longer transitions from neutrality to side-selection. v-AUC, vertical area under the curve, the area underneath the vertical reach trajectory. Thus, the bigger the v-AUC value is the more vertically arched the trajectory is. AA, ascending angle, angle between the latest point before the vertical lift-off and the upmost vertical point in the reach. The bigger the angle, the steeper the vertical ascend. AD, ascending distance, distance from the lowest to the upmost point in the vertical trajectory.

Table S2. PCA weights

Temporal and kinematic measurements	Weights
IT	−0.108
MT	0.674
TT	0.397
PV	0.032
TimePV	−0.010
h-AUC	0.093
SCA	0.389
SCD	0.107
v-AUC	0.794
AA	0.745
AD	0.244

The first PCA component accounted for 21.98% of the kinematic variability in the stimuli set and successfully distinguished between chosen and directed reaches. Inspection of this component loadings shows positive weights (≥0.3) for movement time, total time, side-selection angle, vertical area under the curve (v-AUC), and ascending angle. No negative loadings were relevant (≤ −0.3). IT, initiation time; MT, movement time; TT, total time. Other abbreviations are as in Table S1.

Movie S1. Video clips of actor 1 reaching to left target in the direct condition.

Movie S2. Video clips of actor 1 reaching to right target in the direct condition.

Movie S3. Video clips of actor 1 reaching to left target in the choice condition.

Movie S4. Video clips of actor 1 reaching to right target in the choice condition.

Movie S9. Video clips of actor 3 reaching to left target in the direct condition.

Movie S9

Movie S10. Video clips of actor 3 reaching to right target in the direct condition.

Movie S10

Movie S11. Video clips of actor 3 reaching to left target in the choice condition.

Movie S11

Movie S12. Video clips of actor 3 reaching to right target in the choice condition.

Movie S12

Movie S13. Video clips of actor 4 reaching to left target in the direct condition.

Movie S14. Video clips of actor 4 reaching to right target in the direct condition.

Movie S15.   Video clips of actor 4 reaching to left target in the choice condition.

Movie S16.   Video clips of actor 4 reaching to right target in the choice condition.

Other Supporting Information Files

Dataset S1 (XLS)