The motor cortex is causally related to predictive eye movements during action observation

Claudia Elsner a,*, Alessandro D’Ausilio b, Gustaf Gredebäck a, Terje Falck-Ytter a,c, Luciano Fadiga b,d

a Department of Psychology, Uppsala University, Uppsala, Sweden
b Robotics, Brain and Cognitive Sciences Department, Italian Institute for Technology, Genova, Italy
c Center of Neurodevelopmental Disorders at Karolinska Institutet (KIND), Karolinska Institutet, Stockholm, Sweden
d DSBTA – Section of Human Physiology, University of Ferrara, Ferrara, Italy

ABSTRACT

We examined the hypothesis that predictive gaze during observation of other people’s actions depends on the activation of corresponding action plans in the observer. Using transcranial magnetic stimulation and eye-tracking technology we found that stimulation of the motor hand area, but not of the leg area, slowed gaze predictive behavior (compared to no TMS). This result shows that predictive eye movements to others’ action goals depend on a somatotopical recruitment of the observer’s motor system. The study provides direct support for the view that a direct matching process implemented in the mirror-neuron system plays a functional role for real-time goal prediction.

1. Introduction

When observing other people interacting with the environment, humans systematically fixate action goals ahead of time. Such predictive eye movements emerge early in human development and continue to play a crucial role for collaboration and competition throughout life (Falck-Ytter, Gredebäck, & von Hofsten, 2006; Flanagan & Johansson, 2003; Kanakogi & Itakura, 2011). To date, the mechanisms behind predictive goal-directed eye movements in action observation are unknown.

According to a motor-cognitive view, goal-directed eye movements reflect the activation of effector-specific action plans in the observer. On a general level, we know that others’ action perception is dependent on motor system activation (D’Ausilio et al., 2009; Pobric & Hamilton, 2006; Stadler et al., 2012; Urgesi, Calvomerino, Haggard, & Aglioti, 2007). More specifically, the idea that action plans generate predictions in real time receives support from behavioral data demonstrating similar patterns of eye movements during execution and observation (Flanagan & Johansson, 2003). In line with this finding, several studies have demonstrated that observers take advantage of their own motor abilities (Aglioti, Cesari, Romani, & Urgesi, 2008) and specific motor cues like hand pre-shaping (Ambrosini, Constantini, & Sinigaglia, 2011) to predict other people’s actions. In addition, it was shown that predictive gaze behavior is temporarily affected when observers’ hands are tied behind their backs while observing reaching-grasping but not touching actions, suggesting that effective action prediction also depends on the observer being in a position to perform the same actions or not (Ambrosini, Sinigaglia, & Constantini, 2012). Moreover, Constantini, Ambrosini, and Sinigaglia (2012a) found that object reachability influences observers’ predictive gaze performance, i.e. observers are more predictive when an observed agent is able to reach for a goal object than when an object is out of the agent’s reach. Besides, they demonstrated that compatibility between observed and executed prehension selectively impacts gaze behavior (Constantini, Ambrosini, & Sinigaglia, 2012b), showing that the readiness of the observers’ own motor representations effects proactive eye movements as well.

Despite plenty of studies indicating that action prediction is driven by a recruitment of the observer’s own corresponding action plans, available data are not able to rule out alternative hypotheses, e.g. that predictive eye movements are driven by inferential processes, implemented by non-motor brain areas (Eshuis, Coventry, & Vulchanova, 2009; Falck-Ytter, 2012).
The current study aimed to investigate the underlying mechanisms of predictive eye movements. We showed point-light displays of manual actions to human adults and applied TMS to either the hand or a control area (leg) of the motor cortex while measuring latencies of goal-directed gaze shifts. The motor-cognitive view makes two predictions: (i) TMS over the motor cortex hand area should alter latencies of goal-directed gaze shifts and (ii) effects of TMS should differ between hand area and control sites within the motor cortex. According to non-motor accounts (Eshuis et al., 2009; Southgate, Johnson, & Csibra, 2008), TMS over the motor cortex should not effect predictive gaze.

2. Material and methods

2.1. Subjects

The sample consisted of 21 healthy adults (M=25.8 years, SD=5.3, 10 males), 10 in the TMS-Hand condition (M=24.4 years, SD=3.89, four males) and 11 in the TMS-Leg condition (M=27.1 years, SD=6.15, six males). Three additional subjects had to be excluded from the analysis because of failure to fulfill the inclusion criteria. For instance, they did not follow the experimental instructions as they focused on counting the number of dots during stimulus presentation. All subjects (mainly Italian University students) gave their informed consent for the experimental procedures and received 25€ as compensation for their participation. None reported any history of neurological disease or psychiatric syndrome. All subjects were right-handed (Oldfield, 1971) and had normal vision. The study was approved by the local ethics committee and conducted in accordance with the 1964 Declaration of Helsinki.

2.2. Stimuli and apparatus

Gaze was measured with a Tobii T120 near infrared eye tracker (sampling rate 60 Hz, accuracy 0.5°, monitor size 17 in., Tobii, Sweden, Stockholm). The presented reaching-grasping action was shown as a point-light (PL) display in order to control the visual information presented to the observer and to minimize the involvement of other processes that also become activated when seeing a human hand. The PL action depicting motion and configuration information from 25 moving dots was recorded in a motion capture system (QUALISYS, Sweden, Gothenburg) and converted into a 2D movie in MATLAB (MATHWORKS, U.S.A., Natick). The PL hand contained 19 PL markers (94 mm) attached to the joints of the fingers, plus two PL markers attached on the wrist and four additional point-lights that were placed on the arm (two at the elbow, two at the upper arm). The PL action was integrated in a virtual environment containing a target object (toy dinosaur), a barrier (Elsner, Polck-Ytter, & Gredebäck, 2012) and 25 static dots in the background (in Cinema 4D, MAXON, Germany, Friedrichsdorf; see Fig. 1A). In accordance with D'Ausilio et al. (2009), the background dots (same size and color as the PL markers) were added to increase the difficulty of the task and to avoid ceiling effects. The final stimulus movie showed a PL hand tapping with the fingers and subsequently reaching for and acting on the target object. A gray barrier partly covered the target object, so that the final approach and interaction of the PL hand with the target object was not visible (Fig. 1A). After the PL representation of the hand disappeared behind the barrier, the target object was grasped, lifted, and replaced on the floor by the PL hand. Subjects were presented with the PL action from a lateral perspective showing either reaching from the left to the right or from the right to the left side. Between stimuli presentations, various attention-grabbing movies and several counting backwards tasks were presented as distractors to maintain subjects' attention. Since the examined behavior (gaze prediction) is subject to quick habituation (Elsner et al., 2012; Heinrichs, Elsner, & Gredebäck, 2012), we altered the stimulus movies (showing reaching from the left or from the right) and presented distractor tasks in-between to include more variation and to avoid early habituation.

2.3. Procedure

After written consent, subjects were seated approximately 60 cm in front of the eye tracker in a modified dentist chair to minimize body and head movements. A right tissue cap with a grid of points spaced 1 cm was centered on the vertex to aid coil positioning. The location for either the hand area or the leg area was marked on the cap. An articulated Manifrotto arm (Manifrotto, Italy) maintained the coil in a stable position and stabilized subjects' head throughout the experiment. After a gaze five-point-calibration (Gredebäck, Johnson, & von Hofsten, 2010), subjects were presented with 48 trials of a PL reaching-grasping action directed towards a target object (alternating PL reaching starting from the right or from the left), interleaved with different distraction tasks. Subjects were instructed to only passively observe the stimulus movies. During half of the PL stimulus presentations and in a randomized order, we applied a train of TMS biphasic pulses over either subject’s hand area or leg area of the left motor cortex (see Fig. 1B). Participants were randomly assigned to either the TMS-Hand or TMS-Control (Leg) condition. All subjects were also presented with 24 trials with no TMS, randomly mixed with the TMS ones. After the first block (24 PL stimulus presentations), subjects filled out a questionnaire and were asked to describe the presented PL stimuli and whether they recognized the PL display as a specific object or event. One subject that did not detect the hand was asked to fill out the questionnaire again after the second block.

2.4. TMS

TMS was delivered using a biphasic stimulator (Magstim, Whitland, UK) through a figure-of-eight coil (70 mm). The mapping procedure and threshold estimation was performed via wireless EMG recordings (Aurion, ZeroWire EMG) with a standard tendon-belly montage (Ag/AgCl electrodes). Signal was band-pass filtered at 10–1000 Hz and digitized at 2000 Hz. EMG traces were then acquired using a CED power1401 board and visualized with the software Signal 3.09 (Cambridge Electronic Design Limited, UK). The hand area was found by locating the First Dorsal Interosseus (FDI) muscle whereas the control site was set by locating the Anterior Tibialis (AT) muscle representation. Stimulus location and intensity were defined on a subject-by-subject basis. During the experiment, a short train of TMS pulses was delivered in half of the trials, at the beginning of the reaching trajectory when the maximal acceleration is located in time. The train consisted in three pulses spaced 100 ms between them. The onset of the TMS train was triggered by a photodiode placed in the upper left corner of the screen, detecting the appearance of a white square. The TMS sequence was decided to stimulate the motor cortex long enough to disrupt the underlying contribution of motor cortices to other’s action simulation. Short burst or trains of pulses are often used in online TMS protocols especially when strong temporal hypothesis cannot be formulated (for a review: Sandrini, Umiltà, & Rusconi, 2011; Siebner, Hartwigsen, Kassuba, & Rothwell, 2009).

Subjects in the TMS-Hand condition received TMS pulses at an intensity of 100% of the resting Motor Threshold (mean stimulator output: 51.5; SD: 6.68). Subjects in the TMS-Leg condition received TMS pulses at an intensity of 100% of the AT active Motor Threshold (aMT; mean stimulator output: 60.7; SD: 9.69). Resting Motor Threshold and active Motor Threshold were defined as the intensity of the stimulator output eliciting a Motor Evoked Potentials (MEP) of at least
50 mV in half of 10 consecutive stimulations for the resting Motor Threshold and a MEP of at least 200 mV in half of 10 consecutive stimulations for the active Motor Threshold (Rossini et al., 1994). This choice was motivated by the fact that the FDI resting Motor Threshold is always much smaller than the AT resting Motor Threshold. Therefore, stimulating the leg area at the resting Motor Threshold would have elicited Motor Evoked Potentials (MEP) also in the FDI muscle, introducing a critical spatial specificity issue. On the other hand, it has to be noted that stimulation intensity for the control site was always higher than for the hand area, thus excluding the possibility that control stimulation was not able to excite the neural tissue. In sum, we have direct evidence that we were on the right spot and intensity was tailored to the respective excitability of the hand and leg area at the individual level, but more crucially, spreading to the adjacent area was monitored and avoided. The onset and timing of TMS pulses during stimulus presentation were the same in the Hand and the Leg condition.

2.5. Data analyses

All data analyses were performed in MATLAB. Two areas of interest (AOIs) were used to analyze predictive eye movements. One AOI covered the initial position of the PL hand (plus one visual degree in each direction, Hand AOI). The second AOI covered the target object (plus one visual degree at the right and left side; 17.4 visual degrees vertical extension, Goal AOI). In order to exclude cases where subjects only fixated the right border of the barrier, the right side of the Goal AOI had a distance of one visual degree to the right border of the barrier (Fig. 1A). Within this predictive eye movement paradigm, we measured when subjects fixated the target object relative to the onset of end-effects (Elsner et al., 2012) during the reaching phase (the lifting of the target object; Fig. 1A). Data were included in the analysis when subjects fixated first the Hand AOI and subsequently the Goal AOI (each for more than 200 ms) in a time period between the onset of the reaching action and 1000 ms after the onset of end-effects. A gaze fixation in the Goal AOI before the goal was achieved was measured as a prediction. Latencies of goal-directed gaze shifts were calculated by subtracting the time when subjects first fixated the Goal AOI from the point in time when the PL hand started to lift the target object (onset of end-effects). Thus, positive numbers refer to predictive gaze arriving at the goal AOI before the onset of end-effects. Latencies of goal-directed gaze shifts were aggregated and averaged over all valid TMS and No-TMS trials respectively. An outlier analysis with z-transformations was performed on all data points for each condition. Two data points in each condition were classified as an outlier and excluded from the analyses.

Statistical analysis was performed in SPSS. To verify that subjects used predictive gaze, we run a priori one-sample t-tests on latency differences from all trials to assess if subject’s average gaze behavior was significantly predictive or reactive by comparing latencies of all valid goal-directed gaze shifts to 0 ms, expecting that all latencies should be significantly different from 0. Furthermore, an independent sample t-test compared mean latencies in No-TMS trials between the Hand and the Leg condition to test for baseline differences in subjects’ gaze behavior.

Due to the inherent inter-individual variability in gaze behavior, individual difference scores between mean latency in TMS trials and mean latency in No-TMS trials were calculated and used as dependent variable to evaluate only the TMS-induced modulation by removing baseline differences. In order to test our predictions, we conducted one-sample t-tests comparing difference scores with 0 to assess if gaze behavior differed between TMS and No-TMS trials within the Hand condition and within the Leg condition (two-tailed probabilities corrected for multiple comparisons with α =0.025). An independent samples t-test examined if difference scores between TMS and No-TMS trials differed between the two conditions.

3. Results

Ninety-five percent of all subjects detected the PL display as a hand (assessed by questionnaire, see Elsner et al., 2012) and all but one subjects showed predictive gaze during PL action observation (TMS trials: M = 645.0 ms, SE = 67.8; No-TMS trials: M = 603.3 ms, SE = 74.8). An independent-sample t-test revealed that there was no significant baseline difference in subjects’ gaze behavior between the Hand and the Leg condition in No-TMS trials, t(19) = 1.44, p = .166.

Subjects’ difference scores between TMS and No-TMS trials differed significantly from zero in the Hand condition, M_{TMS-\text{No TMS}} = −137.6 ms, SE = 50.3, t(9) = 2.73, p = .023 (Fig. 2). That is, TMS over the hand area significantly delayed latencies of goal-directed gaze-shifts (TMS trials: M = 575.8 ms, SE = 98.5) compared to trials without TMS (No-TMS trials: M = 713.4 ms, SE = 98.6). In the Leg condition, difference scores were also significantly different from zero, M_{TMS-\text{No TMS}} = 204.6 ms, SE = 48.0, t(10) = 4.26, p = .002. In this condition, we found the reverse effect, i.e. TMS trials: M = 707.8 ms, SE = 93.8, No-TMS trials: M = 503.2 ms, SE = 106.4. Difference scores between the TMS-Hand condition and the TMS-Leg condition differed significantly from each other, t(19) = 4.92, p < .001, d = −2.25, see Fig. 2.

4. Discussion

Other’s action prediction is a pervasive ability, which develops very early in life (Falck-Ytter et al., 2006; Kanakogi & Itakura, 2011). Gaze prediction has been proven to be a powerful tool to measure such capability in young children and adults (e.g. Ambrosi et al., 2011; Constantin et al., 2012; Gredeback & Melinder, 2010; Kochukhova & Gredeback, 2010). However, a critical open question regards the possible contribution of the motor system to these anticipatory processes.

One possibility is that action plans in our motor cortex are functionally related to predictive eye movements. This motor-cognitive view emphasizes the importance of motor-simulation during action observation (Aglioti et al., 2008; Fogassi et al., 2005; Iacoboni & Dapretto, 2006; Rizzolatti & Craighero, 2010). An alternative view suggests that the ability to predict the actions of others’ actions requires inferential processes and a high level of action understanding (Csibra & Gergely, 2007; Eshuis et al., 2009; Gergely & Csibra, 2003).

One of the possible solutions to this debate is the direct measure of brain activities associated to predictive eye movements. However, no neuroimaging or neurophysiological study has yet been performed to study this question. Here we used online trains of TMS pulses to disrupt activity in the hand primary motor cortex to test the effect in gaze prediction during the observation of hand actions. We demonstrate that TMS to this particular area of the primary motor cortex causes a delay in predictive gaze shifts relative to No-TMS trials. These findings support the motor-cognitive view and speak in favor of a functional connection between real-time goal prediction and activation of the observer’s own corresponding action plans. Generally, the mirror-like effect should in principle be generated by a network of areas including ventral premotor, inferior parietal lobule, superior temporal sulcus as well as motor and somatosensory areas (for a review see Iacoboni & Dapretto, 2006; Rizzolatti & Craighero, 2004). However, prior studies (e.g. Fatiga, Fogassi, Pavesi, & Rizzolatti, 1995) have demonstrated that measuring activity of the cortico-spinal tract (via TMS stimulation of the primary motor cortex) was a good proxy for the mirror-neuron system as a whole. More importantly, only the primary motor cortex enables the differential stimulation of two body effectors. This is due to the clear and large somatotopy present in M1 and the ease by which TMS can target different sections of it. Thus, based on the findings in the current study, we argue that stimulation of the motor cortex directly impacts a direct-matching process providing the ability to predict other people’s actions.

In the current study, we find a reversal effect in predictive gaze behavior when stimulating the leg primary motor representation. This may be due to several reasons. In fact, a reversal of effects is frequently reported (Kapur, 1996; Najib & Pascual-Leone, 2011), e.g. when stimulating competing cortical representations (Walsh, Ellison, Battelli, & Cowey, 1998) or might depend on local transient activity (Silvanto, Cattaneo, Battelli, & Pascual-Leone, 2008). Stimulation of the leg area might release the hand cortex from lateral inhibition, resulting in a net facilitation (Kobayashi & Pascual-Leone, 2003). Stimulation on the control site was stronger (due to the inherent lower excitability of the leg area) and thus
more likely to elicit sub-threshold spreading to the adjacent hand area. The exact nature of the reversal effect, although interesting, goes beyond the scopes of the present investigation.

It is important to consider whether our results could reflect a differential effect on an adjacent brain area, in particular the frontal eye fields (FEF) – known to be important for saccadic eye movements. It is unlikely that our results are linked to the FEF since the distance from any of the experimental sites to FEF is more than 4 cm, and at this scalp distance TMS effects drop to below 30% of the actual pulse strength (Siebner et al., 2009). From a functional point of view, and drawing from the original paper on gaze prediction (Flanagan & Johansson, 2003), the presence of strong eye-hand coordination programs allows humans to predict the goal of other people’s manual actions. However, the claim that these tight links enable action prediction when observing others’ actions was based on behavioral measures alone and subsequent studies (Eshuis et al., 2009; Southgate et al., 2008) suggested that other non-motor inferential strategies could account for the same effects. Here, by stimulating the hand area we demonstrate that proactive eye movements use motor strategies as originally suggested by Johansson & Flanagan’s seminal paper.

The underlying mechanisms of predictive eye movements have never been tested in a neurophysiological study and thus this is the first one describing the neural markers of predictive gaze behavior. The establishment of a causal connection between predictive eye movements and motor cortex activity, along with prior demonstrations of real-time action prediction across species, suggests that we are tapping a phylogenetic ancient mechanism at the core of social cognition. The fact that goal prediction is present in both infants and chimpanzees (Myowa-Yamakoshi, Scola, & Hirata, 2012) suggests that this mechanism could provide a common foundation for action understanding and bootstrap social cognition at large.

Additional information

The authors declare no competing financial interests. The contents are original and have not been published or submitted for potential publication elsewhere.

**Acknowledgments**

We would like to thank Laila Craighero for valuable discussions and Eleonora Bartoli and Laura Maffongelli for helping in collecting the data. This work was supported by grants from the Swedish Research Council (VR 2011-1528), Marie-Curie ITN RobotDoC (2010-2013), SIEMPRE EU project (ICT-FET project number 250026), POETICON++ EU project (STREP Project ICT-288382) and the University-ER Region Grant to L.F.

**References**


