Cognition 168 (2017) 65–75

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

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

ARTICLE INFO

Article history:
Received 26 July 2016
Revised 13 June 2017
Accepted 14 June 2017
Available online 23 June 2017

Keywords:
Action prediction
Action reading
Motor interaction
Competitive interaction
Biological motion

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

---

* Corresponding author at: 33 Kirkland Street, Room 744, Cambridge, MA 02138, United States.

E-mail address: mvaziri@fas.harvard.edu (M. Vaziri-Pashkam).

http://dx.doi.org/10.1016/j.cognition.2017.06.014
0010-0277/Published by Elsevier B.V.

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 slowed 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.