Predicting other people’s action goals with low-level motor information

Terje Falck-Ytter
Center of Neurodevelopmental Disorders at Karolinska Institutet, Astrid Lindgren Children’s Hospital, Stockholm; and Department of Psychology, Uppsala University, Uppsala, Sweden

Submitted 24 August 2011; accepted in final form 14 January 2012

Falck-Ytter T. Predicting other people’s action goals with low-level motor information. J Neurophysiol 107: 2923–2925, 2012. First published January 18, 2012; doi:10.1152/jn.00783.2011.—In support for the direct-matching hypothesis, Ambrosini et al. (2011) recently reported that goal-directed saccades during action observation were modulated by manipulations of basic motor information. This finding indicates that motor programs, activated by low-level visual descriptions of others’ actions, are involved in predicting other people’s action goals. Here, I put this result into a broader context, review alternative interpretations, and suggest strategies for future studies.

Proactive eye movements and the direct-matching hypothesis. When you reach out to grasp an object, your action plan includes specific directions for the oculomotor system. For reaching and many other manual actions, predictive eye movements are important for planning and control. In a seminal study by Flanagan and Johansson (2003), it was shown that adults spontaneously perform such eye movements also when the observed actions are performed by other people. The similarity in performance indicates that motor programs for actions guide the oculomotor system in both execution and observation. Supporting this interpretation, control experiments showed that seeing a human actor was important for predictive gaze behavior to occur. That is, when subjects observed apparently self-propelled objects that moved on the same trajectory and bringing about the same end effects as the human actions, goal-directed saccades were not found. Follow-up studies have replicated these findings in adults, demonstrated the same pattern of performance in preverbal infants (Rosander and von Hofsten 2011), and shown that even in situations where multiple potential targets are present and the observers do not know the intention of the agent in advance, gaze still arrives at the correct target ahead of the arrival of the actor’s hand (Rotman et al. 2006). The resemblance of eye movements in execution and observation is just one illustration of the influential idea in neuroscience that the motor system plays an active role in perception of others’ actions (Rizzolatti and Sinigaglia 2010). In the literature, this view is often labeled the “direct-matching hypothesis.”

The role of motor cues for prediction. Most previous studies of predictive eye movements in action observation have contrasted very broad categories of effectors such as a human hand vs. a “self-propelled” moving object (e.g., Flanagan and Johansson 2003). However, if prediction during action observation reflects that action plans are triggered in the observer, gaze performance should be affected by more subtle manipulations of motor information. In a recent study of human adults, Ambrosini et al. (2011) provided the important demonstration that eye movements in action observation are influenced by the configuration of the hand of the actor. They showed subjects multiple trials of reaching-to-grasp actions. In every trial, there were two potential target objects. The authors varied both the size of the to-be-grasped object (small vs. large) and the configuration of the reaching hand (no preshaping vs. correct preshaping). In the no-preshaping trials, the observed hand moved towards and stopped when in contact with one of the two objects, with a constant closed fist configuration. In the preshaping conditions, the hand configuration corresponded to the target object (whole hand grip for large object, precision grip for small object). Ambrosini et al. found that when the observers looked at a moving hand with a precision grip configuration, they tended to saccade towards the small object, and their gaze arrived there before the arrival of the hand. Conversely, when the observed hand approached the two targets with a power grip configuration, predictive performance towards the large object was found. The authors interpret this main finding, that the hand configuration is important for predictive eye movements, as support for the direct-matching hypothesis of predictive eye movements.

Interestingly, the authors also found that the subjects looked predictively towards the largest object when they watched ambiguous actions (no preshaping). This performance was attributed to goal saliency. Another selection bias has previously been observed based on the relative proximity of available targets (when ambiguous, observers tend to first look at the object closest to the reaching hand) (Rotman et al. 2006). Together, in the face of incomplete information from ongoing actions, observers are expected to move their gaze to salient and near targets, and to shift their gaze further only when they realize that the “default” target is incorrect. It is not unlikely that this early and apparently “premature” selection of targets reflects, in part, expectations based on past experience with actions. An alternative hypothesis put forward by Rotman et al. (2006) is that gazing at one potential target is the best vantage point to evaluate the actual target of the reach.

Regardless of their function, these predictive but “incorrect” eye movements found in the no-preshape condition suggest that a clearly functional hand configuration is not necessary to trigger the predictive system. It indicates that seeing a moving hand, regardless of its configuration, can induce a predictive stance in the observer, but that the likelihood of predictive performance (particularly to correct targets) is further increased if functional hand configurations are observed.

Direct matching vs. inference. The direct-matching hypothesis has been defined in different ways (for a critical review of the term, see Csibra 2007). According to Ambrosini et al.
(2011), it states that “observing actions performed by others elicits a motor activation in the brain of the observer similar to that which occurs when she plans her own actions.” Moreover, the authors assume that goal-directed saccades found during observation of manual actions reflect an activation of this observation-action execution matching circuit (Flanagan and Johansson 2003). This circuit is presumed to consist of two main regions: the inferior section of the precentral gyrus plus the posterior part of the inferior frontal gyrus, and the inferior parietal lobule (Rizzolatti and Sinaglia 2010). These areas receive input from the superior temporal sulcus (STS). According to an alternative account (hereafter, “the inference account”), predictive eye movements do not reflect direct matching, but are contingent on that the observer understands “the intention of an agent to achieve a goal” (Eshuis et al. 2009). According to this view, goal prediction reflects a nonmotor cognitive process. Nonmotor action processing could be based on midlevel visual analysis (in the STS; Csibra 2007) and/or higher-level “mentalizing” areas, including the temporoparietal junction, and the midline structures posterior cingulate and medial prefrontal cortex in addition to the STS (de Lange et al. 2008). As pointed out by Ambrosini et al., the two hypotheses are not necessarily mutually exclusive.

To the author’s knowledge, all studies into the nature of predictive eye movements in action observation have used a purely behavioral approach, without direct manipulation (or measurement) of the putative processes driving the behavior. This is a clear limitation, given that the alternative explanations are not easily disentangled on a purely behavioral basis. Of the available studies, most support the direct-matching hypothesis. These indicate that motor cues influence predictive eye movements during action observation (Ambrosini et al. 2011), that eye movements are similar during execution and observation (Flanagan and Johansson 2003), and that action experience probably is a prerequisite for predicting others’ actions (Gredebäck and Kochukhova 2010; Kanakogi and Itakura 2011). Cannon and Woodward (2008) demonstrated that sequential finger tapping, but not backward counting, inhibits predictive gaze during action observation. Supporting the direct-matching hypothesis, this study showed that eye movements to the goal can be slowed when the motor system is performing a concurrent operation and that effortful cognitive processes may have no effect on such eye movements. Thus, prediction of basic manual actions seems to be a rapid, low-level process that can run independently of higher-level cognitive processes.

Supporting the inference account, one study has showed that that prediction of others’ actions can occur with few motor cues (Eshuis et al. 2009). This result is problematic for the direct-matching hypothesis if one assumes that there is only one route to prediction of others’ actions. However, this assumption is unlikely to be correct. It is not unlikely that both lower level visuo-motor matching and higher-level cognitive processes can influence the oculomotor system during action observation, depending on the context (de Lange et al. 2008).

A limitation of the Ambrosini et al. (2011) study is that it does not present data that are incompatible with the inference account. According to an influential version of this this view, people infer the goal of an ongoing action by assessing what end state it would efficiently bring about given the particular constraints of the situation (e.g., Csibra 2007). Interpretation at this level could be accomplished by midlevel visual analysis (e.g., in the STS) and does not necessarily require effortful processing involving other mentalizing areas. Given that the biological constraints (hand configuration) were manipulated in the Ambrosini et al. study, the observed difference between conditions is not unexpected from the point of view of this nonmotor hypothesis (Csibra 2007). Ambrosini et al. acknowledge that “processing of others’ actions based on contextual information might help one to identify the target objects” but argue that this is unlikely given the type of action they showed to the subjects.

There are studies available indicating that inferential processing of actions and action prediction are dissociable phenomena. These studies include nonhuman self-propelled objects that move rationally given the constraints of the situation towards a salient end state. It is well documented that such self-propelled objects are perceived as goal directed, as predicted by the inference account (Csibra 2007). However, human observers typically do not predict the “goal” of such nonhuman moving objects with their gaze (Eshuis et al. 2009; Flanagan and Johansson 2003; see also Paulus et al. 2011). In other words, perceiving that nonhuman agents (self-propelled objects) are directed towards goals does not automatically lead humans to look at these goals in the same predictive fashion as they would if they looked at human actions. Prediction of goals is most typically seen when a human interacts with objects.

Conclusion. Most available evidence is in line with the view that motor areas play a functional role in predictive gaze performance during action observation. The experimental stimuli used by Ambrosini et al. (2011) allowed the authors to demonstrate an effect of motor cues both on timing and accuracy in a context where the correct target could not be identified based on previous knowledge, but must be selected based on “online” information. A correspondence in terms of configuration (e.g., small object, precision grip) is likely to be particularly informative when many potential goals are present and the observer has no prior knowledge about the intention of the actor. Indeed, this situation is frequently encountered in everyday life. The results are in line with the direct-matching hypothesis, but equally important, they add substantial new knowledge about the information humans capitalize on when perceiving an ecologically highly relevant stimulus: the manual actions of other people.

Remarkably, none of the studies evaluating the validity of the direct-matching hypothesis include any direct manipulation (or measurement) of brain activity despite the fact that it has direct and testable neuroanatomical implications. Well-controlled neurophysiological studies are likely to be a crucial next step to enhance our understanding of predictive gaze performance in action observation.

ACKNOWLEDGMENTS

I thank Dr. Ben Kenward and Prof. Gustaf Gredebäck for valuable comments on an earlier version of this article.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: T.F.-Y. drafted manuscript; T.F.-Y. edited and revised manuscript; T.F.-Y. approved final version of manuscript.
REFERENCES


