Title
Action Anticipation and Interference: A Test of Prospective Gaze

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN
1069-7977

Authors
Cannon, Erin N.
Woodward, Amanda L.

Publication Date
2008

Peer reviewed

Action Anticipation and Interference: A Test of Prospective Gaze

Erin N. Cannon (ecannon@umd.edu)
Department of Psychology, University of Maryland
College Park, MD 20742 USA

Amanda L. Woodward (awoodward@psyc.umd.edu)
Department of Psychology, University of Maryland
College Park, MD 20742 USA

Abstract

In the current study we investigate the proposal that one aspect of social perception, *action anticipation*, involves the recruitment of representations for self-produced action. An eye tracking paradigm was implemented to measure prospective gaze to a goal while performing either a motor or working memory task. Results indicate an effect of the motor task, suggesting the interference of a shared motor and action perception representation.

Keywords: Eye tracking; Action Perception; Motor representation; Working Memory.

Introduction

Findings from both non-human primates and adult humans suggest that some aspects of social perception recruit representational structure from systems that guide one's own actions. The discovery of "mirror neurons" in nonhuman primates (DiPelligrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese et al., 1996) and mirror-like responses in human brain imaging studies (e.g., Buccino et al., 2001) indicate that perceiving others' actions activates regions associated with action control. It has been hypothesized that these mirror systems contribute to action understanding and the perception of others' intentions (e.g., Decety & Grezes, 2006, Gallese & Goldman, 1998). In the current study we investigate the proposal that one aspect of social perception, *action anticipation*, involves the recruitment of representations for self-produced action.

The mirror neuron system and predictive gaze

Studies have shown that when people perform an action, such as moving an object from one position to another, their gaze will shift ahead towards the goal or endpoint before the object in hand arrives (Johansson *et al*., 2001, Land, Mennie & Rusted, 1999). Likewise, when observing someone else performing an action, adults also show this prospective gaze shift to the goal, just as when they perform the action themselves (Flanagan & Johansson, 2003). This predictive gaze behavior seems to be specifically elicited by the observation of an agent who produces a well-structured goal-directed action (Flanagan & Johansson, 2003; Falck-Ytter, Gredebäck, & von Hofsten, 2006).

To illustrate, Falck-Ytter et al. (2006) showed adults, 12-month-olds, and six-month-old infants trials of either a human agent moving three objects into a bucket, or self-propelled animate-looking objects flying into a bucket. The eye tracking device pinpointed participants' points-of-gaze while watching each trial type. Therefore, it could be determined whether subjects made predictive looks to the goal, or passively tracked the moving objects. Both adults and 12-month-olds made predictive looks toward the goal (the bucket) when the human agent moved the objects, but not when the objects were self-propelled. Thus, action anticipation only occurred when participants witnessed a goal-directed action. Age differences in the results suggested a link to participants' own motor abilities: Unlike older infants, six-month-olds tested in the human agent condition did not show any anticipatory looking, perhaps because such an action was not in their motor repertoire. The authors suggested six-month-olds' lack of anticipatory responses might reflect their lack of a motor representation for the action of putting balls into a container.

The fact that similar anticipatory gaze patterns occur for both self-produced and observed actions has been taken as evidence that these two responses rest on a shared system. Consistent with this claim, ERP research has found that when an observed action is predictable with respect to the endpoint or goal, motor activation occurs within the observer just prior to the onset of the action (Kilner et al., 2004). This finding implies that the anticipatory nature of one's own motor system might also influence the utilization of shared representations, allowing one to anticipate or predict others' motor behaviors. Although the similarity in patterns of anticipatory gaze for self- and other-produced actions is suggestive of shared representations (Flanagan & Johannsson, 2003), these findings do not specifically test whether anticipatory gaze recruits motor representations.

In the current study, we tested this hypothesis directly by assessing the effects of motor versus verbal concurrent tasks on participants' action anticipation. If visual action anticipation recruits motor processes, then we predict that a concurrent motor task (finger tapping) will interfere with action anticipation, but a working memory task (sub vocal rehearsal) will not interfere or will interfere less strongly with action anticipation. We used the paradigm developed by Falck-Ytter and colleagues (2006) as a measure of action anticipation.

Method

Participants. Forty-five adults participated in this study, 29 females and 16 males. Each was assigned to one of three conditions: *Observation*, *Finger-tapping*, or *Working Memory*. An additional 6 adults were tested but excluded from the analysis due to a failure to collect enough data points, *Observation* (N = 0), *Finger-tapping* (N = 3), and *Working Memory* (N = 3). See inclusion criteria in the Data Reduction and Analysis section below. Participants were recruited from the Psychology department's undergraduate subject pool and received extra credit in their Psychology courses in exchange for their participation.

Apparatus and Stimuli. Data were collected via corneal reflection using a Tobii 1750 (Tobii Technology). This specialized 17" monitor contains near-infrared lights and a camera mounted around the video-stimulus display. Head motion tolerance of the Tobii 1750 is 30 x 16 x 20 cm from a viewing distance of approximately 60 cm. The monitor was attached to a movable arm, easily adjusted to an optimum distance and angle to record each participant's eye movements. Two PC computers were networked with the Tobii 1750; one recorded the gaze data collected from both eyes at a rate of 50Hz; and a second PC ran the Clearview 2.5.1 software (Tobii Technology). This software program recorded the calibration data and integrated the gaze data with the images being viewed. It also allowed the researcher to define time windows of interest and regions within those windows, or areas of interest (AOIs) that were identified in the time-stamped output if a fixation fell within that region.

Participants viewed a video in which an actor was shown sitting at a table with a bucket to his right and three balls to his left. He moved his right hand to grasp the balls and moved them to the bucket, one at a time. The natural sound of the rubber ball hitting the metal bucket when it dropped in could be heard. The total duration of the movie was 12 s, and the time it took to move the three objects until they disappeared into the bucket was 1.08, 1.26, and 1.22 s, respectively.

Procedure. Participants sat approximately 60 cm from the Tobii monitor, in a small research room with black curtains covering the walls. Informed consent was obtained from all participants. First, each participant was given a 9-point calibration in which a blue dot appeared against a white background at nine different points on the screen. There were three between-subject conditions that each required a slightly different instruction. In the *Observation* condition, participants were told they would be observing several trials of a movie involving a person performing an action. They were encouraged to attend to the video and avoid thinking about potential distractions (such as what they are doing later, etc.) during the two-and-a-half minute presentation. In the *Finger-tapping* condition, participants were given the same instructions as above, but were told that after the second movie presentation, the experimenter would instruct them to tap their fingers in one of two sequences during the next trial. The finger tap sequence was meant to be simple, thus the cascade sequence used in previous working memory studies (e.g., Kane & Engle, 2000) was adopted for

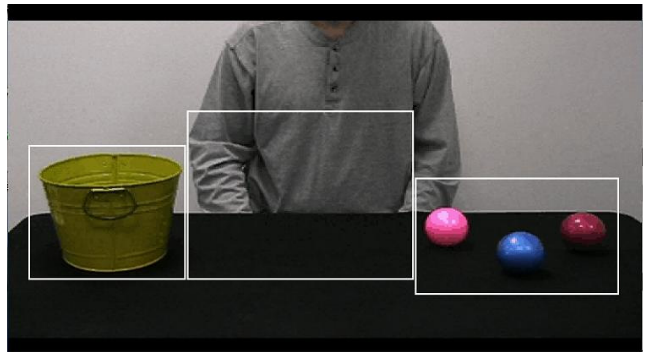

Figure 1. Define Areas of Interest (AOIs) within the scene. Start AOI encompassed the balls, the Trajectory AOI encompassed the region between the balls and the bucket, and the Goal AOI was the region surrounding the bucket.

this task. Participants rested their dominant hand on a clipboard placed on their lap. They were instructed to rest their thumb against the board and either tap from starting with their pinky (little finger) to the index finger, or starting with their index finger in sequence out to the pinky. Participants were instructed to continue repeating the sequence during the entire length of the movie. They were told to go at whatever pace they were comfortable with, but to try to keep a steady rhythm. Speed was not important but accuracy of tap order was important. Finger taps were video taped to record the accuracy of the tap sequences. Participants were given as much time as they wanted to practice the two different finger tap orders prior to starting movie presentations. Participants in the *Working Memory* condition were also given instructions similar to the *Observation* group, with the exception that starting on the third trial, the experimenter would give a sequence of four letters or numbers to repeat, then to sub-vocally rehearse it while watching the movie. At the end the experimenter would say "Go" and they were to repeat the rehearsed sequence. They were told we were interested in the accuracy between the pre- and post-movie responses. Participants were asked to make a conscious effort to sub-vocally rehearse even if they did not feel it was necessary for remembering the sequence. For each of these memory trials a unique sequence of the letters "R" "C" "M" and "L" were given, or the numbers "6" "9" "5" and "3." Number or letter sequences alternated over trials. There were a total of nine movie presentations. The first two presentations all participants watched, and thus served as the baseline trials. The following seven trials participants either watched, finger-tapped, or sub-vocally rehearsed based on the randomly assigned condition (*Observation*, *Finger-Tap* or *Working Memory*, respectively). Each movie was 12 seconds in length and followed by a four second attention-getter: a black screen with a toy moving at the center and making a sound. This was used to promote overall attention and to signal to the participant the start of the movie. In the WM condition, an additional blank screen appeared for 5s immediately following a movie presentation. This served as the response window for the participant to repeat the rehearsed sequence, then the experimenter to give a new sequence to be immediately repeated aloud and sub-vocally rehearsed during the proceeding movie. The combined length of all trials including attention getters was approximately 140 s in the *Observation* and *Finger-Tap* conditions, and 175 s in the *Working Memory* condition.

Predictions. Based on the work previously discussed, we first predicted that adults in the *Observation* condition would anticipate the arrival of the balls into the bucket, just as Falck-Ytter et al. (2006) found. If motor representations are recruited in action anticipation, then the finger tapping manipulation was predicted to disrupt anticipatory looking. The working memory condition provided an initial test of whether anticipatory looking is generally disrupted by a concurrent cognitive load, or rather selectively vulnerable to concurrent tasks that tax related motor systems (i.e., manual motor tasks). However, if neither of these tasks interfered with the anticipatory looking, then it would suggest this is a perceptual or automatic behavior.

Data Reduction and Analysis. Three areas of interest (AOIs) were defined in the movies, and are made visible in Figure 1. Participants were unaware of these regions of interest as they were not visible during viewing. The Start AOI covered the area where the target action began. The Trajectory AOI covered the area of space the person's hand moved through. Finally, the area encompassing the goal, the bucket, made up the Goal AOI. The dependent measure of interest was *Timing* of gaze arrival to the Goal AOI with respect to the ball's arrival to the Goal AOI. For each instance of the target action, data was collected at the moment the ball began to move at pick-up, until 1000ms after its disappearance into the bucket (the Goal AOI). During this time interval, gaze had to fall within the Start AOI followed by a look to the Goal AOI in order for a single data point to be included in the analysis. This inclusion criteria window was based on that used by Falck-Ytter et al. (2006). The window requirement of an initial look to the Start AOI ensured that anticipatory looks to the Goal were in fact anticipatory of the action in question and were not saccades launched to the goal prior to the action's initiation (Engel, Anderson, & Soechting, 1999).

The timing data was obtained by subtracting the time of the gaze arrival into the Goal AOI from the time of the ball's arrival into the Goal AOI. Thus, predictive looking times (in ms) resulted in positive numbers, and reactive looking times to the goal resulted in negative numbers. There were a total of 27 possible data points for each participant, i.e., three iterations to the bucket in each of the 9 trials. We adopted a conservative subject inclusion criteria of a minimum of 33% data points obtained in both the baseline and test trials (two data points in baseline and 7 at test). In the data presented here, the average number of data points used to obtain each participant's mean score for baseline trials was 5.13, 4.67, and 5.00 (out of 6) for *Observation*, *Finger-tap*, and *WM* conditions, respectively. For test trials, an average of 16.33, 15.27, and 13.73 (out of 21) data points were obtained in the *Observation*, *Finger-tap*, and *WM* conditions, respectively. The data points were averaged across events within the baseline and test trials, resulting in one aggregated baseline data point and one test trial data point per subject that were used in the analyses.

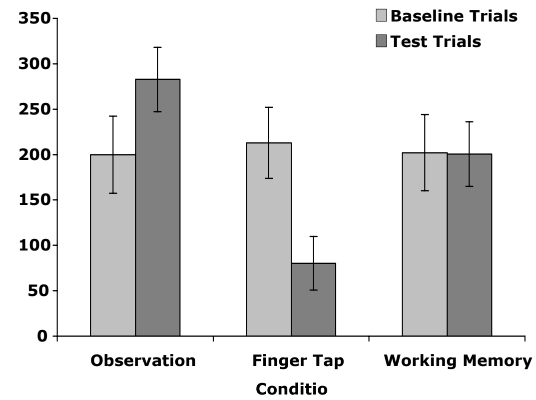

Figure 2. Mean Gaze Arrival to the Goal AOI relative to the Ball Arrival for all trials across conditions

Results

Accuracy of both the finger tapping and the working memory tasks was very high (each averaging over 95% accuracy), suggesting these were both easy tasks. Initial analyses did not find any effects of sex, so this variable was excluded from the subsequent analyses. Therefore, a Trial Type (baseline or test) x Condition (*Observation, Finger-tap, WM*) mixed design ANOVA was conducted on the timing of Gaze Arrival to the Goal AOI with respect to the ball's arrival. The means are displayed in Figure 2. There was a significant interaction of Trial Type x Condition $F(2, 42) = 6.17$, $p < .01$. Baseline trials remained stable across the conditions. However, the average gaze arrival to the Goal AOI on test trials declined in the *Finger-tap* condition. We then conducted a one-way ANOVA of Condition on Gaze Arrival to the Goal AOI on test trials only, and found the test trials significantly varied as a function of Condition $F(2, 44) = 9.14$, $p < .01$. Post hoc analyses (Bonferroni) revealed the test trials in the *Observation* condition were significantly different than those in the *Finger-tap* condition, $p < .001$, but not from the test trials in the *Working Memory* condition, $p = .28$. Critically, the gaze arrival timing in the *WM* condition was also significantly different than in the *Finger-tap* condition, $p < .05$. Both the *Observation* and *WM* conditions yielded faster gaze arrival timings to the Goal AOI than in the *Finger-tap* condition. A one-way ANOVA of Condition on Gaze Arrival during the baseline trials revealed no significant differences.

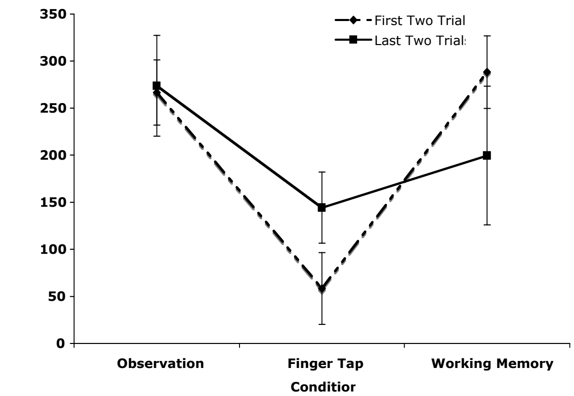

Figure 3. Mean Gaze Arrival to the Goal AOI relative to the Ball Arrival for the first and last two trials across conditions.

Single sample t-tests within each level of Trial Type and Condition on the timing of Gaze Arrival to the Goal AOI were tested against 0 ms to examine whether gaze behavior was significantly predictive or reactive. A 0 ms difference is considered a conservative measure given the typical reactive saccade latency in adults is 200 ms (Engel et al., 1999). Thus a gaze arrival at time 0ms would indicate the launch of the saccade to the goal occurred approximately 150-200 ms prior to arrival. Gaze arrival to the Goal AOI at all levels of trial types and across all conditions was found to be significantly predictive (greater than 0), $p < .001$ for the *Observation* and *WM* groups and the baseline trials of the *Finger-tap* group, and $p < .05$ for the test trials in the *Finger-tap* group. Thus, while it appears the act of finger-tapping did disrupt anticipation by significantly decreasing the timing, it did not obliterate prospective gaze entirely by resulting in reactive saccades. Participants in the *Finger-tap* group did anticipate to the Goal AOI, but their timing was significantly impaired in comparison to the other groups.

Lastly, we tested for learning across the trials. We were interested in whether increased familiarity to the actions (repeated trials) increased the velocity of the anticipation to the Goal AOI. Thus, a Test Pair (first two test trials v. last two test trials) x Condition ANOVA was conducted over average trial scores. There was effect of Condition $F(2, 34) = 5.81$, $p < .01$, and also a Test Pair x Condition interaction, $F(2, 34) = 3.62$, $p < .05$. Means are displayed in Figure 3. Follow-up one-way ANOVAs within each Condition indicated no effects of Test Pair in the *Observation* and the *Working Memory* conditions. However, there was a significant effect of Test Pair within the *Finger-tap* condition, $F(1, 14) = 4.68$, $p < .05$, suggesting some improvement of anticipatory gaze over trials. This finding reveals that the motor task might have been more taxing on the system initially, but over the course of 7 test trials (i.e., with practice) the task became more automatic, thus freeing up the shared motor resources and resulting in anticipation improvement over trials.

Conclusion

In this study we investigated the role of a divided attention task on anticipating the actions of another person. The results indicate that performing a manual motor task does interfere with the anticipation of another's actions, whereas a working memory task does not. This finding is the first to systematically link the role of one's own motor system to one's anticipatory gaze of another person's actions. These findings are consistent with the mirror neuron literature, and thus are supportive of their role in predictive gaze.

One caveat to these findings is the loss of usable data points in both the *Finger-Tap* and *Working Memory* conditions. The data loss was often due to inattention to the actions themselves (i.e., failure to gaze at the balls and bucket within the allotted time window, staring at the person or solely at the bucket). Based on participant feedback, it was often reported by those who were inattentive to the action that they spent most of their effort trying to concentrate on the diversion task. Studies of working memory have suggested strong individual differences in behavioral inhibition between groups with high and low working memory (WM) span (Rosen & Engle, 1998). Thus, a question for future work is to investigate these individual differences with respect to potential effects on prospective gaze.

Similarly, future work should also address if expertise at motor tasks involving finger articulation, such as playing an instrument, influences the degree of interference on prospective gaze. Our result of improvement in anticipation from the first two test trials to the final two trials was not predicted by assumptions of the mirror neuron system. Rather, we expected performance to remain stable over trials, as was the case in the Observation and WM conditions. Investigating the role of motor expertise will advance our knowledge about the systems involved in goal or action anticipation.

In conclusion, the mirror neuron system is a potential candidate for the neural system's modulation of anticipating others' actions, but more work is needed to explore explicitly goal anticipation as a possible function of this system. We take caution in making the claim that what we investigated here was 'goal' anticipation. Although others suggest the MNS is dedicated to the detection of goals and others' intentions (e.g., Gallese & Goldman, 1998), and Falck-Ytter et al. (2006) claimed 'goal' anticipation occurred in their paradigm, we feel more work needs to be done to test the selectivity of anticipation to 'goal' objects. We note that this paradigm conflated goals with regular patterns of movement. Thus, the effect may be about action anticipation rather than goal anticipation, per se. It will be important to conduct future studies in which this confound is eliminated.

Acknowledgments

We would like to thank all those who participated in this study. We would also like to thank Michael Dougherty, Claes von Hofsten, and Gustaf Gredebeck for their useful advice in implementing this paradigm. The work presented here was supported by grants to the second author (ALW) from the National Institutes of Health (HD35707) and the National Science Foundation (0634796).

References

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13*, 400-404.

Decety, J. & Grezes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research, 1079*, 4-14.

DiPelligrino, G., Fadiga, L. Fogassi, V. Gallese & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*, 176-180.

Engel, K. C., Anderson, J. H. & Soechting, J. F. (1999). Oculomotor tracking in two dimensions. Journal of *Neurophysiology, 81*, 1597-1602.

Falck-Ytter, T., Gredebäck, G. & von Hofsten (2006). Infants predict other peoples' action goals. *Nature Neuroscience, 9*, 878-879.

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

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Science, 8*, 396-403.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Science, 2*, 493-501.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by PET: 2. Observation compared with imagination. *Experimental Brain Research, 112*, 102-111.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping, 12*, 1-19.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *JEP: Learning, Memory, and Cognition, 26*, 336-358.

Kilner, J. M., Vargas, C., Duval, S., Blackemore, S-J. & Sirigu, A. (2004) Motor activation prior to observation of a predicted movement, *Nature Neuroscience, 7*, 1299-1301.

Land, M., Mennie, N. & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28*, 1311–1328

Rosen, V. A. & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39*, 418-436.