

Action plans used in action observation

J. Randall Flanagan¹ & Roland S. Johansson²

¹Department of Psychology and Centre for Neuroscience Studies, Queen's University, Kingston, Ontario K7L 3N6, Canada

²Section for Physiology, Department of Integrative Medical Biology, Umeå University, SE-90187 Umeå, Sweden

How do we understand the actions of others? According to the direct matching hypothesis, action understanding results from a mechanism that maps an observed action onto motor representations of that action^{1–4}. Although supported by neurophysiological^{1,5–13} and brain-imaging^{3,14–18} studies, direct evidence for this hypothesis is sparse. In visually guided actions, task-specific proactive eye movements are crucial for planning and control^{19–22}. Because the eyes are free to move when observing such actions, the direct matching hypothesis predicts that subjects should produce eye movements similar to those produced when they perform the tasks. If an observer analyses action through purely visual means, however, eye movements will be linked reactively to the observed action. Here we show that when subjects observe a block stacking task, the coordination between their gaze and the actor's hand is predictive, rather than reactive, and is highly similar to the gaze–hand coordination when they perform the task themselves. These results indicate that during action observation subjects implement eye motor programs directed by motor representations of manual actions and thus provide strong evidence for the direct matching hypothesis.

We tested the hypothesis that patterns of eye–hand coordination are similar when performing and observing a block stacking task. Previous studies of similar tasks have shown that there is a robust coupling between gaze and hand movements. Gaze leads the hand to blocks to be grasped and landing sites where blocks will be subsequently placed, and is rarely directed to the moving hand or block^{20,22–24}. These eye movements support hand-movement planning and control²² and may be viewed as part and parcel of the overall motor program for the task¹⁹.

The task involved moving three wooden blocks in the coronal plane. The blocks were all 2 cm in height and depth but had widths of 2, 3 and 4 cm. The locations of the blocks, as viewed by the subject, are shown in Fig. 1a. Before stacking, the blocks were aligned end to end at the right edge of a horizontal work surface and the task was to stack them, from the widest to the narrowest, at the left edge of the work surface. The same group of subjects both performed the task and observed an actor performing the task in counter-balanced order, although no order effects were observed in any of our analyses. The experimenter demonstrated the task once to each subject before data collection. In the action observation task, the actor sat across from the subject. Both the actor and the subjects used the tips of the right index finger and thumb to grasp the objects. In both conditions, the task was performed at a preferred rate and then at a faster rate.

Figure 1a shows all of the fixation points, from all subjects and trials, recorded in the action task performed at the preferred rate. The median path of the distal pad of the index finger is also shown. In agreement with previous reports, almost all fixations were directed towards sites of contact^{20,22,24}. Thus, the fixations were directed towards the grasp sites of the blocks to be picked up and the landing sites where the blocks were subsequently placed. The spatial distribution of fixation points in the corresponding action observation task (Fig. 1b) was very similar to that observed in the action task. In both tasks, gaze was predominantly directed to contact sites and subjects did not fixate the moving block or hand.

In Fig. 1c and d, the spatiotemporal coordination between gaze and hand movements in action and action observation, respectively, is examined by plotting the horizontal (x) gaze and hand positions as a function of time. Each red trace represents the x position of gaze during a single fixation (which can vary over time), and all fixations from all subjects and trials are shown. The blue and black curves represent the median x positions of gaze and of the distal pad of index finger, respectively. The median vertical (y) position of the index finger is shown in Fig. 1e and f. To preserve phase information when combining data from different trials, we first segmented each trial into phases and then normalized the time base of each phase to the median duration of that phase²².

In the action task (Fig. 1c, e), subjects fixated each forthcoming grasp and landing site well before the index finger arrived within the vicinity of the site. By contrast, gaze exited the grasp and landing sites at about the same time as the index finger exited these contact sites. This pattern of eye–hand coordination matches our previous work on object manipulation²². A similar pattern was observed in the action observation task (Fig. 1d, f). The sole exception was the first grasp site. In the action observation task, gaze exited the first grasp site slightly after the hand started to move away with the block. Although the lead of gaze over the hand appeared greater in the action task, in both tasks the gaze clearly anticipated forthcoming grasp and landing sites.

To quantify the coordination between eye and hand movements, we determined for each trial the times at which gaze and the index finger entered and exited each grasp and landing site zone. We defined the centre of each grasp zone as the position of the index finger when it reached a minimum in the vertical while grasping. The same method was used to define the centre of each landing site, except that we added a 1-cm downward offset to the measured minimum position of the index finger because the landing surface

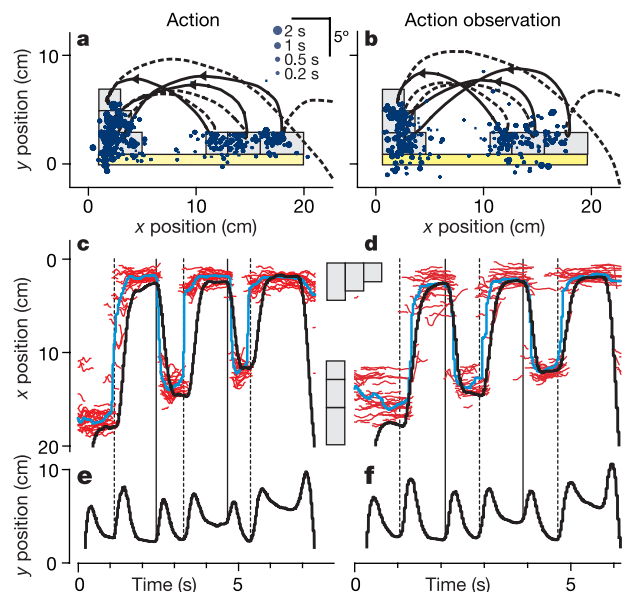


Figure 1 Gaze–hand coordination in action and action observation. **a, b**, Gaze positions at the end of periods between saccades (blue circles scaled to fixation duration), median hand path with and without a block in hand (unbroken and broken black lines, respectively), and approximate block positions before (right) and after (left) stacking. **c, d**, Median horizontal (x) positions of gaze (blue) and the index finger (black) as a function of time. Red traces represent the x position of gaze for all periods between saccades. Approximate x locations of the blocks are shown. **e, f**, Median vertical (y) position of the index finger. **c–f**, Time normalized to the median duration of all trials (Methods). Broken and unbroken vertical lines indicate the times at which the index finger exited grasp and landing sites, respectively.

was about 1 cm below the fingertip when placing blocks. Gaze and the index finger were deemed to be inside a given contact zone when they were within 2 cm (3°) and 0.5 cm (0.75°) of the centre position, respectively. The 3° radius for gaze was selected as it represents the size of the functional fovea in natural manipulation tasks²².

We analysed data from all contact sites with two exceptions. For gaze entries we excluded the first grasp site because gaze could arrive at this site well before the subject initiated the stacking task, and for gaze exits we excluded the last landing site because gaze could remain there until well after the task was completed. In both action and action observation, gaze arrived at a given contact zone before the hand, and the lead of gaze entry over hand entry increased with trial duration.

To assess the dependence of gaze entry times on trial duration, task and the interaction between duration and task, we carried out a stepwise linear regression analysis. The best regression model was one that included separate intercepts for the action and action observation tasks ($P < 0.001$), but had a common slope (-0.082 , $P < 0.001$) relating entry time to trial duration. On average, gaze arrived 150 ms earlier, relative to the index finger, in the action task. Unlike the gaze entry times, the gaze exit times did not vary with trial duration in either task, but stepwise linear regression analysis indicated separate intercepts for action and action observation ($P < 0.001$). On average, gaze departed the contact zone 72 ms before the index finger in the action task and 38 ms after the index finger in the action observation task. In both action and action observation, however, gaze shifts away from the contact zones were anticipatory or proactive. Had such gaze shifts been triggered by visual information related to hand movement away from the contact zone, we would expect gaze exits to lag behind hand exits by at least 100 ms (ref. 25).

To assess whether proactive eye movements require observing an interaction between the hand of the agent of the action and an object¹⁻⁴, we tested an additional group of subjects who observed the block stacking task being performed by the actor, no part of whom could be seen by the subjects (see Methods). Although subjects also

directed gaze to the block to be grasped and landing sites in this task (Fig. 2a), the pattern of eye–hand coordination was very different to that seen in the action and action observation tasks.

First, eye movements were linked reactively, rather than proactively, to the course of events. Gaze did not arrive at forthcoming block landing sites ahead of the block (Fig. 2b), and gaze exits from the grasp sites were markedly delayed (~ 200 ms) relative to the times at which blocks started to move (Fig. 2b and c, broken vertical lines). Second, fixations were distributed widely in the workspace (Fig. 2a); in fact, gaze frequently tracked the blocks when they moved to the stacking area. This is seen in Fig. 2b by the close match between the x position of the index finger when it moved to the landing site (unbroken black lines) and the x positions of gaze between saccades (red lines). Also note that the median x position of gaze (blue line) closely overlapped or lagged the x position of the index finger. Tracking was not observed when the actor's hand returned to pick up the next block, a movement that was invisible to the subject (Fig. 2b, c, broken segments of the black curves).

Figure 2d illustrates the tracking behaviour during block observation in the work plane by showing the gaze positions in the work plane between saccades. The longer red traces show periods of marked tracking, and the gaze position during these periods closely matched the path of the index finger (Fig. 2a). Figure 2e shows the corresponding plot for data combined from the action and action observation tasks, which were also performed at the preferred rate. Despite the fact that data from twice as many trials are shown, very few tracking periods can be observed.

To quantify the gaze tracking, we computed, for each subject, the total distance travelled by gaze (between saccades) that occurred during hand movements with the block, excluding periods when either the fingertip or gaze were within 2 cm (3°) of a grasp or landing site. We then expressed this as a percentage of the total distance travelled by the index finger with the block in hand and outside the contact zones. A between-subjects analysis of variance (ANOVA) indicated that the percentage of tracking was significantly greater ($F_{1,14} = 21.88$, $P < 0.001$) in the block observation task (mean \pm s.d., 22.2 ± 10.8) than in the action and action observation tasks in which subjects viewed the hand (3.96 ± 1.9 ; data combined from both tasks).

The central advance of this study is the demonstration that during action observation humans implement, instinctively, motor programs equivalent to those used in action. Specifically, subjects activate highly similar eye motor programs when performing and observing the same task. In addition, these programs operate almost in phase with the corresponding programs run by the actor, because the eye–hand coordination was essentially the same in the two tasks. Thus, as postulated by the direct matching hypothesis^{4,26}, a part of the motor system of the observer resonated with that of the actor.

The tight linkage, in time and space, between actions of the hand in manipulation and eye movements²² implies that the control program for a particular action includes directions for the oculomotor and visual processing systems¹⁹. Therefore, the task-specific eye movements that we observed during action observation are probably linked to parts of the neural processes that account for planning and control of the manual action. We can thus infer that such processes function in action observation, which supports the hypothesis that action understanding is based on a direct matching mechanism that maps the visual representation of the observed action onto a motor representation of the same action.

By contrast, our results provide little support for the alternative view that the observer captures the actor's behaviour through purely visual analysis of the elements that form the action, such as moving objects or a moving hand (reviewed in refs 4, 27, 28). First, rather than being reactively linked to events of the scenery, gaze specifically predicted forthcoming contact events in action observation. This proactive gaze behaviour is in line with the event-based predictive

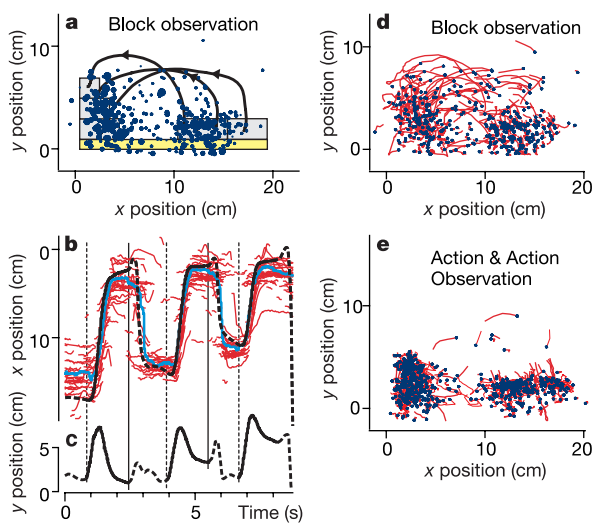


Figure 2 Gaze behaviour in the block observation task. **a–c**, Plots corresponding to those shown in Fig. 1; the unbroken black lines show the median path of the moving block and the broken segments in **b** and **c** mark periods when the actor's hand, invisible to the subjects, returned to pick up the next block. **d, e**, Gaze movements in the work plane during periods between saccades (red traces) for all subjects and trials in the block observation task (**d**) and for the action and action observation tasks combined (**e**). Blue circles indicate gaze positions at the start of each such period.

control strategies typical for manipulation²⁹. Second, in action observation, as in action, gaze was rarely directed towards the moving hand, as might be expected if hand motion were being visually analysed³⁰. Third, when subjects observed blocks moving without hand involvement, the gaze pattern differed from that engaged in action and action observation by being reactively coupled to the events. This result matches findings showing that the part of the motor representation to which the mirror system belongs is activated only when the observer views an object-oriented goal-directed task and not when the observer views its components^{1,2,18}. □

Methods

Subjects

Nine women and seven men aged from 19 to 30 years participated after providing written informed consent. All subjects were healthy, were right-handed, had normal vision and were naive as to the purpose of the experiment.

Block stacking task

The wooden blocks were located on a 19-cm wide work surface (Fig. 1a) centred 39 cm in front of the eyes and 8 cm below eye level. Eight subjects both performed the block stacking task and observed an actor performing the task who sat across from them. In both tasks, nine stacking trials performed at a preferred rate were followed by nine trials performed at a faster rate. In each block of nine trials, six were performed with an obstacle in the form of a vertically oriented block located in the centre of the workspace. These trials were interspersed among three trials without an obstacle. We focused our analysis on the latter trials. Each stacking trial was followed by an unstacking trial and between all trials subjects moved the hand to a parking position located 29 cm below and 4 cm to the right of the work surface.

Eight additional subjects observed the same block stacking task performed at the preferred rate by the same actor who was hidden from view. In this condition, the actor was positioned below the work surface and reached up to grasp tabs attached to the back surface of each block. The actor wore black gloves and a black outfit so as not to be visible against the black drape positioned behind the work surface in all conditions. Subjects were given no instructions on where to look in any of the three block stacking tasks.

Gaze position of the right eye in the plane in which the blocks were moved was recorded using an infrared eye tracking system and the position of the tip of the right index finger of the subject or actor were recorded with miniature electromagnetic sensors attached to the nail. A haptic calibration procedure was used to estimate the position on the distal pad of the index finger when in contact with a block. A bite bar was used to stabilize the head. The apparatus, calibration procedures and the accuracy and resolution of all measures have been described in detail elsewhere²².

Data processing

To preserve phase information when combining data from different trials for plotting, we first segmented each trial into seven contiguous phases based on when the index finger crossed a vertical line located at $x = 6$ (see Fig. 1a); the start and end of hand movement defined the beginning and end of the first and last phases, respectively. We then normalized the time base of each phase to the median duration of that phase. We defined trial duration, for use in statistical analysis, as the time from the end of the first phase to the start of the last. We detected the occurrences of saccades based on a filter applied to the gaze position signals combined with a threshold criterion as described²².

Received 6 May; accepted 13 June 2003; doi:10.1038/nature01861.

1. Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. Action recognition in the premotor cortex. *Brain* **119**, 593–609 (1996).
2. Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* **3**, 131–141 (1996).
3. Iacoboni, M. *et al.* Cortical mechanisms of human imitation. *Science* **286**, 2526–2528 (1999).
4. Rizzolatti, G., Fogassi, L. & Gallese, V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Rev. Neurosci.* **2**, 661–670 (2001).
5. di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180 (1992).
6. Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* **73**, 2608–2611 (1995).
7. Hari, R. *et al.* Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc. Natl Acad. Sci. USA* **95**, 15061–15065 (1998).
8. Nishitani, N. & Hari, R. Temporal dynamics of cortical representation for action. *Proc. Natl Acad. Sci. USA* **97**, 913–918 (2000).
9. Strafella, A. P. & Paus, T. Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *NeuroReport* **11**, 2289–2292 (2000).
10. Gangitano, M., Mottaghy, F. M. & Pascual-Leone, A. Phase-specific modulation of cortical motor output during movement observation. *NeuroReport* **12**, 1489–1492 (2001).
11. Umiltà, M. A. *et al.* I know what you are doing: a neurophysiological study. *Neuron* **31**, 155–165 (2001).
12. Avikainen, S., Forss, N. & Hari, R. Modulated activation of the human SI and SII cortices during observation of hand actions. *NeuroImage* **15**, 640–646 (2002).
13. Kohler, E. *et al.* Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846–848 (2002).
14. Grafton, S. T., Arbib, M. A., Fadiga, L. & Rizzolatti, G. Localization of grasp representations in humans

- by positron emission tomography. 2. Observation compared with imagination. *Exp. Brain Res.* **112**, 103–111 (1996).
15. Rizzolatti, G. *et al.* Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.* **111**, 246–252 (1996).
16. Decety, J. *et al.* Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain* **120**, 1763–1777 (1997).
17. Iacoboni, M. *et al.* Reafferent copies of imitated actions in the right superior temporal cortex. *Proc. Natl Acad. Sci. USA* **98**, 13995–13999 (2001).
18. Koski, L. *et al.* Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb. Cortex* **12**, 847–855 (2002).
19. Land, M. F. & Furneaux, S. The knowledge base of the oculomotor system. *Phil. Trans. R. Soc. Lond. Biol.* **352**, 1231–1239 (1997).
20. Land, M., Mennie, N. & Rusted, J. The roles of vision and eye movements in the control of activities of daily living. *Perception* **28**, 1311–1328 (1999).
21. Land, M. F. & McLeod, P. From eye movements to actions: how batsmen hit the ball. *Nature Neurosci.* **3**, 1340–1345 (2000).
22. Johansson, R. S., Westling, G., Bäckström, A. & Flanagan, J. R. Eye-hand coordination in object manipulation. *J. Neurosci.* **21**, 6917–6932 (2001).
23. Ballard, D. H., Hayhoe, M. M., Li, F. & Whitehead, S. D. Hand-eye coordination during sequential tasks. *Phil. Trans. R. Soc. Lond. Biol.* **337**, 331–338 (1992).
24. Ballard, D. H., Hayhoe, M. M. & Pelz, J. B. Memory representations in natural tasks. *J. Cogn. Neurosci.* **7**, 66–80 (1995).
25. Becker, W., (ed. Carpenter, R. H. S.) 95–137 (Macmillan, London, 1991).
26. Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. Resonance behaviors and mirror neurons. *Arch. Ital. Biol.* **137**, 85–100 (1999).
27. Perrett, D. I. *et al.* Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* **6**, 87–113 (1989).
28. Allison, T., Puce, A. & McCarthy, G. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* **4**, 267–278 (2000).
29. Johansson, R. S. & Cole, K. J. Sensory-motor coordination during grasping and manipulative actions. *Curr. Opin. Neurobiol.* **2**, 815–823 (1992).
30. Mataric, M. J. & Pomplun, M. Fixation behavior in observation and imitation of human movement. *Brain Res. Cogn. Brain Res.* **7**, 191–202 (1998).

Acknowledgements We thank G. Westling and A. Bäckström for engineering and software support. This study was supported by the Canadian Institutes of Health Research, the Swedish Medical Research Council, the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine, and the 5th Framework Program of European Union.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.F. (flanagan@psyc.queensu.ca).

.....
Prediction of auditory spatial acuity from neural images on the owl's auditory space map

Avinash D. S. Bala, Matthew. W. Spitzer & Terry T. Takahashi

Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403, USA

.....
 The owl can discriminate changes in the location of sound sources as small as 3° and can aim its head to within 2° of a source^{1,2}. A typical neuron in its midbrain space map has a spatial receptive field that spans 40°—a width that is many times the behavioural threshold³. Here we have quantitatively examined the relationship between neuronal activity and perceptual acuity in the auditory space map in the barn owl midbrain. By analysing changes in firing rate resulting from small changes of stimulus azimuth, we show that most neurons can reliably signal changes in source location that are smaller than the behavioural threshold. Each source is represented in the space map by a focus of activity in a population of neurons. Displacement of the source causes the pattern of activity in this population to change. We show that this change predicts the owl's ability to detect a change in source location.

We measured spatial discrimination behaviour by using the