PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

rstb.royalsocietypublishing.org

Research

Cite this article: Flanagan JR, Rotman G, Reichelt AF, Johansson RS. 2013 The role of observers' gaze behaviour when watching object manipulation tasks: predicting and evaluating the consequences of action. Phil Trans R Soc B 368: 20130063.
http://dx.doi.org/10.1098/rstb.2013.0063

One contribution of 17 to a Theme Issue 'Attentional selection in visual perception, memory and action'.

Subject Areas:
cognition, neuroscience

Keywords:
gaze behaviour, action–observation, object manipulation

Author for correspondence:
J. Randall Flanagan
e-mail: flanagan@queensu.ca

The role of observers' gaze behaviour when watching object manipulation tasks: predicting and evaluating the consequences of action

J. Randall Flanagan[1,2], Gerben Rotman[2], Andreas F. Reichelt[2] and Roland S. Johansson[3]

[1]Department of Psychology, and [2]Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada K7L 3N6
[3]Section for Physiology, Department of Integrative Medical Biology, Umeå University, 901 87 Umeå, Sweden

When watching an actor manipulate objects, observers, like the actor, naturally direct their gaze to each object as the hand approaches and typically maintain gaze on the object until the hand departs. Here, we probed the function of observers' eye movements, focusing on two possibilities: (i) that observers' gaze behaviour arises from processes involved in the prediction of the target object of the actor's reaching movement and (ii) that this gaze behaviour supports the evaluation of mechanical events that arise from interactions between the actor's hand and objects. Observers watched an actor reach for and lift one of two presented objects. The observers' task was either to predict the target object or judge its weight. Proactive gaze behaviour, similar to that seen in self-guided action–observation, was seen in the weight judgement task, which requires evaluating mechanical events associated with lifting, but not in the target prediction task. We submit that an important function of gaze behaviour in self-guided action observation is the evaluation of mechanical events associated with interactions between the hand and object. By comparing predicted and actual mechanical events, observers, like actors, can gain knowledge about the world, including information about objects they may subsequently act upon.

1. Introduction

The planning and control of manipulation tasks is centred on mechanical events that mark transitions between consecutive action phases and represent subgoals of the overall task. For example, when lifting, contact between the digits and object marks completion of the reach phase and the breaking of contact between the object and surface marks completion of the lift phase. By comparing predicted and actual sensory signals related to these events, actors can launch corrective actions if a mismatch occurs and confirm or update hypotheses about the world, such as hypotheses about the weights of objects [1,2]. The sensorimotor representations underlying manipulation tasks include specifications for task-specific eye movements and use of vision [3,4]. Thus, actors direct their gaze to the spatial targets of the unfolding task (e.g. an object to be lifted) and often maintain gaze at these locations until the subgoal is competed (e.g. the object lifts-off) such that events associated with action phase outcomes can be monitored in central vision [5,6].

When people observe manipulation tasks, their gaze behaviour is similar to that of the actor [6–9]. For example, when watching an actor pick up and replace a series of objects, observers direct their gaze to each object before the actor's hand arrives and keep gaze on the object until the hand departs for the next object [9]. This similarity appears to support the idea that observers implement, in real time, sensorimotor representations of the observed task [6,10–13]. However, the functional role of observers' eye movements remains

Royal Society Publishing — Informing the science of the future

unclear. The aim of this paper is to investigate the function of the gaze behaviour naturally generated by observers of manipulation tasks. We focus on two hypothesized functions. The first is that observers' gaze behaviour stems from processes involved in the prediction of the target of the actor's movement (i.e. which object the actor intends to manipulate), which is related to the movement selected by the actor [7,13,14]. The second is that observers direct their gaze to objects being manipulated so that they can evaluate sensory events related to mechanical interactions between the actor's hand and the object, events that are related to the completion, or outcome, of each phase of the actor's action [9,15].

Participants first completed a block of trials in which they simply watched an actor reach for and lift either of two candidate objects. After this condition with natural or self-guided action–observation, participants were given an explicit task to perform while watching the actor perform the same movements. One group completed the target prediction task in which they were asked to indicate, as early as possible during the unfolding action, which object the actor would grasp and lift. The other group completed the weight judgement task in which they were asked to judge the weight of the object, which involves evaluating the mechanical interaction between the actor's hand and object. If the primary function of observer's gaze behaviour in self-guided action–observation is prediction of the targets of the actor's movement, eye movements in the target prediction task should be similar to that seen in self-guided action–observation. Conversely, if the primary function is evaluation of the mechanical events associated with interactions between the actor's hand and objects, gaze behaviour in the weight judgement task should be similar to that seen in self-guided action–observation.

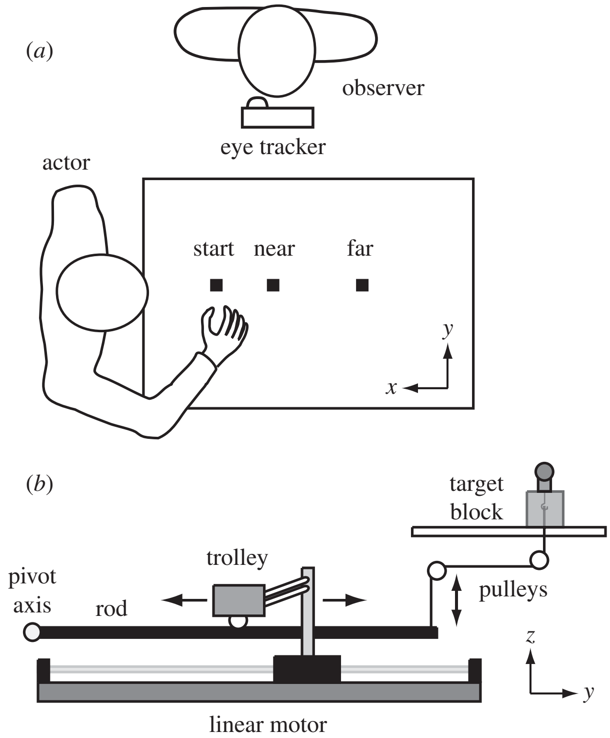

Figure 1. Experimental setting. (*a*) Observers viewed the actor from the side. In all trials, the actor first picked up and replaced the start block and then picked up and replaced either the near or far block before returning the hand to the start position shown. The black squares show the locations of the blocks. The position of the observer's gaze and the positions of the tips of the actor's index finger and thumb were recorded. (*b*) Device used to control the effective weights of the near and far blocks in the weight judgement task. A computer-controlled linear motor was used to position a pair of trolleys along parallel rotating rods attached, via a string running through pulleys, to hooks located in the centre of the blocks. A single rod, trolley and block are shown.

2. Material and methods

(a) Subjects

Sixteen participants took part after providing informed written consent. Seven women and three men participated in the target prediction experiment, and four women and two men participated in the weight judgement experiment. The participants, who received payment for their participation, were undergraduate and graduate students with normal vision. The local university ethics board approved the experiments, which complied with the Declaration of Helsinki.

(b) Equipment

Participants sat with their forehead resting against a fixed headband. A small bite bar further reduced head movements. An eye-tracking device (RK-726PCI, ISCAN Inc., Burlington, VT, USA), mounted on a frame below the headband, recorded the gaze position of the participant's right eye at 240 samples s^{-1} in a defined work plane that corresponded to a coronal plane located 100 cm in front of the eyes (figure 1*a*). The gaze calibration procedure has been described elsewhere [16]. Two small position/angle sensors (FASTRAK, Polhemus, Colchester, VT, USA) attached to the nails of the index finger and thumb recorded the position of the actor's hand at a sample rate of 120 samples s^{-1}.

In the target prediction task and its corresponding self-guided action–observation condition, three wooden blocks (cubes of height 2 cm), located in the work plane, were placed on a horizontal surface located in front of the actor. These included a start block, a near block and a far block whose centres were located 25, 40 and 50 cm, respectively, from the torso of the actor. In the weight judgement task and its corresponding self-guided action–observation condition, the start, near and far blocks were located 25, 40 and 60 cm from the torso. The start block was a solid cube (height 3.5 cm) composed of the opaque black polyoxymethylene plastic Delrin. The target blocks consisted of Delrin shells (hollow cubes 5 cm in height with an open bottom) with handles on top (figure 1*b*) instrumented with force–torque sensors (Nano 17 F/T, ATI Industrial Automation, Garner, NC, USA) that enabled us to measure the forces applied by the actor's hand when lifting. To set the effective weights of the near and far blocks, we used a linear motor system (figure 1*b*). (Note that we used slightly different block locations in the weight judgement experiment, which was carried out after the target prediction experiment, because we took advantage of pre-existing holes in the tabletop, through which the strings attached to the objects passed; figure 1*b*). Both the actor and observer wore headphones and listened to pink noise so that they could not hear the linear motor moving. The linear motor was moved in every trial even when the weight was not changed.

(c) Procedure

A single person, naive with respect to the research hypotheses, was the actor throughout. He was asked to perform the task at a comfortable speed and to be consistent. Before the start of a trial, the actor rested his right forearm on the table with his hand positioned on the right side next to the start block (figure 1*a*). At the start of each trial, the actor picked up and replaced (at the same location)

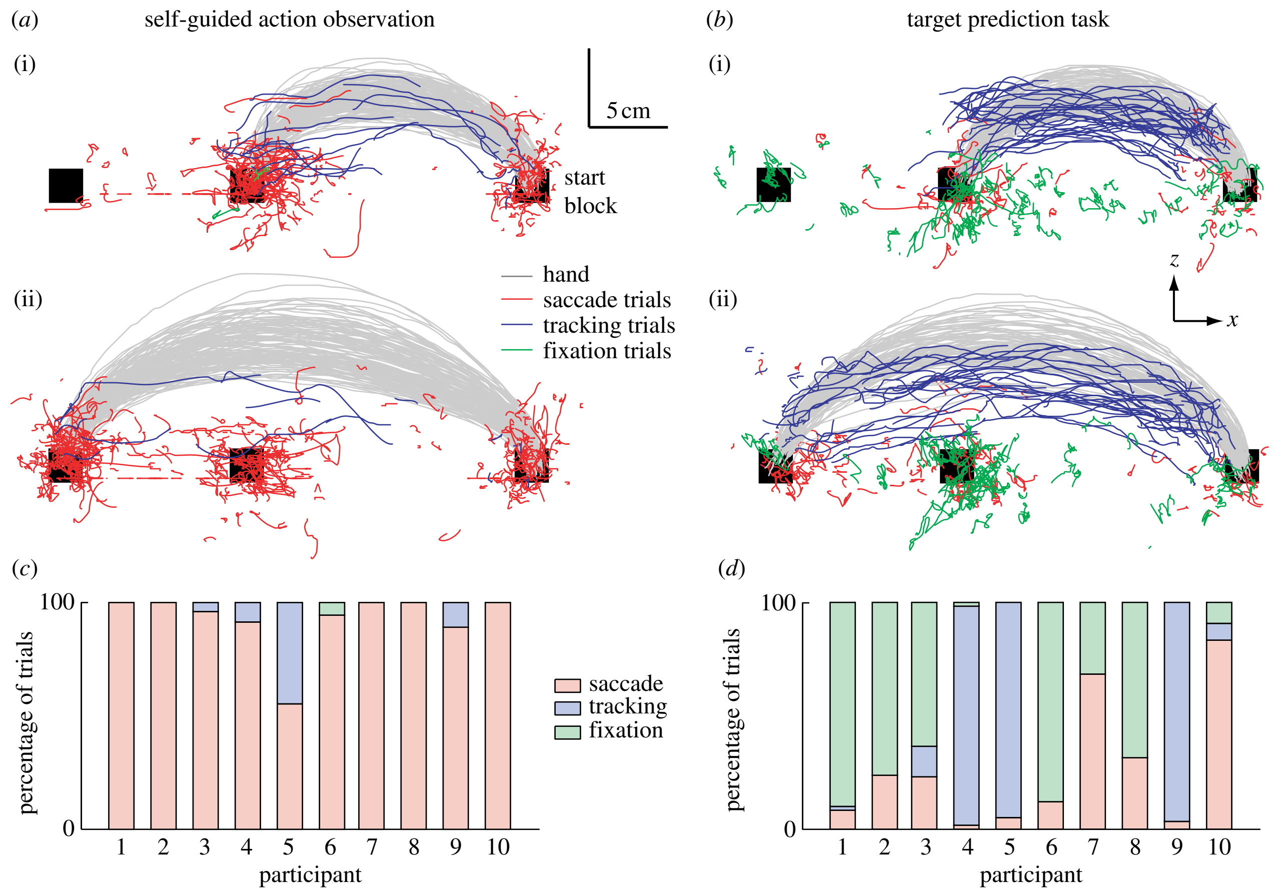

Figure 2. Gaze behaviour in the target prediction task and corresponding self-guided action–observation condition. (*a*,*b*) Gaze and hand paths during periods of fixation and tracking (i.e. excluding saccades) in the plane of movement as viewed by the observer. Data from all trials by all participants are superimposed. For each trial, data are shown from the time at which the start block was replaced, just prior to the onset of the hand movement away from the start block, to the offset of the hand movement to the target block. Separate plots are shown for hand movements from the start block to the (i) near and (ii) far targets. Gaze paths are coloured red, blue or green depending on whether they are from saccadic, tracking or fixation trials, respectively, and hand paths are coloured grey. (*c*,*d*) Percentages of saccadic, tracking and fixation trials per participant in the action–observation (*c*) and target prediction (*d*) tasks.

the start block and then picked up and replaced (again at the same location) either the near block or the far block before returning the hand to the rest position. The blocks were lifted about 10 cm above the table. The actor grasped the blocks (or handles attached to the blocks) from above and contacted the near and far sides (from his viewpoint) with the distal pads of the thumb and index finger, respectively. A visual cue, visible to the actor at the start of each trial but not to the observer, instructed the actor about which target block to lift. The onset of this cue also served as a go signal.

In self-guided action–observation trials, participants were asked just to watch the actor. In both the target prediction task and the weight judgement task, participants held, in each hand, a rod equipped with a button that could be pressed by the thumb. In the target prediction task, participants were instructed to press the right- or left-hand button, as quickly as possible, to indicate whether the actor was going to pick up the near or far block, respectively. Thus, there was spatial congruency between the response button and the location of the predicted target block. In the weight judgement task, participants were told that, in occasional trials, the target block (near or far) would be twice as heavy and that the actor would not know when these trials would occur. Participants were told to press the left- or right-hand button, as quickly as possible, to indicate whether the target block lifted by the actor was the heavy or standard weight, respectively. No instructions about where to look were given in any of the tasks.

For each participant, the target prediction task involved 30 trials, which were performed after 30 trials of self-guided action–observation. The weight judgement task involved 56 trials which were also performed after 30 trials of self-guided action–observation. In all conditions, the order of near- and far-target trials was randomized. In the weight judgement task, there were four unpredictable heavy weight trials (two for each target object) separated by 10–18 trials.

(d) Data analysis

Hand and gaze positions were smoothed using a fourth-order, low-pass Butterworth filter with cut-off frequencies of 14 and 25 Hz, respectively. We used the speed of hand and eye movements in the work plane to determine the onset and offset times of the actor's hand movements, between blocks, and saccades made by the observers. For hand movements and saccades, we used thresholds of 0.1 and 1.0 $m\,s^{-1}$, respectively. The average position of the thumb and index finger position sensors defined the position of the hand. Block lift-off was deemed to have occurred when the vertical velocity of the hand exceeded 0.1 $m\,s^{-1}$. In the target prediction task and its corresponding self-guided action–observation condition, we determined the time when the actor's hand contacted a block based on the rate of change of grasp aperture, the distance between the thumb and index finger positions. Contact was deemed to have occurred when the absolute rate of change dropped below 0.05 $m\,s^{-1}$. In the weight judgement task and its corresponding self-guided action–observation condition, the time when the actor's hand contacted the near and far blocks was determined from forces applied to the force sensors. Contact was deemed to have occurred when the force on either sensor exceeded 0.1 N. For the purposes of plotting gaze paths (figures 2 and 3), a fourth-order, low-pass Butterworth filter with a cut-off frequency of

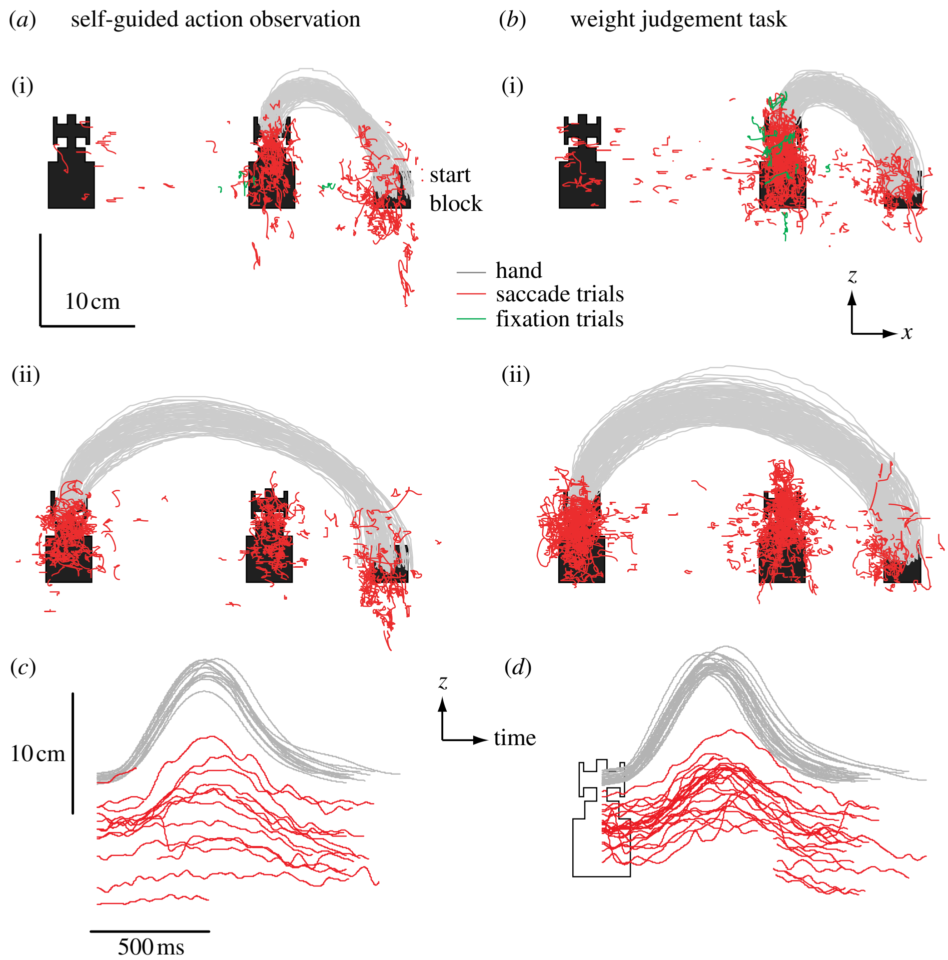

Figure 3. Gaze behaviour in the weight judgement task and corresponding self-guided action–observation condition. (*a*,*b*) Gaze and hand paths during periods of fixation and tracking (i.e. excluding saccades) in the plane of movement as viewed by the observer. Data from all trials by all participants are shown. For each trial, data are shown from the time at which the start block was replaced, just prior to the onset of the hand movement away from the start block, to the offset of the hand movement to the target block. Separate plots are shown for hand movements from the start block to the (i) near and (ii) far targets. Gaze paths are coloured red or green depending on whether they are from saccadic or fixation trials, respectively, and hand paths are coloured grey. (No tracking trials were observed.) Although the start block was fixated in almost all trials, some of the fixations do not appear in the figure because gaze shift away from the start block before the start block was replaced. (*c*,*d*) Vertical hand (grey) and gaze (red) positions, as a function of time, during the lift and replace of the target block in the weight judgement (*c*) and self-guided action observation (*d*) tasks. The continuous changes in vertical gaze position during fixation periods indicate that gaze tended to track the actor's lift and replace motion. Data from all saccadic trials in which the far block was the target, and weighed 3 N, are shown for a representative participant. Note that a larger number of such trials were collected in the weight judgement task than in the corresponding self-guided action–observation condition. For each trial, data are shown for the time period during which the actor's grip force exceeded 0.5 N, with an additional 50 ms before and after.

14 Hz was applied to the raw gaze position signals, as opposed to the 25 Hz cut-off frequency used for data analysis.

In a given trial, gaze behaviour was classified as saccadic if it consisted of saccades and fixations during the actor's reaching movement towards a target block. Gaze behaviour was classified as fixation if participants fixated a single location throughout the reaching movement. During such fixations, the precise gaze position could drift (see green traces in figure 2*b*, each of which shows the position of gaze during a single fixation trial) and very small saccades could occasionally occur. However, large saccades between blocks, as seen in saccadic trials, did not occur. Finally, gaze behaviour was classified as tracking if gaze tracking was observed for over 60% of the duration of the reach movement. Tracking trials could include catch-up saccades in addition to pursuit eye movement. Gaze was deemed to be tracking when gaze velocity exceeded 1.0 m s^{-1} in the absence of a saccade being detected. When gaze tracking was observed, gaze tracked the actor's moving hand (figure 2*b*). Overall, a very clear distinction was seen between saccadic and tracking trials. In tracking trials, the percentage of hand movement duration during which gaze tracked the hand was, on average, 91.7% (s.d. = 7.8%). By contrast, very few periods of tracking were seen in saccadic trials (figures 2*a*,*b* and 3*a*,*b*). Repeated measures ANOVA were used to compare various measures across conditions, and an α level of 0.05 was considered significant. Finally, gaze behaviour was only analysed in trials in which good gaze recordings were obtained. This included 92.5% of all trials.

3. Results

(a) Target prediction task

In the target prediction task and the corresponding self-guided action–observation condition, participants observed, from the side, an actor lifting wooden blocks located in the actor's mid-sagittal plane (figure 1*a*). In each trial, the actor first picked up and replaced the start block closest to him and then picked up and replaced one of the other blocks (near or far) before returning the hand to the rest position next to the start block. The actor was cued at the start of

each trial to pick up either the near or far block. The order of trials in which either the near or far block was lifted was randomized. After completing 30 trials involving self-guided action–observation without any specific instructions, the participants completed 30 trials of the target prediction task in which they were asked to indicate, as soon as possible, whether the actor would pick up the near or far block by pressing buttons held in the right or left hand, respectively. We asked for a quick response because we wanted to identify the gaze behaviour associated with predicting, as opposed to simply reporting, the actor's target, and because target prediction during natural action–observation occurs quickly [9].

The actor's performance was similar in the target prediction task and the corresponding self-guided action–observation condition, as assessed by the duration of the hand movement from the start block to the target block and the maximum vertical displacement of the hand during this movement. Hand movement duration was longer ($F_{1,9} = 325$, $p < 0.001$) for the far block (618 ± 19 ms; mean $\pm$ s.e.) than the near block (465 ± 10 ms), but that there was no effect of task and no interaction ($p > 0.05$ in both cases). Likewise, the maximum vertical hand displacement was greater ($F_{1,9} = 368$; $p < 0.001$) for the far block (7.1 ± 0.2 cm) than for the near block (4.7 ± 0.1 cm), but there was no effect of task and no interaction ($p > 0.05$ in both cases).

Depending on the observer's gaze behaviour, we classified each trial as saccadic, fixation or tracking, and there was little ambiguity in classifying trials (see §2). During self-guided action–observation, the vast majority of trials were saccadic (see figure 2*a*, which shows all of the data from all participants). A small number of tracking trials were observed and one of the nine participants produced almost all of these trials (figure 2*c*). The gaze behaviour in saccadic trials was similar to that observed in a previous study using a similar task [9]. Participants first directed their gaze to the start block ahead of the actor's hand approaching the block. Irrespective of whether the near or far block was the target, in the majority of the saccadic trials (91.7%), gaze then shifted to the near block shortly after the actor's hand started to move away from the start block after it had been replaced (93 ± 14 ms). When the near block was the actor's target, gaze remained at the near block until the hand arrived, whereas when the far block was the target, participants made a second saccade that shifted gaze to the far block ahead of the arrival of the actor's hand. The second saccade commenced when the actor's hand had moved 4.6 ± 0.9 cm past the near block in the horizontal (x) direction, which occurred 405 ± 13 ms after the onset of the actor's hand movement away from the start block. In the occasional saccadic trials in which gaze shifted from the start block to the far block, gaze remained there when the far block was the target but shifted to the near block, ahead of the hand, when the near block was the target. Overall, these results are consistent with previous results showing that when simply observing block manipulation tasks, people make predictive saccades to the objects involved and rarely track the moving hand [6,7,9,17].

By contrast, in the target prediction task, there were numerous fixation and tracking trials in addition to saccadic trials (see figure 2*b*, which shows data from all trials from all participants). Eight participants mainly produced fixation (five) or tracking (three) trials, whereas two participants mainly produced saccadic trials (figure 2*d*). The location of the single fixation in the fixation trials was generally close to one of the blocks and most often close to the near block. In tracking trials, on average, gaze tracked the hand for 91.7% (s.d. = 7.8%) of the duration of the hand movement. Across participants, the ratio of saccadic, tracking and fixation trials during the target prediction task (26 : 31 : 43) clearly differed ($\chi_2 = 93.0$; $p < 0.001$) from the corresponding ratio during self-guided action–observation (92 : 7 : 1).

Participants accurately predicted the target block in 90% of the trials in the target prediction task. This success rate did not depend on the location (near or far) of the target block ($F_{1,9} = 2.08$; $p = 0.18$). The time of correct button presses relative to the start of the actor's hand movement away from the start block was 422 ± 30 ms and did not depend on the location of the target block ($F_{1,9} = 0.72$; $p = 0.42$). When the near block was the target, correct button presses occurred when the actor's hand was 1.2 ± 0.8 cm ahead of the near block in the horizontal (x) direction. When the far block was the target, the actor's hand was 4.7 ± 1.3 cm past the near block in the horizontal. Neither prediction success rate ($\chi_2 = 0.04$; $p = 0.98$) nor the timing of button presses ($F_{2,562} = 0.58$; $p = 0.56$) depended on gaze behaviour as categorized in terms of saccadic, fixation and tracking trials (data were pooled across participants because not all participants exhibited all gaze behaviours).

(b) Weight judgement task

After completing 30 trials involving self-guided action–observation, participants completed 56 trials of the weight judgement task in which they were asked to indicate, as soon as possible, whether the target block lifted by the actor was the standard weight or the heavy weight, by pressing buttons held in the right or left hand, respectively. Both the actor and observer wore soundproof earphones and could neither see nor hear the linear motor system. Handles instrumented with force sensors were mounted on top of the near and far blocks and measured the forces applied by the actor's thumb and index finger (figures 1*b* and 3). In the weight judgement task, the target block weighed 3 N in standard weight trials, and in occasional heavy weight trials (one in 14 randomly selected trials), the weight was doubled to 6 N. In the corresponding self-guided action–observation condition, the weight was set to 3 N throughout to ensure that this condition represented natural action–observation, devoid of possible influences of unexpected events.

The actor's hand movements, as assessed by the duration of the hand movement from the start block to the target block and the maximum vertical displacement of the hand during this movement, were similar in the weight judgement task and the corresponding self-guided action–observation condition. Hand movement duration was longer ($F_{1,5} = 2178$, $p < 0.001$) for the far block (578 ± 7 ms) than for the near block (363 ± 4 ms), but there was no effect of task and no interaction ($p > 0.05$ in both cases). Likewise, maximum vertical hand displacement was greater ($F_{1,5} = 713$; $p < 0.001$) for the far block (13.8 ± 0.4 cm) than for the near block (8.9 ± 0.2 cm), but there was no effect of task and no interaction ($p > 0.05$ in both cases). We also compared the actor's lift and replace movements of the target object in the 3 N trials of the weight judgement task and the corresponding self-guided action–observation condition. The maximum height of the lift was slightly greater ($F_{1,5} = 38.9$; $p = 0.002$) for the far block (9.07 ± 0.21 cm) than for the near block (8.63 ± 0.22 cm), but

there was no effect of task and no interaction ($p > 0.05$ in both cases). The duration that the actor contacted the target block, defined as the time period over which grip force exceeded 0.5 N, was on average 695 ± 16 ms. ANOVA failed to reveal an effect of target block or task, and there was no interaction ($p > 0.05$ in all three cases).

In both the weight judgement task (figure 3*b*) and the corresponding self-guided action–observation condition (figure 3*a*), gaze behaviour was very similar to that previously seen during natural or self-guided action–observation [6,9] (see also figure 2*a*). (Note that figure 3*a*,*b* show data from all trials from all participants.) That is, the vast majority of trials in both tasks were saccade trials. No tracking trials were observed, and less than 2.5% of the trials were fixation trials in both tasks. Gaze shifted from the start block to the near block, ahead of the hand, in 87% and 90% of the trials during self-guided action–observation and the weight judgement task, respectively. When the far block was the target, gaze then inevitably shifted to the far block, again arriving before the hand.

During the weight judgement task, gaze shifted away from the start block (to either the near or far block) earlier than during self-guided action–observation ($F_{1,5} = 24.24$; $p = 0.004$). Specifically, gaze shifted away from the start block 621 ± 116 ms and 111 ± 101 ms before the hand started to move away from the start block in the weight judgement task and during self-guided action–observation, respectively. There was no effect of target block ($F_{1,5} = 5.23$; $p = 0.071$) and no interaction between task and target block ($F_{1,5} = 0.006$; $p = 0.943$) on the time of this gaze shift. Although gaze shifted from the start block earlier in the weight judgement task than during self-guided action–observation, in the weight judgement task, participants still fixated the start block until well after the actor had lifted it off the support surface.

On average, observers pressed the button 385 ± 48 ms after the actor first contacted the target block and this time did not differ significantly between the near and the far blocks ($p > 0.05$). Thus, the button press almost always occurred during the lift and replace movement and well before the actor released the block. Very consistent gaze behaviour was observed during the lift and replace movement in both the weight judgement task and the corresponding self-guided action–observation condition. Specifically, gaze continued to be directed towards the location of the target block for the duration of the lift and replace movement, and there was a tendency for gaze to move up and down with the vertical motion of the target block (figure 3*c*,*d*). However, the amplitude of the gaze movement during the lift and replace movement was typically smaller than the amplitude of the block movement. Small saccades were also observed during the lift and replace movement, occurring in about 30% of all trials in both the self-guided action–observation condition and the weight judgement task (see interrupted curves in figure 3*c*,*d*).

As expected, the period from when the actor contacted the target block to the instance it lifted off the support surface was substantially longer ($F_{1,5} = 210$; $p < 0.001$) in the heavy weight trials (596 ± 37 ms) compared with the standard weight trials (228 ± 12 ms) [18–20]. In addition, the maximum height of the lift was smaller ($F_{1,5} = 70.2$; $p < 0.001$) in heavy weight trials (6.82 ± 0.36 cm) than in standard weight trials (9.06 ± 0.24 cm). Consistent with previous findings [21–24], all participants were able to use visual information related to the lift, such as the time between contact and lift, lift height and speed and even hand shape, to judge object weight. Specifically, participants correctly judged the block as being heavy in all heavy weight trials and correctly judged the block as being the standard weight in 76 ± 8% of the standard weight trials. Most of the trials in which participants incorrectly judged the standard weight to be the heavy weight were performed before the occurrence of the first heavy weight trial. In an additional experiment not reported here, a new set of participants performed the weight judgement task where the ratio of light (i.e. standard) and heavy weight trials was 1 : 1. The gaze behaviour seen in this weight judgement 1 : 1 task was strikingly similar to that seen in the weight judgement task reported above. Moreover, participants correctly judged the weight of the target block in over 90% of the trials.

4. Discussion

Prediction of the consequences of action is a fundamental component of sensorimotor control [25,26]. In manipulation tasks, the prediction of sensory events associated with mechanical events marking the completion of action phases, such as the digits contacting an object and object lift-off during lift tasks, are especially important for dexterous control. By comparing predicted and actual sensory events, including tactile, visual and auditory events, actors can monitor task progression, launch appropriate corrective actions if a mismatch occurs and confirm or update hypotheses related to the task, including hypotheses about the weights of objects being manipulated [1,2]. For example, when lifting objects, people increase vertical load force to a target level corresponding to the expected weight and generate a prediction of when they will receive sensory feedback signifying lift-off. If the object is heavier than expected, then the object does not lift off at the expected time and the resulting mismatch between predicted and actual sensory events triggers a corrective action that involves probing increases in load force [18]. This mismatch also updates expectations about object weight such that in a subsequent lift, the increase in load force is more accurately targeted to the correct weight [18–20].

During object manipulation tasks, actors typically direct their gaze to the spatial targets of action phases, such as the object to be grasped and lifted or the location where an object will be placed [3,6,16,27]. This gaze behaviour helps guide the hand to the target [28–30] and also permits monitoring, in central vision, of mechanical events that represent goals of action phases [4,5,31,32]. We suggest that the similar eye movements produced by observers of manipulation tasks [6–9,17,33,34], also serve to monitor, in central vision, mechanical events that mark the completion of action phases, and that this involves neural processes resembling those used for prediction and evaluation of these events during action. When participants were explicitly asked to judge the weight of the target block, which involves comparing predicted and actual sensory information related to the lift, they directed their gaze proactively to the blocks as in self-guided action–observation. By contrast, when explicitly asked to predict which block the actor will lift, most participants fixated a single location or tracked the actor's hand. This indicates that target prediction does not require generation of proactive eye movements. Of course, because target prediction is required for proactive eye movements, such prediction is part of self-guided

action–observation. However, the converse is not true; our findings indicate that observers do not produce this gaze behaviour in order to predict the target.

Our results indicate that in the target prediction task, observers do not implement sensorimotor representations that fully simulate the task performed by the actor. Had observers implemented such representations in the target prediction task, proactive eye movements, specified by these representations [3,4], would be expected. Instead, our findings appear to be more consistent with the idea that, in the target prediction task, observers relied on inferential processes not involving complete action simulation [35].

The finding that participants, in both the weight judgement task and self-guided action–observation, directed their gaze to each block handled by the actor, including the start block, suggests that action–observation involves an urge to monitor all available behaviour of others. By continuously predicting and evaluating mechanical events associated with the completion of action phases observers can calibrate the performance of the actor. This, in turn, upholds the quality of the knowledge observers can gain about the state of the world, including information about properties of objects in the environment. For example, by monitoring how quickly and how high an actor lifts objects in general, the observer can more reliably detect when a particular object is heavier or lighter than expected by the actor.

It has been proposed that actors generate sensory predictions based on an efferent copy of the motor commands that is passed through an internal forward model of the controlled system, which includes the body and objects in the environment with which the body interacts [25]. A similar process could underlie the generation of certain sensory predictions in observers. That is, the observer could simulate covertly the motor commands of the observed action and pass an efference copy of these commands through a forward model of the controlled system [36–38]. This idea is broadly consistent with the results of numerous studies showing that sensorimotor areas, activated when performing action tasks, are also recruited when observing the same tasks [11,12,14]. For example, observing a video of an actor lifting an object increases the excitability of the representation of muscles involved in lifting in motor cortex, and this increases scales with the required force (i.e. object weight) [24,39].

Recently, Cannon & Woodward [40] reported that requiring participants to perform a sequential finger tapping task while observing an actor perform an object manipulation task appears to disrupt the predictive gaze behaviour normally observed [7]. This raises the question whether, in the target prediction task, having to prepare and execute a button press during the actor's reaching movement might have interfered with participants' gaze behaviour, leading to gaze tracking as opposed to proactive saccades. Although we cannot rule out this possibility, we would emphasize that participants were also required to perform a button press in the weight judgement task and yet exhibited proactive saccades. Moreover, the gaze tracking behaviour seen in the target prediction task was initiated well before participants pressed the button indicating which block was the target.

In this study, we considered two possible functions of observers' eye movements when watching manipulation tasks: predicting the target objects of the actor's action and evaluating mechanical events associated with interactions between the actor's hand and objects. A third possible function, that we deliberately did not address, relates to the ability to understand or infer the higher level goals or intentions of others through observation of their actions [13,14]. Indeed, we designed the experiments to minimize this component and note that there are many natural situations in which observers watch an action while already knowing the actor's high-level goals (e.g. when watching a dealer deal cards). However, we suggest that mechanisms supporting evaluation of mechanical events linked to the completion of action phases and those supporting prediction of higher level goals interact dynamically. That is, evaluating the mechanical events of sequentially generated action phases presumably facilitates understanding the actor's higher level goals. At the same time, inferences about the actor's intentions presumably shape observers' predictions about mechanical events.

Acknowledgements. We thank Martin York and Sean Hickman for technical support and Justin Caldwell who served as the actor in the study.

Funding statement. This work was supported by the Canadian Institutes of Health Research, the Human Frontier Science Programme, the Swedish Research Council (project 08667) and the Strategic Research Programme in Neuroscience at the Karolinska Institute.

References

1. Johansson RS, Flanagan JR. 2009 Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* **10**, 345–359. (doi:10.1038/nrn2621)
2. Flanagan JR, Bowman MC, Johansson RS. 2006 Control strategies in object manipulation tasks. *Curr. Opin. Neurobiol.* **16**, 650–659. (doi:10.1016/j.conb.2006.10.005)
3. Land MF, Furneaux S. 1997 The knowledge base of the oculomotor system. *Phil. Trans. R. Soc. Lond. B* **352**, 1231–1239. (doi:10.1098/rstb.1997.0105)
4. Hayhoe M, Ballard D. 2005 Eye movements in natural behavior. *Trends Cogn. Sci.* **9**, 188–194. (doi:10.1016/j.tics.2005.02.009)
5. Bowman MC, Johansson RS, Flanagan JR. 2009 Eye–hand coordination in a sequential target contact task. *Exp. Brain Res.* **195**, 273–283. (doi:10.1007/s00221-009-1781-x)
6. Flanagan JR, Johansson RS. 2003 Action plans used in action observation. *Nature* **424**, 769–771. (doi:10.1038/nature01861)
7. Falck-Ytter T, Gredeback G, von Hofsten C. 2006 Infants predict other people's action goals. *Nat. Neurosci.* **9**, 878–879. (doi:10.1038/nn1729)
8. Kanakogi Y, Itakura S. 2011 Developmental correspondence between action prediction and motor ability in early infancy. *Nat. Commun.* **2**, 341. (doi:10.1038/ncomms1342)
9. Rotman G, Troje NF, Johansson RS, Flanagan JR. 2006 Eye movements when observing predictable and unpredictable actions. *J. Neurophysiol.* **96**, 1358–1369. (doi:10.1152/jn.00227.2006)
10. Cisek P, Kalaska JF. 2004 Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* **431**, 993–996. (doi:10.1038/nature03005)
11. Rizzolatti G, Fogassi L, Gallese V. 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670. (doi:10.1038/35090060)
12. Rizzolatti G, Craighero L. 2004 The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)

13. Kilner JM, Frith CD. 2008 Action observation: inferring intentions without mirror neurons. *Curr. Biol.* **18**, R32–R33. (doi:10.1016/j.cub.2007.11.008)
14. Rizzolatti G, Fabbri-Destro M. 2008 The mirror system and its role in social cognition. *Curr. Opin. Neurobiol.* **18**, 179–184. (doi:10.1016/j.conb.2008.08.001)
15. Burke CJ, Tobler PN, Baddeley M, Schultz W. 2010 Neural mechanisms of observational learning. *Proc. Natl Acad. Sci. USA* **107**, 14 431–14 436. (doi:10.1073/pnas.1003111107)
16. Johansson RS, Westling G, Backstrom A, Flanagan JR. 2001 Eye–hand coordination in object manipulation. *J. Neurosci.* **21**, 6917–6932.
17. Gesierich B, Bruzzo A, Ottoboni G, Finos L. 2008 Human gaze behaviour during action execution and observation. *Acta Psychol.* (*Amst.*) **128**, 324–330. (doi:10.1016/j.actpsy.2008.03.006)
18. Johansson RS, Westling G. 1988 Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Exp. Brain Res.* **71**, 59–71. (doi:10.1007/BF00247522)
19. Flanagan JR, Beltzner MA. 2000 Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nat. Neurosci.* **3**, 737–741. (doi:10.1038/76701)
20. Flanagan JR, Bittner JP, Johansson RS. 2008 Experience can change distinct size–weight priors engaged in lifting objects and judging their weights. *Curr. Biol.* **18**, 1742–1747. (doi:10.1016/j.cub.2008.09.042)
21. Shim J, Carlton LG. 1997 Perception of kinematic characteristics in the motion of lifted weight. *J. Mot. Behav.* **29**, 131–146. (doi:10.1080/00222899709600828)
22. Bingham GP. 1987 Kinematic form and scaling: further investigations on the visual perception of lifted weight. *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 155–177. (doi:10.1037/0096-1523.13.2.155)
23. Hamilton AF, Joyce DW, Flanagan JR, Frith CD, Wolpert DM. 2007 Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychol. Res.* **71**, 13–21. (doi:10.1007/s00426-005-0032-4)
24. Alaerts K, Swinnen SP, Wenderoth N. 2010 Observing how others lift light or heavy objects: which visual cues mediate the encoding of muscular force in the primary motor cortex? *Neuropsychologia* **48**, 2082–2090. (doi:10.1016/j.neuropsychologia.2010.03.029)
25. Wolpert DM, Flanagan JR. 2001 Motor prediction. *Curr. Biol.* **11**, R729–R732. (doi:10.1016/S0960-9822(01)00432-8)
26. Shadmehr R, Smith MA, Krakauer JW. 2010 Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* **33**, 89–108. (doi:10.1146/annurev-neuro-060909-153135)
27. Ballard DH, Hayhoe MM, Li F, Whitehead SD. 1992 Hand–eye coordination during sequential tasks. *Phil. Trans. R. Soc. Lond. B* **337**, 331–338 (discussion 8–9). (doi:10.1098/rstb.1992.0111)
28. Prablanc C, Pelisson D, Goodale MA. 1986 Visual control of reaching movements without vision of the limb. I. Role of retinal feedback of target position in guiding the hand. *Exp. Brain Res.* **62**, 293–302. (doi:10.1007/BF00238848)
29. Saunders JA, Knill DC. 2004 Visual feedback control of hand movements. *J. Neurosci.* **24**, 3223–3234. (doi:10.1523/JNEUROSCI.4319-03.2004)
30. Crawford JD, Medendorp WP, Marotta JJ. 2004 Spatial transformations for eye–hand coordination. *J. Neurophysiol.* **92**, 10–9. (doi:10.1152/jn.00117.2004)
31. Land MF, Mennie N, Rusted J. 1999 The roles of vision and eye movements in the control of activities of daily living. *Perception* **28**, 1311–1328. (doi:10.1068/p2935)
32. Flanagan JR, Terao Y, Johansson RS. 2008 Gaze behavior when reaching to remembered targets. *J. Neurophysiol.* **100**, 1533–1543. (doi:10.1152/jn.90518.2008)
33. Eshuis R, Coventry KR, Vulchanova M. 2009 Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychol. Sci.* **20**, 438–440. (doi:10.1111/j.1467-9280.2009.02317.x)
34. Webb A, Knott A, Macaskill MR. 2010 Eye movements during transitive action observation have sequential structure. *Acta Psychol.* (*Amst.*) **133**, 51–56. (doi:10.1016/j.actpsy.2009.09.001)
35. Brass M, Schmitt RM, Spengler S, Gergely G. 2007 Investigating action understanding: inferential processes versus action simulation. *Curr. Biol.* **17**, 2117–2121. (doi:10.1016/j.cub.2007.11.057)
36. Wolpert DM, Doya K, Kawato M. 2003 A unifying computational framework for motor control and social interaction. *Phil. Trans. R. Soc. Lond. B* **358**, 593–602. (doi:10.1098/rstb.2002.1238)
37. Oztop E, Wolpert D, Kawato M. 2005 Mental state inference using visual control parameters. *Cogn. Brain Res.* **22**, 129–151. (doi:10.1016/j.cogbrainres.2004.08.004)
38. Wolpert DM, Diedrichsen J, Flanagan JR. 2011 Principles of sensorimotor learning. *Nat. Rev. Neurosci.* **12**, 739–751. (doi:10.1038/nrn3112)
39. Alaerts K, Senot P, Swinnen SP, Craighero L, Wenderoth N, Fadiga L. 2010 Force requirements of observed object lifting are encoded by the observer's motor system: a TMS study. *Eur. J. Neurosci.* **31**, 1144–1153. (doi:10.1111/j.1460-9568.2010.07124.x)
40. Cannon EN, Woodward AL. 2008 Action anticipation and interference: a test of prospective gaze. In *Proc. 30th Annual Conf. Cogn. Sci. Soc.* (eds BC Love, K McRae, VM Sloutsky), pp. 981–984. Austin, TX: Cognitive Science Society.