J Neurophysiol 96: 1358–1369, 2006.
First published May 10, 2006; doi:10.1152/jn.00227.2006.

Eye Movements When Observing Predictable and Unpredictable Actions

Gerben Rotman,[1] Nikolaus F. Troje,[1] Roland S. Johansson,[2] and J. Randall Flanagan[1]
[1]*Department of Psychology and Centre for Neuroscience Studies, Queen's University, Kingston, Ontario Canada; and* [2]*Section for Physiology, Department of Integrative Medical Biology, Umeå University, Umeå, Sweden*

Submitted 2 March 2006; accepted in final form 8 May 2006

Rotman, Gerben, Nikolaus F. Troje, Roland S. Johansson, and J. Randall Flanagan. Eye movements when observing predictable and unpredictable actions. *J Neurophysiol* 96: 1358–1369, 2006. First published May 10, 2006; doi:10.1152/jn.00227.2006. We previously showed that, when observers watch an actor performing a predictable block-stacking task, the coordination between the observer's gaze and the actor's hand is similar to the coordination between the actor's gaze and hand. Both the observer and the actor direct gaze to forthcoming grasp and block landing sites and shift their gaze to the next grasp or landing site at around the time the hand contacts the block or the block contacts the landing site. Here we compare observers' gaze behavior in a block manipulation task when the observers did and when they did not know, in advance, which of two blocks the actor would pick up first. In both cases, observers managed to fixate the target ahead of the actor's hand and showed proactive gaze behavior. However, these target fixations occurred later, relative to the actor's movement, when observers did not know the target block in advance. In perceptual tests, in which observers watched animations of the actor reaching partway to the target and had to guess which block was the target, we found that the time at which observers were able to correctly do so was very similar to the time at which they would make saccades to the target block. Overall, our results indicate that observers use gaze in a fashion that is appropriate for hand movement planning and control. This in turn suggests that they implement representations of the manual actions required in the task and representations that direct task-specific eye movements.

INTRODUCTION

An important idea in psychology is that the perception of action, including speech, involves the action system (Liberman and Whalen 2000). The idea is that observers understand or decode the actions of other people by activating their own action system at some subthreshold level. This idea has gained support in neuroscience with a number of neurophysiological (Avikainen et al. 2002; di Pellegrino et al. 1992; Fadiga et al. 1995; Gallese et al. 1996; Gangitano et al. 2001; Hari et al. 1998; Kohler et al. 2002; Nishitani and Hari 2000; Strafella and Paus 2000; Umiltà et al. 2001) and imaging studies (Decety et al. 1997; Grafton et al. 1996; Iacoboni et al. 1999, 2001; Rizzolatti et al. 1996) showing that brain circuits active during action are also active during action observation. These findings have given rise to the direct matching hypothesis that posits that, when people observe action, they implement covert action plans that, in real time, match the action plans executed by the actor (for review, see Rizzolatti et al. 2001).

Recently, we examined action plans used in action observation by recording observers' eye movements while they watched an actor perform an object manipulation task (Flanagan and Johansson 2003). When people perform object manipulation tasks themselves, they use task-specific eye movements that support hand movement planning and control (Hayhoe and Ballard 2005; Johansson et al. 2001; Land and Furneaux 1997; Land et al. 1999). In particular, through saccadic gaze shifts, subjects fixate forthcoming grasp sites, obstacles, and landing sites where objects will be subsequently grasped, moved around, and placed, respectively (Johansson et al. 2001). Given that the eyes are free to move when observing such tasks, the direct matching hypothesis predicts that people will produce similar eye movements when observing and performing the task. We confirmed this prediction by showing that when people observe a familiar block-stacking task, the coordination between their gaze and the actor's hand is very similar to the gaze-hand coordination when they perform the task themselves. In both cases, observers proactively shifted gaze to forthcoming grasp and landing sites. Thus observers' gaze predicts forthcoming task events and does not simply follow the visual events as they unfold. These findings suggest that during the observation of object manipulation tasks, people implement task-specific eye movement programs that are directed by representations of the manual actions required in the task (Flanagan and Johansson 2003). As such, these results provide strong support for the direct matching hypothesis.

Support for the notion that action observation involves prediction of forthcoming actions has been provided by Cisek and Kalaska (2004). They showed in monkeys that once the observer obtains a cue indicating the forthcoming action most task-related neurons in dorsal premotor cortex—a region involved in movement selection and planning—are activated in advance of an observed action. Likewise, motor circuits are activated in human imaging experiments when information specifying a particular action is provided to the observers (Jeannerod et al. 1995; Johnson et al. 2002; Ramnani and Miall 2004).

In our previous study (Flanagan and Johansson 2003), we examined a block-stacking task that was completely predictable. That is, the actor showed the task to observers before data collection and repeated the same task several times. The main objective of this study was to compare observers' eye movements during predictable and unpredictable movement phases when they watch an object manipulation task. Because many, if not most, actions cannot be predicted, in advance, by observers, any general theory of action observation or action understanding ought to account for such actions. We used a task in which observers watched an actor reach for, lift, and

Address for reprint requests and other correspondence: J. R. Flanagan, Dept. of Psychology, Queen's Univ., Kingston, Ontario K7L 3N6, Canada (E-mail: flanagan@post.queensu.ca).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

replace first one block and then a second. A cue, available only to the actor, indicated which of the two blocks to pick up first. Thus the observer could predict, by simple deduction, which block would be reached for and lifted second but did not have advance information about which block would be targeted first. We expected that when the target block cannot be predicted in advance, observers' eye movements would not match those of the actor in real time. However, we hypothesized that observers would nevertheless implement task-specific eye movements as quickly as possible based on information provided by the kinematics of the actor's movement. That is, we predicted that during the first hand movement, observers would proactively direct their gaze to the block ahead of the actor's hand. We also predicted that during the second hand movement, observers would exploit knowledge of the task to predict the target block and that this would result in earlier gaze shifts to the target block. Confirmation of these predictions would support the notion that observers implement eye movements directed by representations of the manual actions required in the task but would also show that these representations need not be time-locked to those used by the actor.

To further assess the observers' ability to use visual cues about the actor's movements to predict the goal of the upcoming action, we carried out an additional experiment in which observers tried to predict the goal based on partial viewing of the actor's movements. Observers watched computer animations (based on measured kinematics) of the actor performing the task and were asked to guess which of the two blocks the actor was reaching for. We varied the duration of the viewed animation (relative to onset of the reach movement) and related this to observers' saccade behavior in the on-line experiment.

METHODS

Participants

Thirty-two undergraduate students participated in three experiments after providing their informed consent and receiving payment for their participation. Nine participated in *experiment 1*, 10 in *experiment 2*, and 13 in *experiment 3*. All participants had normal or corrected-to-normal vision and did not report or exhibit any obvious neurological deficits. The local university ethics board approved the experiments, which complied with the Declaration of Helsinki. Data from three of the participants in *experiment 2* were excluded from further analysis. One of these participants did not move her eyes and kept staring at one position in the middle of the table, a second looked around the room watching other things than the scene with the blocks, and the third fixated the actor's hand at all times, unlike all of the remaining participants, who rarely, if ever, fixated the hand.

Apparatus and stimuli

In *experiments 1* and *2*, we recorded participants' eye movements while they observed an actor perform a block manipulation task. The observers sat at a table with their forehead resting against a fixed headband. An infrared video-based eye-tracking device (RK-726PCI pupil/corneal tracking system, ISCAN, Burlington, VT), mounted below the headband, recorded the gaze position of the right eye in a defined work plane at 240 Hz. A small bite bar was used to further minimize head movements.

In one experimental condition (*experiment 1*; Fig. 1*A*), the observers viewed the actor from the side and the defined work plane was the actor's midsagittal plane (or *xz* plane). In another condition (*experiment 2*; Fig. 1*B*), the actor was viewed from the front and the defined work plane was the horizontal plane of the tabletop (or *xy* plane). In either condition, the observer could not see the actor's face, but could see his body from the shoulders down. The observers were instructed to simply watch the actor picking up small wooden blocks that were lying on the table. In both conditions, a motion capture system (Vicon, Oxford Metrics, Oxford, UK) recorded the movements of the actor's right hand and arm by tracking the positions of 13 reflective markers (Fig. 1). Arm markers were placed on the acromiom process of the shoulder, mid-upper arm, elbow joint, and the radial and ulnar processes of the wrist. For both the thumb and index finger, markers were placed at the metacarpophalangeal joint, the proximal and distal interphalangeal joints, and the tip. With these markers, we could calculate the position and orientation of the upper arm, forearm, hand and the phalanges of the index finger, and thumb.

In *experiment 3*, observers watched computer animations of the actor picking up the blocks. We tried to make these animations resemble the settings of *experiments 1* and *2* as much as possible. Only part of the movement was displayed and by pressing one of two buttons on a keyboard, participants indicated which block they believed that the animated character would pick up. The animations, made with Poser (Poser 5.0.2, Curious Labs, Santa Cruz, CA), were based on kinematic recordings of the actor's movements in *experiments 1* and *2*. Specifically, we assigned the measured movement of the actor's right arm, hand, index finger, and thumb to the corresponding body parts of the animated character, and inferred the exact block positions from the actor's hand movement. Although the participants could see the eyes of the animated character, the eye and head did not

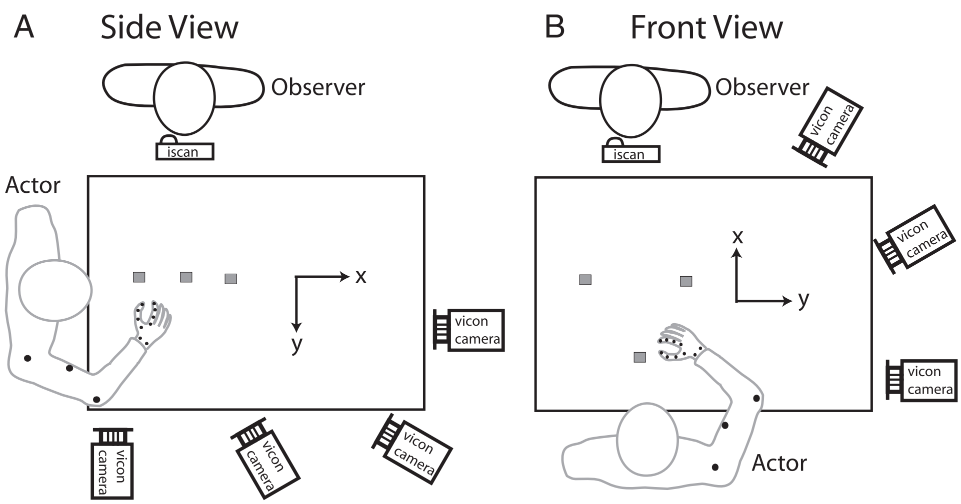

FIG. 1. Schematic drawings of the experimental settings used in *experiments 1* (*A*) and *2* (*B*), in which observers viewed the actor from the side and front, respectively. Gray squares show locations of the start block, closest to the actor, and the 2 target blocks. Locations of 13 reflective markers attached to the actor's arm and hand and optical cameras used to track positions of these markers are shown. A video-based eye tracking system recorded the observer's gaze in either the actor's sagittal (side view) or horizontal (front view) plane.

move and thus did not convey any information about the target. The animations were projected onto a wall and scaled to the size of the real actor. In addition, the animated actor resembled the real actor by sitting behind a table (Fig. 2). The animations were played at normal time. Note that we did not measure eye movements in *experiment 3*. Our goal in including this experiment was to compare observers' ability to predict target blocks, as revealed by natural gaze behavior (*experiments 1* and *2*) with their performance in a task in which they were explicitly asked to predict the goal. It is quite possible that gaze behavior in the latter task would be quite different from in action observation because the stimuli and the demands of the two tasks are very different.

Procedure

EXPERIMENTS 1 AND 2. There were always three blocks on the table (Fig. 1). The one located closest to the actor served as a start block and the other ones as target blocks. Immediately after a block was lifted, the actor replaced it in the same location. Between trials, the actor rested his hand about 5 cm to the side of the start block. In each trial, the actor picked up and replaced the start block, one of the target blocks, the start block again, and then the other target block before returning his hand to the vicinity of the start block. The observer did not know, in advance, which target block the actor would pick up first. Because the observers could not see the actor's face, they could not use the actor's gaze to predict his actions.

One actor performed all of the trials in *experiments 1* and *2* and was naïve with respect to the specific hypotheses under study. A visual signal, only available to the actor, informed him which of the target blocks to pick up first. The actor picked up the blocks using a precision grip with the tips of the thumb and index finger contacting the near and far sides of the block, respectively.

EXPERIMENT 1. This experiment consisted of two sessions. In session 1, the blocks were arranged along a line straight ahead of the actor (Fig. 1*A*) and located on top of a box so that they would be in the middle of the observers' field of view (Fig. 2*A*). The start block was located about 20 cm from the actor's trunk, and the centers of the first and second target blocks were located 25 and 40 cm from the start block, respectively. The start block had a size of $2 \times 2 \times 2$ cm. The target blocks were either $2 \times 2 \times 2$ or $2 \times 2 \times 10$ cm and will be referred to as the 2- and 10-cm target blocks. We used three different target block layouts: 2–2, 2–10, or 10–2 (closest target block to furthest target block). The 10-cm block was lying flat with its long axis in line with the three blocks such that a large grip aperture was required to lift it. In total, session 1 consisted of 120 trials with the three different block layouts each presented in 40 trials. We changed the layout every 10 trials in a randomized order. The order in which the two target blocks were picked up was also randomized but subject to the constraint that in each set of 10 trials, each target block would be picked up first in 5 trials.

The second session of *experiment 1* was similar to the first with the exceptions that there was no 2–2 layout and that the orientation of the 10-cm block was different. In the second session, the 10-cm block was standing with the long axis vertical (as in Fig. 2*A*) and the actor grasped near the top of the block. There was no 2–2 layout in session 2. The block layouts with the vertical 10-cm block in the near and far positions are referred to as 10up-2 and 2-10up, respectively.

We included different block layouts to manipulate both the trajectory of the hand and the grip aperture (distance between the tip of the index finger and the tip of the thumb). The aim was to examine whether observers exploit these kinematic features to predict the goal of the actor's first movement to guide their gaze.

EXPERIMENT 2. In this experiment, the observers had a front view of the actor (Fig. 1*B*) and only 2-cm^3 blocks were used and arranged in a triangular configuration. The target blocks were located 25 cm further away from the actor than the start block (along the midline) and were displaced 18 cm to the left and the right. Each subject observed 72 trials. The target block lifted first was randomized, but within each set of 12 trials, both the left and the right target block was picked up first in 6 trials.

EXPERIMENT 3. Naïve participants observed animations of only part of the actor's reaching movement, and the task was to guess which block the actor was reaching for. Participants received no feedback about correct responses. By varying the duration of the animated reach, we sought to determine how much of the actor's movement the observers needed to see to be able to predict the target block. We compared these times with the timing of the saccades that the subjects in the first two experiments directed to the target blocks. Because we were unable to animate the start block, the start of the reach for a target block was preceded by a pantomimed pickup and replace of the start block performed by the animated actor.

For each of the five block layouts in *experiment 1* and the one in *experiment 2*, we randomly selected six movements for animation, three for each target block. For each animated movement, we constructed five animations of different duration that all started with the animated actor's hand in the rest position. The animations ended 5, 10, 15, 20, or 25 frames (at 60 Hz; 83–417 ms) after the hand started moving from the start block toward one of the two target blocks. Thus there were a total of 180 different animations (6 layouts, 2 target blocks, 3 movements per target block, 5 durations). At the end of the animation, the screen went blank.

For each of the six block layouts, we constructed three sets of animations, each set consisting of five animations of different duration based on a reach to the near block and similarly five for the far block. The 10 animations within a set were each shown three times, so that a set contained 30 trials (2 reach targets, 5 durations, 3 repetitions). Both the order of trials within a set and the order of sets were randomized.

Data analysis

From the hand and eye movement data collected in *experiments 1* and *2,* we determined hand and eye movement onset and offset times based on velocity criteria. For hand and gaze movements, we used

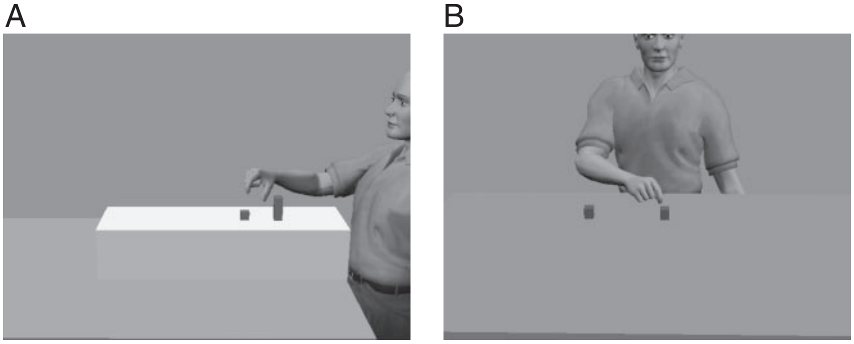

FIG. 2. Frames from the animations used in *experiment 3*. *A*: frame from an animation based on a hand movement from *experiment 1*, *session 2*. *B*: frame based on a hand movement from *experiment 2*.

thresholds of 0.1 and 1 m/s, respectively. We defined the position of the hand in the work plane as the average position of the markers located at the tips of the index finger and the thumb. We examined the timing of gaze movements relative to hand movement as well as the frequency of gaze shifts to either target block. Repeated-measures ANOVA were used to compare various measures across conditions and an α level of 0.05 was considered significant.

RESULTS

Experiment 1

The observers' gaze occasionally tracked the hand or was directed away from the general vicinity of the actor. However, such behavior occurred in <1% of the trials and only in some subjects. These trials were not further analyzed. The general pattern was that observers proactively fixated the start block and target blocks ahead of the actor's hand contacting them. Figure 3*A* shows observer's gaze position and actor's hand position and corresponding velocity records for a trial in which the actor picked up the near block first and then the far block in the 2–2 layout of the side-view session (see Fig. 1*A*). The observer fixated each block before the actor's hand arrived and shifted gaze to the next block shortly after the actor started to move his hand away from the current block toward the next block. In this trial, the observer's gaze correctly anticipated the target blocks and the gaze behavior is very similar to the behavior seen when the sequence in which blocks will be lifted is known to the observer (Flanagan and Johansson 2003). Figure 3*B* shows a single trial in which the actor first picked up the far block (again with the 2–2 layout). In this trial, the observer's gaze initially shifted to the near block (again shortly after the actor's hand first moved from the start block) and shifted to the far block at around the time the hand passed over the near block. As will be shown below, this pattern was quite typical; when the actor first lifted the far block, observer frequently fixated the near block during the early part of the reach.

Figure 4 shows the average number of saccades that observers made in different situations in *experiment 1*. When the hand went from the start block to the near block (Fig. 4, *top row*), a saccade was made from the start block to the near block in the vast majority of trials. This was the case for the first and second hand movements and for all block layouts. Across the 10 cases

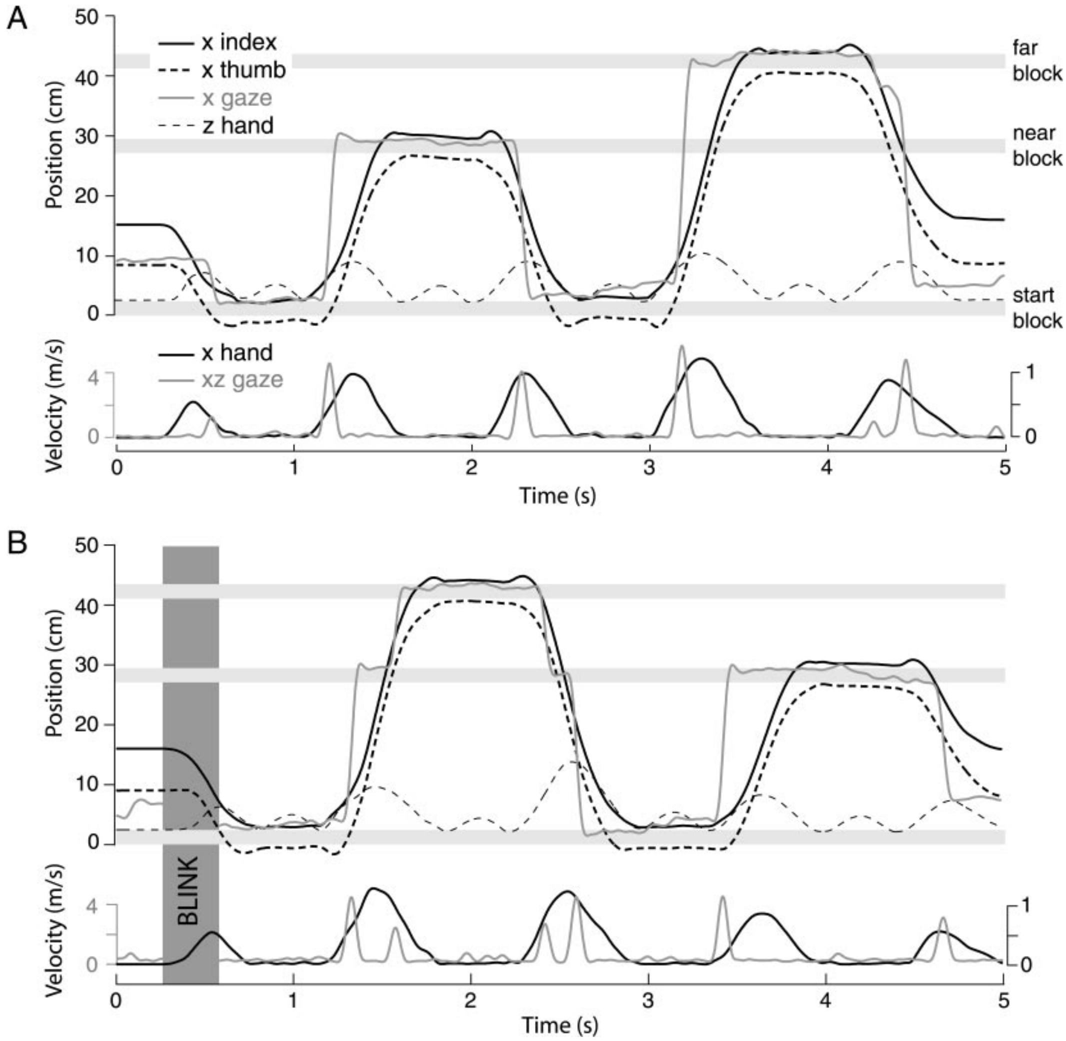

FIG. 3. Single position and tangential velocity records from 2 side-view trials. *x*-positions of index finger, thumb, and gaze, *z*- (vertical) position of the hand, *x*-velocity of the hand, and tangential velocity of gaze in the *xz* plane are shown. Separate calibration scales are provided for the hand and gaze velocities. Horizontal bars show locations of the 3 blocks.

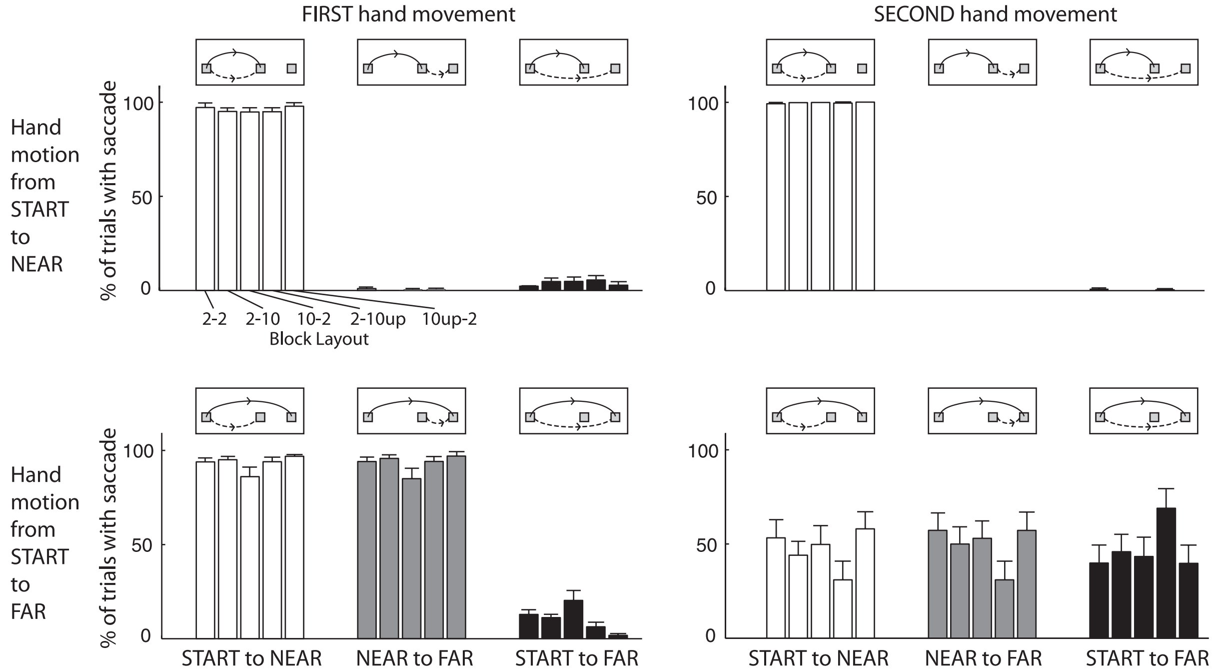

FIG. 4. Frequency, expressed as a percentage, of different types of saccades observed in *experiment 1*. Frequency is shown as a function of the hand movement target (near or far block), hand movement number (1st or 2nd), and block layout. Small drawings above bars schematically show hand (solid line) and gaze (broken line) movement. Order of block layouts is in all cases as is indicated in the *top left panel*. Bars represent averages based on participant means and error bars represent SE.

(2 movements, 5 block layouts), the percentage of trials in which a saccade was made to the near block ranged from 95.3 to 100%. This percentage was higher for the second hand movement ($F_{1,8}$ = 15.58; P = 0.004). Block layout did not influence the percentage ($F_{4,32}$ = 0.45; P = 0.77), and there was no interaction between movement number and block layout ($F_{4,32}$ = 0.59; P = 0.67).

When the hand went from the start block to the far block (Fig. 4, *bottom row*), the pattern of saccades differed between the first and second hand movement. During the first hand movement, two saccades were made in almost every trial—a saccade from the start block to the near block followed by a saccade from the near block to the far block. Few saccades (1.7–13.3% across the 5 layouts) went straight from the start block to the far block. During the second hand movement, observers shifted their gaze directly from the start block to the far block in 47.8% of the trials on average, and this percentage was similar across block layouts ($F_{4,32}$ = 1.36; P = 0.27; note that only 7 of the 9 observers produced initial saccades to the far target in all 5 block layouts). Although observers more frequently shifted their gaze directly to the far block when this block was the goal of the second hand movement (compared with the 1st), they nevertheless first fixated the near block in about one half of the trials. This occurred despite the fact that observers should have been able to deduce where the hand would go.

Figure 5*A* shows onset times of saccades made in *experiment 1*, relative to the start of the actor's hand movement. First, consider saccades during the first hand movement (*left column*). On average, the initial saccades from the start block to the near block (white bars) were initiated 54 ± 10 (SE) ms after hand movement, and the saccade onset time did not depend on whether the hand goal was the near or far block ($F_{1,8}$ = 4.78; P = 0.06) or on the block layout ($F_{4,32}$ = 1.47; P = 0.96). When the far block was the goal of the first hand movement, observers made the second saccade from the near block to the far block (gray bars) on average 330 ms after hand movement onset. However, the timing of the second saccade depended on the block layout ($F_{4,32}$ = 11.49; P < 0.001). As can be readily appreciated in Fig. 5*A*, these second saccades occurred relatively earlier and later for the 2-10up and 10up-2 layouts, respectively, compared with the three other layouts.

Next, consider saccades during the second hand movement (Fig. 5*A*, *right column*). Initial saccades from the start block to the near block were initiated earlier ($F_{1,8}$ = 12.16; P = 0.008) when the far block was the goal (mean 7 ± 11 ms) than when the near block was the goal (mean 42 ± 10 ms). When the far block was the goal, initial saccades to the near block were initiated earlier ($F_{1,8}$ = 16.85; P = 0.003) during the second hand movement compared with the first. The second saccades from the near block to the far block (mean 267 ± 6 ms) also occurred earlier ($F_{1,8}$ = 65.60; P < 0.001) during the second hand movement. However, as in the first hand movement, the timing of these second saccades varied with block layout ($F_{4,32}$ = 8.90; P < 0.001), with earlier and later saccades observed for the 2-10up and 10up-2 layouts, respectively. No reliable effect of movement number was observed when the near block was the goal ($F_{1,8}$ = 4.58; P = 0.07). When the second hand movement was directed to the far block, an appreciable number of saccades shifted gaze directly from the start block to the far block (see Fig. 4). The average onset time of these saccades (Fig. 5*A*, black bars) was 83 ± 15 ms, and there was no effect of block layout ($F_{4,24}$ = 1.94; P = 0.14). These times were different from the onset time of saccades to the near block during the first hand movement ($F_{1,8}$ = 18.39; P = 0.003).

Figure 5*B* shows the *x*-position of the hand, relative to the center of the start block, at the time of saccade onset. (Recall

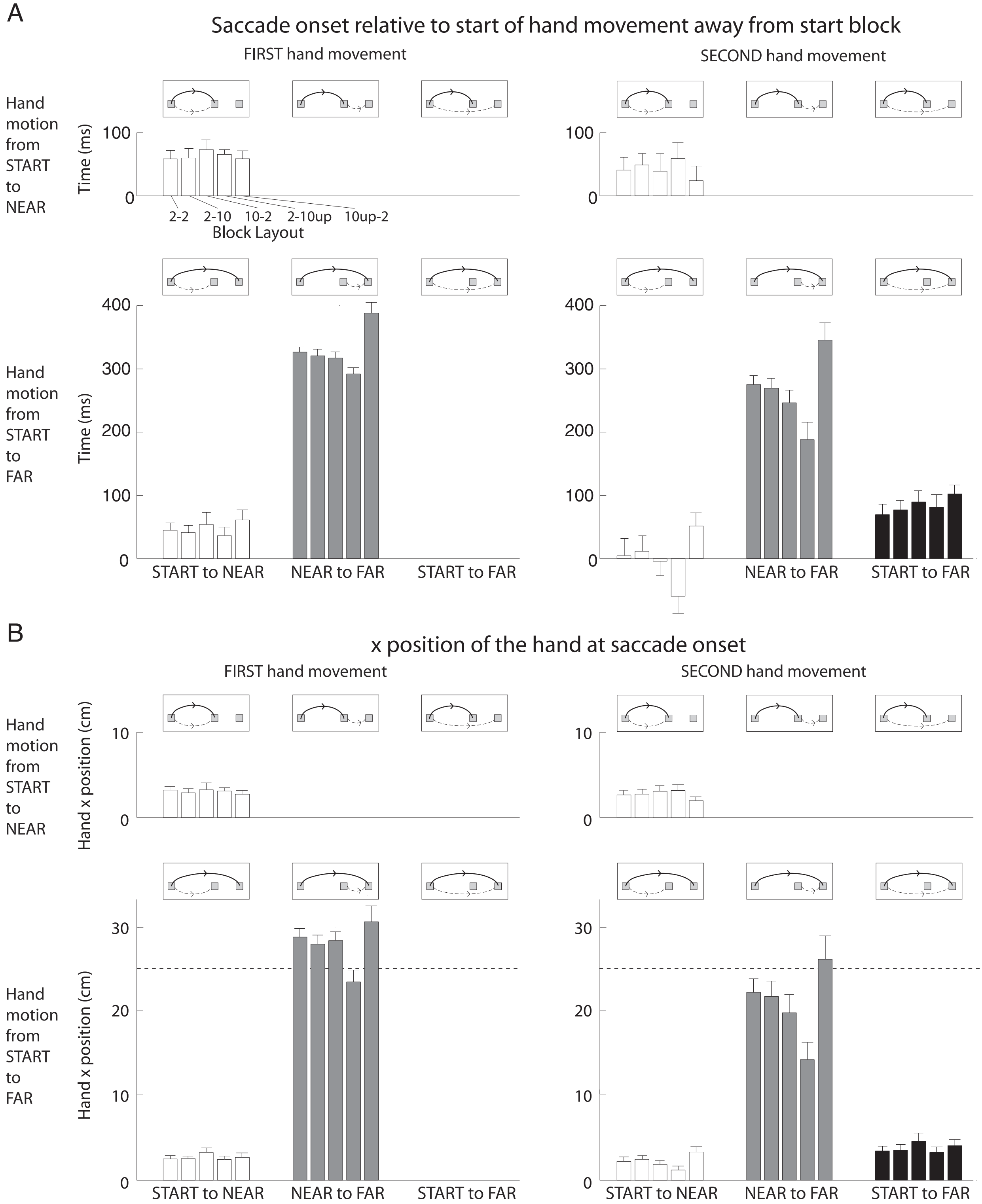

FIG. 5. Onset times of saccades made in *experiment 1* (*A*) and the *x*-position of the hand at these times (*B*). These variables are shown as a function of the hand movement target (near or far block), hand movement number (1st or 2nd), and block layout. Small drawings above bars schematically show hand (solid line) and gaze (broken line) movement. Order of block layouts is in all cases as is indicated in the *top left panel* in *A*. Bars represent averages based on participant means and error bars represent SE.

that the *x*-axis is aligned with the blocks.) Overall, the pattern of results across conditions is similar to that observed for saccadic onset times. On average, initial saccades to the near block (white bars) were initiated when the hand had traveled 2.84 ± 0.37 and 2.40 ± 0.36 cm for the first and second hand movements, respectively; a difference that was reliable ($F_{1,8}$ = 7.39; P = 0.03). There was no reliable effect of hand goal on these distances ($F_{1,8}$ = 4.25; P = 0.07) and no interaction between movement number and hand goal ($F_{1,8}$ = 0.42; P = 0.45). The effect of layout was reliable ($F_{4,32}$ = 4.14; P =

0.008), and there was an interaction between hand movement number and layout ($F_{4,32}$ = 3.21; P = 0.025). In particular, during second hand movements directed to the far block, the hand had traveled a relatively short and long distance for the 2-10up and 10up-2 layouts, respectively, by the time the first saccade to the near block was initiated. A three-way interaction among movement number, goal, and layout was observed ($F_{4,32}$ = 4.85; P = 0.004).

When the reach goal was the far target, second saccades from the near to the far block (gray bars) were initiated, on average, when the hand had traveled 27.82 and 20.92 cm for the first and second hand movement, respectively; a difference that was reliable ($F_{1,8}$ = 64.7; $P < 0.001$). Note that the near block was located ~25 cm from the start block (Fig. 5*B*, dashed horizontal lines). Thus during the first hand movement, second saccades to the far block were initiated roughly when the hand passed over the near block. The distance traveled by the hand at the onset of these second saccades depended on block layout ($F_{4,32}$ = 5.51; P = 0.004), as was the case at the onset of the initial saccades, but there was no interaction between hand movement number and layout ($F_{4,32}$ = 1.46; P = 0.24). For both hand movements, the distance was smaller and larger for the 2-10up and 10up-2 layouts, respectively, than for the three other layouts. When gaze shifted directly from the start block to the far block, which only happened during the second hand movement to the far block (black bars), the average *x*-position of the hand was 3.80 ± 0.68 cm at saccade onset, and this was not influenced of block layout ($F_{4,24}$ = 0.83; P = 0.52).

We included different block layouts to manipulate both the trajectory of the hand and the grip aperture to examine whether observers exploit these kinematic features when directing their gaze. We will next address this question. Figure 6*A* shows average hand paths for first hand movements directed to the near and far blocks in the 10up-2, 2-10up, and 2–2 layouts. To obtain these paths, hand movements were time normalized (100 samples), and the average *x*- and *z*-positions of the hand were computed for each time sample. The average hand position at the onset and offset of the initial saccade to the near block and at the onset of the subsequent saccade from the near to the far block during the hand movements to the far block are indicated by vertical lines. Thus the distance between the first and second lines represents the hand travel during the first saccade and the distance between the second and third lines represents the hand travel during the fixation of the block. The dots along each hand path mark 50-ms intervals, based on the average movement time, and lines connect corresponding time marks.

Several points regarding Fig. 6*A* can be emphasized. First, given that the delay between saccade onset and the visual information driving the saccade is about 100 ms at a minimum (Caspi et al. 2004; Lisberger et al. 1975; Pare and Hanes 2003), it is clear that observers could determine that the far block was the target well before the actor's hand passed over the near block. Note that 100 ms represents a minimum estimate, and it is possible that, with our particular experimental set-up, the time period is longer. If so, then observers may have been able to determine the target even earlier. (This point will be developed further below when considering the results of *experiment 3*.) The open circles on the hand paths to the far target show the average position of the hand 100 ms before the saccade to the far block. Second, the earliest saccades from the near block to the far block occurred with the 2-10up layout, which had the greatest separation between paths. This suggests that observers were able to exploit knowledge about the two hand paths when determining the target. However, it should be noted that the latest saccades were observed in the 10up-2 layout even though the separation in hand paths was greater than in the 2–2 layout. Third, the timing of the first and second saccades indicates that observers' primarily based their decisions regarding the target on visual information obtained while fixating the near target. On average, these fixation durations were 257, 186, and 220 ms for the 10up-2, 2-10up, and 2–2 layouts, respectively (with

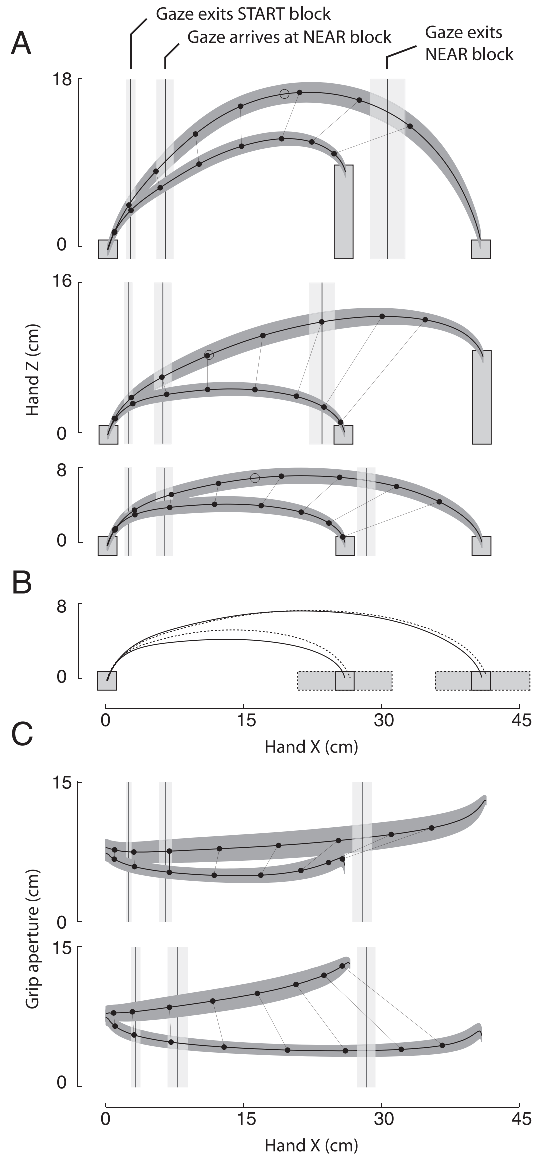

FIG. 6. *A*: average hand paths for 1st hand movements directed to the near and far blocks in the 10up-2, 2-10up, and 2–2 layouts. Height of each gray region represents ±SD in *z*. Vertical lines represent average hand position at saccade onset and offset and width of the gray bars represent ±SE (n = 9). Open circles indicate position of the hand 100 ms before gaze exits the NEAR block. Dots along each path mark 50-ms intervals. *B*: average hand paths for the reaches in the 2–10 and 10–2 layouts; solid and dotted curves show average paths to the 2- and 10-cm blocks, respectively. *C*: average grip aperture curves in the 2–10 (*top*) and 10–2 (*bottom*) layouts. Vertical lines and bars, gray regions, and dots correspond to those shown in *A*.

corresponding SEs of 17, 10, and 15 ms). In our previous work on object manipulation, we found that actors fixated obstacles for ~90 ms and argued that these obstacle fixations were too brief to allow visual processing to influence the subsequent saccade (Johansson et al. 2001). In contrast, the longer fixations at the near block in this study would allow some 85–160 ms of visual processing to influence the next saccade (again assuming that 100 ms represents the minimum time between visual processing and saccade onset).

Figure 6*B* shows the average hand paths directed to the 2- (solid) and 10-cm (dotted) blocks in the 2–10 and 10–2 layouts. The hand paths for both the near and far blocks did not vary appreciably as a function of block width. Figure 6*C* shows average grip aperture, as a function of the hand *x*-position, for reaches to the near and far blocks with the 2–10 (*top*) and 10–2 (*bottom*) layouts. The difference between the two aperture curves was clearly far greater for the 10–2 layout than for the 2–10 layout. However, this did not translate into a difference in the timing of the saccade from the near block to the far block (see also Fig. 5). Together, the results shown in Fig. 6 indicate that observers did not exploit visual information related to grip aperture when deciding whether or not to shift their gaze to the far target, even though this information was available for the 10–2 layout.

There are a number of possible features of the hand trajectory that observers may have used when predicting the target object. For example, they could have processes the height of the hand path—a spatial feature—or the speed of the hand—a temporal feature because both of these features distinguish, on average, the hand paths to the near and far target block (Fig. 6). To explore this issue, we checked whether the variation in the timing of saccades, from the near block to the far block, correlated with movement duration or the maximum height of the hand during the first hand movement. We chose these two dependent variables because they should reflect the temporal and spatial properties of the hand trajectories, respectively. Separate regressions were carried out for each observer and for each block in the side view condition. The slope of the least squares regression line predicting saccade onset time based on movement duration was significant in only one observer and in only one layout (2-10up, 1.01 ms/ms, $P = 0.002$). No significant slopes were found for the regression lines predicting saccade onset time based on the maximum height of the hand.

Experiment 2

In *experiment 2*, observers viewed the actor from the front, and the target blocks were located to the right and left of the actor's (and observer's) midline (Fig. 1). In this experiment, observers almost exclusively made saccades to the block that the actor was about to pick up. This was the case during both the first and second hand movements. On average, across observers, saccades to the nontarget or "incorrect" block were made in <2% of all first hand movements and <0.3% of all second hand movements. Thus observers were clearly able to judge which block the actor was going to pick by the time they initiated the saccade away from the start block. Saccades were initiated later ($F_{1,6} = 9.9$; $P = 0.02$) during the first hand movement (mean 156 ± 16 ms) than during the second hand movement (mean 77 ± 26 ms). On average, the hand had traveled 7.6 ± 0.41 cm in the horizontal plane at the time of saccade onset during the first hand movement and 3.4 ± 0.37 cm during the second hand movement—a difference that was reliable ($F_{1,6} = 123.2$; $P < 0.001$).

Experiment 3

In the third experiment, participants watched videos of an animated actor reaching toward one of the two target blocks. In this experiment, the participants had to indicate which block the actor was reaching for after viewing only part of the hand movement trajectory. We varied the duration of viewing to determine how much of the movement the observers had to see to be able to accurately discriminate between reaches to either target block. We compared these time values with the timing of saccades, directed to the target block during first hand movements, in the first two experiments. If these saccades are initiated as soon as possible based on kinematic cues, participants in *experiment 3* should be accurate when provided with similar kinematic cues but inaccurate when provided with less kinematic information (i.e., when viewing less of the actor's movement).

Figure 7 shows the percentage of correct responses as a function of viewing time (black dots) relative to the start of the reach. Separate panels are shown for the single front view layout and each of the five side view layouts. Each side view panel also shows a cumulative frequency plot of onsets times for saccades from the near block to the far block during first hand movements, recorded in *experiment 1*. We selected these particular saccades because observers only directed their gaze to the far target (during 1st hand movements) once they knew this target was the goal based on kinematic cues. The front view panel shows a cumulative frequency plot of onset times of initial saccades bringing gaze from the start block to the target block during the first hand movement, recorded in *experiment 2*. Again, we selected these saccades because they were generated only when observers determined the target block based on kinematic cues. The solid vertical line in each panel represents the average saccadic onset time (gray bars represent ±SE; data from *experiments 1* and *2*), and the dashed vertical line located 100 ms to the left provides an estimate of when, on average, visual information could have last been used to trigger the saccade. We will refer to this time as the "final decision time."

First consider the side view condition. After viewing 333 ms of the animated hand movement, the average percentage of correct responses varied from 79 to 95% across the five different block layouts. The average times at which observers in *experiment 1* initiated a second saccade from the near block to the far block ranged from 292 to 387 ms in those block layouts. We can relate timing of the saccades made by observers in *experiment 1* to the percentage of correct guesses made by participants in *experiment 3* if we assume that the observers in *experiment 1* generated saccades to the far block as soon as they were certain that this block was the target. Because the distributions of saccade times are approximately normal and because observers only made saccades to the far block when they were certain it was the target (as revealed by the fact that they never made erroneous saccades to the far block), we can infer that, at the average final decision time, observers in *experiment 1* knew that the far block was the target in about one half the trials. Thus if participants in *experiment 3* were to

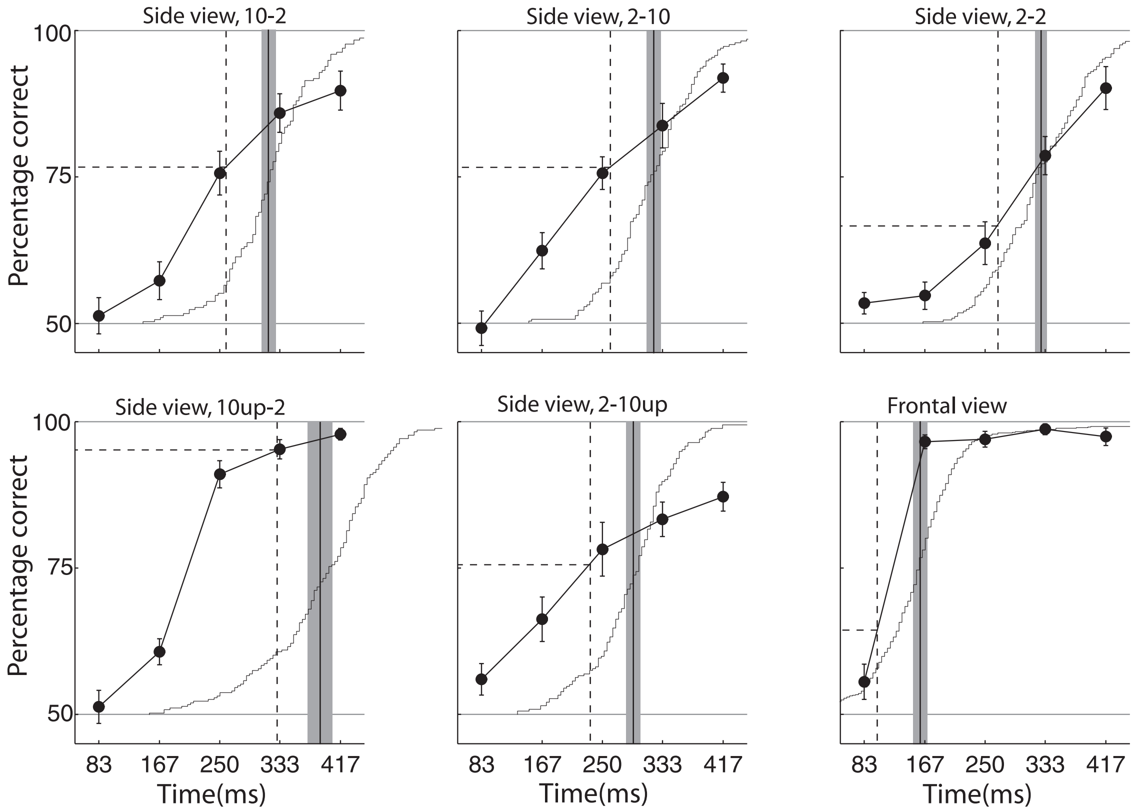

FIG. 7. Percentage of correct predictions of which block the animated actor was going to pick up as a function of the duration of the viewed hand movement. Each dot represents an average of participant means and vertical lines represent ±SE. Different panels show results for the 5 block layouts in the side view condition and the single layout in the front view condition. Each panel shows cumulative frequency distribution of onset times of saccades from the near block to the far target block (side view) or from the start block to the target block (front view) during the 1st hand movement. Vertical black line and gray bar indicate mean and SE of saccade onset times based on participant averages. Dashed vertical line is drawn 100 ms to the left of mean onset time.

view the hand up to the final decision time, they should receive sufficient information to correctly determine the target in about one half the trials. Thus we would expect that these participants would guess correctly (100%) in one half the trials and be at chance (50%) in the remaining trials. In other words, on average, we would expect them to guess correctly in 75% of the trials if they viewed the animated actors hand moving until the final decision time.

The dashed horizontal lines in Fig. 7 show the percentages of correct guesses (estimated using linear interpolation between measured values) at the average final decision time. These percentages were 67, 70, 61, 93, and 70 for the 10–2, 2–10, 2–2, 10up-2, and 2-10up layouts, respectively, and the average percentage was 73%. The fact that this is very close to 75% provides support for our assumption that observers in *experiment 1* generated saccades to the far block as soon as they knew that this block was the goal of the first hand movement. However, it should be stressed that there were clear differences across block layouts in the estimated percentage of correct guesses at the average final decision time. This suggests that observers in *experiment 1* may have relied on cues that were not available in the animations, at least for some block layouts.

When viewing the animated actor from the front, participants in *experiment 3* were able to accurately predict (mean 97 ± 1.1%) the target block after viewing 167 ms of the hand movement, but were roughly at chance when viewing only 83 ms of the movement. About one half of the saccades were initiated before the 167-ms mark (which was close to the average saccade onset time of 156 ms), but few were initiated within 83 ms of the start of hand movement. In this front view condition, the estimated final decision time occurred at a time when participants were at chance in identifying the target block. Again, this suggests that observers in *experiment 2* may have used different sensorimotor mechanisms and cues to control their eye movements compared with those used by the participants in the animation experiments. Nevertheless, the results of these animation experiments suggest that observers in *experiments 1* and *2* shifted their gaze to the target block more or less as soon as they could determine the correct block based on vision of the actor's hand. Indeed, on balance, the results indicate that these observers were extremely proficient at evaluating the actor's hand movement in real time.

Actors

Although this study focuses on action observation, we examined gaze behavior in eight actors performing the block manipulation tasks used in *experiments 1* and *2*. We also examined a variant of the task in *experiment 1* where the actors viewed the three aligned blocks from the side (i.e., from the same viewpoint as the observers in *experiment 1*). Overall, the results were very similar to those that we have reported previously (Flanagan and Johansson 2003; Johansson et al. 2001). When the blocks were aligned from right to left, the actors almost exclusively made saccades between the start block and the target block, and they hardly ever fixated the near (or middle) block when it was not the target. When the blocks were aligned ahead, the actors sometimes fixated the near or middle block when the far block was the target. These fixations may be similar to the optional obstacle fixations reported by Johansson et al. (2001).

DISCUSSION

We have previously shown that when watching an actor perform a familiar block-stacking task, observers generate proactive eye movements that are strikingly similar to those produced by the actor (Flanagan and Johansson 2003). We suggested that this common gaze behavior arises because both the actor and the observer implement similar motor representations of the manual task—or "action schema" (Land and Furneaux 1997)—that direct these task-specific eye movements. Although observers' manual actions may be inhibited, their eye movements are not.

The aim of this study was to probe observers' gaze behavior in a situation where the goal of the actor's movement could not be predicted in advance and had to be determined from the kinematics of the actor's movement. We showed that observers use gaze proactively under these conditions. Several key results support this conclusion. First, in all conditions, observers invariably directed their gaze to the target block ahead of the actor's hand such that their gaze was fixating the target block when the actor's hand contacted it. Second, observers shifted their gaze away from the start block shortly after the actor's hand released the start block and initiated a reach. In the front view condition, observers were able to shift their gaze directly from the start block to the target block because they could quickly determine the movement goal based on vision of the actor's hand movement. However, in the side view condition, observers were unable to determine the movement goal so quickly and adopted a different strategy. During the first hand movement—and often during the second—observers shifted their gaze from the start block to near block. From this vantage point, they assessed the actor's hand movement and shifted their gaze onto the far block if they determined that the far block was the target. The third result supporting the notion that observers use gaze proactively comes from the third experiment in which participants were asked to guess the target block after viewing only a part of the actor's hand movement, as depicted using an animated character. Comparing the results from this experiment with observers' saccade onset times suggests that observers initiated target-directed saccades about as soon as they were able to predict the hand goal.

The question arises as to why, in the side view condition, observers shifted their gaze to the near block during the actor's first hand movement rather than maintain fixation at the start block until they were certain which block was the target, as in the front view condition. One possibility is that the observers selected the near block as the default target. Because the near block was located "en route" to the far block, either this strategy would bring gaze to the goal or, when the far block was the target, closer to the goal. Moreover, it would reflect that the observer engaged gaze behavior that is similar to that of the actor, with gaze exiting the start block shortly after the hand. Another possibility is that observers were better able to discriminate between the alternative hand paths from the vantage point of the near block. However, Fig. 6*A* suggests that the segment of the hand trajectory analyzed during the fixation of the near block lies roughly between the start block and the near block.

During the second hand movement, observers should have been able to infer the goal because the actor always picked up one block and then the other. In the front view condition, observers shifted their gaze from the start block to the target block about 90 ms earlier during the second hand movement compared with the first and presumably did not rely on visual motion cues from the hand to determine the goal. This clearly shows that observers exploited their cognitive knowledge of the task to generate proactive eye movements similar to those observed in actors (Flanagan and Johansson 2003; Johansson et al. 2001). That is, even though they know the goal, observers still fixated the start block initially and only shifted their gaze to the target block at around the time, or shortly after, the hand started to move away from the start block. In the side view condition, when the second hand movement was directed to the far block, observers shifted their gaze directly to the goal in about one half the trials. In these trials, gaze was shifted, on average, 84 ms after the onset of hand movement and therefore most likely before the goal could be reliably determined based on vision of the hand movement. Thus once again, observers exploited cognitive knowledge in these trials. In the remaining trials in which the second hand movement was directed to the far block, observers made an initial saccade to the near block. Both this initial saccade and the second saccade to the far block occurred earlier than during first hand movements to the far target, suggesting some influence of cognitive knowledge. It is possible that, in these double saccade second hand movement cases, observers treated the near object as an obstacle. In our previous work on gaze behavior in object manipulation tasks, we showed that actors often fixate obstacles in the path of the hand (Johansson et al. 2001). When an obstacle is fixated, gaze arrives at the obstacle ahead of the hand and departs at around the time that the hand (or block in hand) is closest to the obstacle. We have also shown that observers often fixate obstacles they watch an actor move around (Flanagan and Johansson 1999).

We previously argued (Flanagan and Johansson 2003) that the similarity of actors' and observers' gaze behavior when performing and observing familiar manual tasks, respectively, provides direct support for the direct matching hypothesis put forward by Rizzolatti et al. (2001). This hypothesis holds that observers of action implement covert action plans that, in real time, match those executed by the actor. We would argue that the present results are entirely consistent with this view. Although the gaze behavior of our observers necessarily differed from that of an actor in a number of ways, striking similarities remained. In particular, observers still used proactive eye movements to direct gaze to forthcoming targets or action goals and thus obtained information about the goal that would be beneficial in guiding and controlling manual action. Clearly, many of the details of the action plan implemented by the observer cannot be specified at the same time as the actor specifies them. However, observers seem to specify these details as quickly as possible, based on vision of the actor's movement, and generate task-specific eye movements that support the action plan.

In this study, we did not allow observers in *experiments 1* and *2* to view the actor's eyes. Previous studies have shown that observing another person's gaze is important in establishing joint attention, that is, the observer's attention shifts rapidly and automatically to the direction of the other person's gaze (Friesen et al. 2004), and that humans have a tendency to imitate others' gaze direction (Ricciardelli et al. 2002). This raises the question as to whether, in our experiments, observers would have fixated the actor's eyes if given the opportunity. We believe this is unlikely because, in our previous study on gaze behavior in action observation (Flanagan and Johansson 2003), the actor's eyes were visible but observers essentially never fixated the actor's face. Of course, it is possible that when the actor's goal cannot be predicted in advance, observers might exploit gaze direction. However, given that observers fixated the start block to determine when the task started, it is not clear whether it would be advantageous to shift gaze to the eyes.

In our previous work on block stacking, we showed that observers' gaze behavior is reactive, rather than proactive, when an unseen hand manipulates the blocks (Flanagan and Johansson 2003). Specifically, observers tended to track the moving blocks with their gaze. We suggested that observers did not use proactive eye movements because they could not implement representations of the manual task that called for such eye movements. However, it could be argued that proactive gaze behavior was not observed because observers were unable to predict when the block would start and stop moving and that the unpredictable motion of the blocks captured gaze. If that were the case, one might expect to see gaze tracking in this task. That is, one might predict that observers would tend to track the hand—at least during the first hand movements—and use gaze reactively. However, gaze tracking of the actor's hand occurred only in 1 of the 19 participants recruited for *experiments 1* and *2*.

A fundamental question is whether the putative activation of action plans or schema, in the observer, is required for understanding action. Presumably, observers would appreciate something about viewed actions even if they did not activate action schemas. For example, although observers may not activate action schemas when observing blocks being moved by an unseen hand (Flanagan and Johansson 2003), they would presumably be able to describe what they viewed. Thus the question arises as to why observers might activate action plans. To answer this question, one first must consider what action plans, in the actor, involve. In controlling object manipulation tasks, the sensorimotor system generates specific predictions about the sensory feedback it will receive related to discrete mechanical events (e.g., object lift-off) and compares these predictions with actual sensory feedback (Johansson and Westling 1984, 1987, 1988). If a mismatch occurs, rapid corrections are made. Thus action plans in manipulation tasks include sensory predictions that are tightly linked to motor commands. Moreover, we have recently suggested that an important role of gaze in manipulation tasks is to capture key mechanical events so that visual information about these events can be compared with predictions and correlated with other sensory information (e.g., tactile, auditory, and proprioceptive) marking these events (Johansson et al. 2001). Thus observers, like actors, may activate action plans or schema so that they can generate and evaluate predictions about task-related events. If so, observers, again like actors, should be sensitive to mismatches between predicted and actual (i.e., observed) events. For example, if an observer were to watch an actor lift an object that was heavier than expected (by the actor and observer), they would notice that the object does not lift off at the expected time. This information could be used by the observer to learn about the environment (i.e., that the object is heavier than expected).

In summary, we showed that, even when observers do not know in advance the goal of the actor's movement, they nevertheless engage gaze in a proactive fashion by fixating the goal ahead of the actor's hand. To do so, observers analyze the actor's hand trajectory and shift gaze to the target as soon as they are certain where the hand is going. When the goal of the actor's movement can be deduced based on the rules of the task (i.e., during the 2nd hand movement), observers exploit this information and shift their gaze to the goal earlier. Thus observers use both kinematic cues and knowledge of the task to produce proactive eye movements. Regardless of the information being used, we believe that observers implement representations or action plans of the manual task being performed by the actor. However, the details of these plans differ depending on whether observers know the goal of the movement in advance or must determine the goal from kinematic cues. It follows that the representations or action plans engaged by observers (especially during the 1st movement) need not precisely match those of the actor in terms of specific details and timing.

ACKNOWLEDGMENTS

We thank M. York and S. Hickman for technical support.

GRANTS

This work was supported by the Canadian Institutes of Health Research, the Human Frontier Science Program, the Swedish Research Council, and the 6th Framework Program of the European Union.

REFERENCES

Avikainen S, Forss N, and Hari R. Modulated activation of the human SI and SII cortices during observation of hand actions. *Neuroimage* 15: 640–646, 2002.

Caspi A, Beutter BR, and Eckstein MP. The time course of visual information accrual guiding eye movement decisions. *Proc Natl Acad Sci USA* 101: 13086–13090, 2004.

Cisek P and Kalaska JF. Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* 431: 993–996, 2004.

Decety J, Grezes J, Costes N, Perani D, Jeannerod M, Procyk E, Grassi F, and Fazio F. Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain* 120: 1763–1777, 1997.

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, and Rizzolatti G. Understanding motor events: a neurophysiological study. *Exp Brain Res* 91: 176–180, 1992.

Fadiga L, Fogassi L, Pavesi G, and Rizzolatti G. Motor facilitation during action observation: a magnetic stimulation study. *J Neurophysiol* 73: 2608–2611, 1995.

Flanagan JR and Johansson RS. Gaze-hand coordination subserving motion planning in object manipulation. *Soc Neurosci Abstr* 29: 50.18, 1999.

Flanagan JR and Johansson RS. Action plans used in action observation. *Nature* 424: 769–771, 2003.

Friesen CK, Ristic J, and Kingstone A. Attentional effects of counterpredictive gaze and arrow cues. *J Exp Psychol Hum Percept Perform* 30: 319–329, 2004.

Gallese V, Fadiga L, Fogassi L, and Rizzolatti G. Action recognition in the premotor cortex. *Brain* 119: 593–609, 1996.

Gangitano M, Mottaghy FM, and Pascual-Leone A. Phase-specific modulation of cortical motor output during movement observation. *Neuroreport* 12: 1489–1492, 2001.

Grafton ST, Arbib MA, Fadiga L, and Rizzolatti G. Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Exp Brain Res* 112: 103–111, 1996.

Hari R, Forss N, Avikainen S, Kirveskari E, Salenius S, and Rizzolatti G. Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc Natl Acad Sci USA* 95: 15061–15065, 1998.

Hayhoe M and Ballard D. Eye movements in natural behavior. *Trends Cogn Sci* 9: 188–194, 2005.

Iacoboni M, Koski LM, Brass M, Bekkering H, Woods RP, Dubeau MC, Mazziotta JC, and Rizzolatti G. Reafferent copies of imitated actions in the right superior temporal cortex. *Proc Natl Acad Sci USA* 98: 13995–13999, 2001.

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, and Rizzolatti G. Cortical mechanisms of human imitation. *Science* 286: 2526–2528, 1999.

Jeannerod M, Arbib MA, Rizzolatti G, and Sakata H. Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci* 18: 314–320, 1995.

Johansson RS and Westling G. Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Exp Brain Res* 56: 550–564, 1984.

Johansson RS and Westling G. Signals in tactile afferents from the fingers eliciting adaptive motor responses during precisions grip. *Exp Brain Res* 66: 141–154, 1987.

Johansson RS and Westling G. Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Exp Brain Res* 7: 59–71, 1988.

Johansson RS, Westling G, Backstrom A, and Flanagan JR. Eye-hand coordination in object manipulation. *J Neurosci* 21: 6917–6932, 2001.

Johnson SH, Rotte M, Grafton ST, Hinrichs H, Gazzaniga MS, and Heinze HJ. Selective activation of a parietofrontal circuit during implicitly imagined prehension. *Neuroimage* 17: 1693–1704, 2002.

Kohler E, Keysers C, Umilta MA, Fogassi L, Gallese V, and Rizzolatti G. Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297: 846–848, 2002.

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

Land MF. Motion and vision: why animals move their eyes. *J Comp Physiol* 185: 341–352, 1999.

Land MF and Furneaux S. The knowledge base of the oculomotor system. *Philos Trans R Soc Lond B Biol Sci* 352: 1231–1239, 1997.

Liberman AM and Whalen DH. On the relation of speech to language. *Trends Cognit Sci* 4: 187–196, 2000.

Lisberger SG, Fuchs AF, King WM, and Evinger LC. Effect of mean reaction time on saccadic responses to two step stimuli with horizontal and vertical components. *Vision Res* 15: 1021–1025, 1975.

Nishitani N and Hari R. Temporal dynamics of cortical representation for action. *Proc Natl Acad Sci USA* 97: 913–918, 2000.

Pare M and Hanes DP. Controlled movement processing: superior colliculus activity associated with countermanded saccades. *J Neurosci* 23: 6480–6489, 2003.

Ramnani N and Miall RC. A system in the human brain for predicting the actions of others. *Nat Neurosci* 7: 85–90, 2004.

Ricciardelli P, Bricolo E, Aglioti SM, and Chelazzi L. My eyes want to look where your eyes are looking: exploring the tendency to imitate another individual's gaze. *Neuroreport* 13: 2259–2264, 2002.

Rizzolatti G, Fadiga L, Matelli M, Bettinardi V, Paulesu E, Perani D, and Fazio F. Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp Brain Res* 111: 246–252, 1996.

Rizzolatti G, Fogassi L, and Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* 2: 661–670, 2001.

Strafella AP and Paus T. Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. *Neuroreport* 11: 2289–2292, 2000.

Umiltà MA, Kohler E, Gallese V, Fogassi L, Fadiga L, Keysers C, and Rizzolatti G. I know what you are doing. a neurophysiological study. *Neuron* 31: 155–165, 2001.