

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Grasping Others' Movements: Rapid Discrimination of Object Size From Observed Hand Movements

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1560080> since 2016-04-19T17:53:11Z

Published version:

DOI:10.1037/xhp0000169

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:

Questa è la versione dell'autore dell'opera:

Journal of Experimental Psychology: Human Perception and Performance

Advanceonline publication. 2016 Apr 14

DOI: [10.1037/xhp0000169](https://doi.org/10.1037/xhp0000169)

The definitive version is available at:

La versione definitiva è disponibile alla URL:

<http://dx.doi.org/10.1037/xhp0000169>

<http://dx.doi.org/10.1037/xhp0000169.supp>

Grasping others' movements: rapid discrimination of object size from observed hand movements

Caterina Ansuini^(a), Andrea Cavallo^(b), Atesh Koul^(a), Alessandro D'Ausilio^(a), Laura Taverna^(a), Cristina Becchio^{(a,b)*}

a) Department of Robotics, Brain and Cognitive Sciences, Fondazione Istituto Italiano di Tecnologia

b) Department of Psychology University of Torino

*Correspondence concerning this article should be addressed to Cristina Becchio, RBCS, Fondazione Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genova, Italy. E-mail: cristina.becchio@unito.it

Abstract: During reach-to-grasp movements, the hand is gradually molded to conform to the size and shape of the object to be grasped. Yet the ability to glean information about object properties by observing grasping movements is poorly understood. In this study, we capitalized on the effect of object size to investigate the ability to discriminate the size of an invisible object from movement kinematics. The study consisted of 2 phases. In the first action execution phase, to assess grip scaling, we recorded and analyzed reach-to-grasp movements performed toward differently sized objects. In the second action observation phase, video clips of the corresponding movements were presented to participants in a two-alternative forced-choice task. To probe discrimination performance over time, videos were edited to provide selective vision of different periods from 2 viewpoints. Separate analyses were conducted to determine how the participants' ability to discriminate between stimulus alternatives (Type I sensitivity) and their metacognitive ability to discriminate between correct and incorrect responses (Type II sensitivity) varied over time and viewpoint. We found that as early as 80 ms after movement onset, participants were able to discriminate object size from the observation of grasping movements delivered from the lateral viewpoint. For both viewpoints, information pickup closely matched the evolution of the hand's kinematics, reaching an almost perfect performance well before the fingers made contact with the object (60% of movement duration). These findings suggest that observers are able to decode object size from kinematic sources specified early on in the movement.

Keywords: action prediction, reach-to-grasp, kinematics, time course, object size.

Grasping movements are necessarily constrained by the intrinsic properties of the object, such as its size and shape. Interestingly, although some kinematic parameters, such as wrist height, conform only gradually to the geometry of the object, others are fully specified early on in the movement (Ansuini et al., 2015; Danckert, Sharif, Haffenden, Schiff, & Goodale, 2002; Ganel, Freud, Chajut, & Algom, 2012; Glover & Dixon, 2002; Holmes, Mulla, Binsted, & Heath, 2011; Santello & Soechting, 1998). For example, wrist velocity discriminates small and large objects at 10% of movement duration. Similarly, grip aperture, which is scaled in flight to conform to the size and shape of the target (Jeannerod, 1984), already discriminates object size at 10% of movement duration (e.g., Ansuini et al., 2015; Holmes et al., 2011). This raises the question of whether, during the observation of object-directed actions, observers may use kinematic information to anticipate the intrinsic properties of objects well before the fingers make contact with the object surface.

In rich and changing environments, our actions and those of others are often partly obscured from view. Yet even if the final part of the action is not visible, observers are in many cases still able to predict the goal object of an actor's reach (e.g., Southgate, Johnson, El Karoui, & Csibra, 2010; Umiltà et al., 2001). Imagine observing an agent walking toward you and picking up something on the ground. What is he up to do? Does he pick up a rock or a nut? Quick discrimination of the object goal would be highly advantageous for understanding others' intention and reacting appropriately.

In this study, we examined the possibility that advance information gleaned from the observation of the earliest phases of a grasping movement can be used to discriminate the size of an object occluded from view. Converging evidence indicates that observers are able to extract object size information from the observation of grasping movements (Ambrosini, Costantini, & Sinigaglia, 2011; Campanella, Sandini, & Morrone, 2011). For example, visual kinematics contributes to calibration of the visual sense of size when grasping movements are displayed from an egocentric perspective (Campanella et al., 2011). How rapidly object information is gleaned from the observation of others' movements, however, is unclear.

To determine the timing of object information pickup, we need to know (a) "when" diverse aspects of the movement are specified, that is, when kinematics

discriminate correctly the properties of the target; and (b) the time courses at which diverse kinematics parameters evolve. Unfortunately, in action observation studies, kinematics data are most often estimated based on video displays rather than quantitatively assessed within the same study (e.g., Ambrosini et al., 2011; Ambrosini, Pezzulo, & Costantini, 2015; Thioux & Keysers, 2015). Ambrosini et al. (2011), for example, recorded eye movements while participants observed an actor reaching for and grasping either a small or a large object. In a control condition, the actor merely reached for and touched one of the two objects with his close fist, that is, without preshaping to the target features. Results showed higher accuracy and earlier saccadic movements when participants observed an actual grasp than when they observed a mere touch. For small objects, in particular, the gaze proactively reached the object to be grasped about 200 ms earlier in the preshape condition than in the no-shape condition. This clearly indicates that, as apparent from the video, hand preshaping provided observers with enough information to proactively saccade toward the object to be grasped. “When” this information was specified in the observed movement kinematics, however, cannot be quantified from video display. This limits inferences about how quickly observers pick up the detail of the observed manual action in their own eye motor program (see also Rotman, Troje, Johansson, & Flanagan, 2006).

Along similar lines, lack of execution data makes it difficult to determine the time at which we would expect observers to be able to judge object properties. Ambrosini et al. (2015), for example, using an action observation paradigm requiring integration of multiple conflicting sources of evidence (gaze direction, arm trajectory, and hand preshaping), report that explicit predictions of the action target disregarded kinematic information during the early phases of a grasping movement (at 20% and 30% of movement duration) in favor of gaze information. The authors assume that this is because, at these earlier time periods, hand preshaping did not transmit information about the size of the object to be grasped. However, it is also possible that despite the availability of size information in the presented stimuli, observers did not have the necessary sensitivity to pick it up (Runeson & Frykholm, 1983). Again, lacking an appropriate characterization of the observed movement kinematics, it is impossible to determine which of these factors — unavailability of stimulus information or lack of attunement — is responsible for the reported findings.

To overcome these limitations, in the present study, we first quantified the effect of object size over time by measuring kinematics during execution of grasping movements toward differently sized objects (action execution phase). Next, using videos of the same grasping actions, we probed observers' ability to discriminate object size from the observation of grasping movements (action observation phase). Participants viewed reach-to-grasp movements toward an occluded object. To determine the timing of advance information pickup, that is, how rapidly observers were able to predict object size, we manipulated information availability by presenting reach-to-grasp movements under different levels of temporal occlusion (Abernethy, Zawi, & Jackson, 2008).

Additionally, we introduced a viewpoint manipulation to investigate whether viewpoint affects discriminatory performance. Previous studies investigating action observation most often displayed actions from a third-person viewpoint, consistent with the detached observation of a movement involving someone else (Becchio, Sartori, & Castiello, 2010). During social interaction, however, others' actions may be perceived more commonly from a second person viewpoint, consistent with the observation of a movement directed toward oneself (Schilbach, 2010; Schilbach et al., 2013). To identify commonalities and differences between second- and third-person viewpoint in information pickup, we presented grasping movements from a frontal view (consistent with second-person viewpoint) and from a lateral view (consistent with third-person viewpoint).

Separate analyses were conducted to determine the participants' ability to predict object size (i.e., to discriminate between stimulus alternatives; Type I sensitivity) and their metacognitive ability to monitor their own predictive performance (i.e., to discriminate between their own correct and incorrect predictive judgments; Type II sensitivity). Metacognition varies across individuals (Fleming, Weil, Nagy, Dolan, & Rees, 2010) and can be dissociated from task performance through pharmacological (Izaute & Bacon, 2005), neural (Rounis, Maniscalco, Rothwell, Passingham, & Lau, 2010), and task-based (Song et al., 2011) manipulations. In healthy individuals, however, metacognitive judgments are usually predictive of task performance (Fleming & Dolan, 2012). We expected, therefore, that Type II sensitivity would scale with Type I sensitivity over time and viewpoint.

Method

The study consisted of an *action execution* phase, aimed at measuring movement kinematics and quantifying the availability of stimulus information over time, and a subsequent action observation phase, aimed at assessing the ability to discriminate object size over time from two different perspectives. A detailed description of the methods and results of the action execution phase is provided in a previous publication from our laboratory (Ansuini et al., 2015). Here, we only briefly describe the most pertinent details.

Action Execution Phase

To create the stimulus material to be used in the action observation phase, we filmed agents performing reach-to-grasp movements toward differently sized objects. Specifically, participants ($N = 15$) were asked to reach toward, grasp, lift, and move an object to an area located 50 cm to the left of the object's initial position (i.e., at about 43 cm in front of agent's midline). The target object was either a grapefruit (diameter = about 10 cm; weight = about 354 g; referred to as "large object") or a hazelnut (diameter = about 1.5 cm; weight = about 2 g; referred to as "small object"). Participants were requested to grasp the objects at a natural speed using their right hand. Each agent performed a total of 60 trials in six separate blocks of 10 trials (three blocks of 10 trials for each of the two sizes).

Movements were recorded using a near-infrared camera motion capture system (frame rate = 100 Hz; Vicon System) and filmed from a frontal and a lateral viewpoint using two digital video cameras (Sony Handy Cam 3-D and Canon Alegria, 25 frames/s). In the frontal viewpoint, the video camera was located in front of the agent, at about 120 cm from the hand start position. In the lateral viewpoint, the video camera was placed at about 120 cm from the hand start position with the camera view angle directed perpendicularly to the agent's midline. Video camera position and arrangement were kept constant for the entire duration of the study. From both viewpoints, the hand was in full view from the beginning up to the end of the movement.

To assess the availability of stimulus information over time, the following kinematic variables were computed:

1. *Movement duration* (ms), defined as the time interval between reach onset (i.e., the first time point at which the wrist velocity crossed a 20 mm/s threshold and remained above it for longer than 100 ms) and offset (the first time at which the wrist velocity dropped below a 20 mm/s threshold);
2. *Wrist velocity* (mm/sec), defined as the module of the velocity of the wrist marker (indicated as *rad* in Figure 1a);
3. *Wrist height* (mm), defined as the z-component of the wrist marker;
4. *Grip aperture* (mm), defined as the distance between the marker placed on the thumb tip and that placed on the tip of the index finger (see Figure 1a; *thu₄* and *ind₃*, respectively).

These variables were expressed with respect to the original frame of reference (i.e., the frame of reference of the motion capture system, termed as global frame of reference; F_{global}). In addition, to provide a better characterization of the hand joint movements, we also analyzed kinematics parameters expressed with respect to a local frame of reference centered on the hand (i.e., F_{local} , Figure 1a; for a similar method, see Carpinella, Mazzoleni, Rabuffetti, Thorsen, & Ferrarin, 2006, and Carpinella, Jonsdottir, & Ferrarin, 2011).

Within F_{local} , we computed the following variables:

1. *x-, y-, and z-thumb* (mm), defined as x-, y-, and z-coordinates for the thumb with respect to F_{local} ;
2. *x-, y-, and z-index* (mm), defined as x-, y-, and z-coordinates for the index with respect to F_{local} ;
3. *x-, y-, and z-finger plane*, defined as x-, y- and z-components of the thumb-index plane, that is, the three-dimensional components of the vector that is orthogonal to the plane.

This plane is defined as passing through *thu0*, *ind3*, and *thu4*, with components varying between +1 and -1 (please refer to Figure 1a). This variable provides information about the abduction and adduction movement of the thumb and index

finger irrespective of the effects of wrist rotation and of finger flexion and extension.

In order to allow comparison of hand postures across trials and participants, all kinematics variables were expressed with respect to normalized (from 10% to 100%, at increments of 10%) rather than absolute movement durations (780 ms and 745 ms, for small and large objects, respectively).

ANOVA techniques performed, with Object Size (small vs. large) and Time (10 levels; from 10% to 100% in 10 steps) as within-subjects factors, revealed that although some kinematic features, such as wrist height, showed little modulation until 60% of movement duration, other features were specified early on in the movement (refer to Table S1 of the online supplemental materials for details). Figure 1b shows the evolution of the grip aperture over reaching. Since the very beginning of the movement, the grip aperture was greater for large than for small objects. Classification analysis based on a support vector machine algorithm confirmed that, already at 10% and 20% of movement duration, the predictive capability of kinematic features was above chance level (63.14% and 70.49%, respectively). Accuracy rates increased as time progressed (95.81% at 50% of movement), achieving approximately 100% of accuracy at 60% of the movement (accuracy rates ranging from 99.47% at 60% to 99.97% at 100% of movement duration). Similarly, we found that *F* scores - a simplified Fisher criterion, which is suitable to estimate the discriminative power of single features as well as of group of features (Yang, Chevallier, Wiart, & Bloch, 2014) - gradually increased across time intervals (Figure 1c and d).

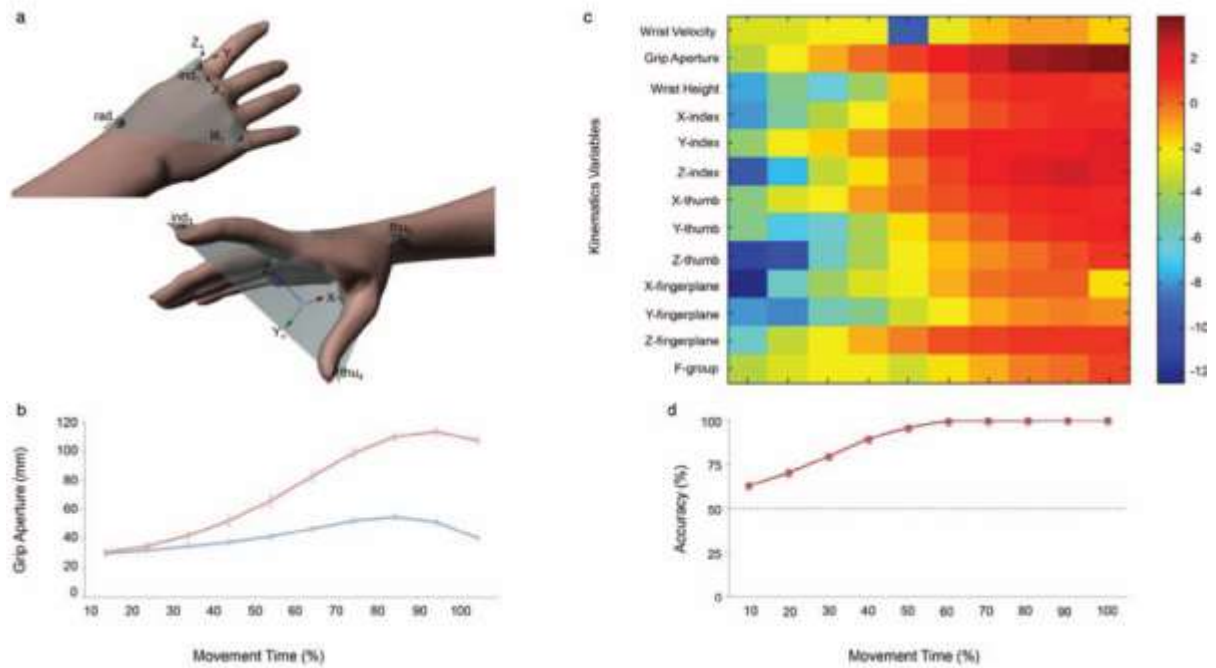


Figure 1. Schematic overview of the results for the movement recording phase. (a) Frontal and lateral views of the hand markers (height _ 0.3 cm) model used to characterize kinematics of reach-to-grasp movements toward small and large objects. To provide a better characterization of the hand joint movements, in addition to the frame of reference of the motion capture system (Fglobal), a local frame of reference centered on the hand was established (i.e., Flocal). Flocal had its origin in the *ind1* marker, whereas vectors (*ind1 – lit1*) and (*ind1 – rad*) defined the metacarpal plane of the hand (refer to the colored triangle). In this frame of reference, the *x*-axis had the direction of the vector (*ind1 – lit1*) and pointed ulnarly; the *z*-axis was normal to the metacarpal plane, pointing dorsally; and the *y*-axis was calculated as the cross-product of *z*- and *x*-axes, pointing distally. (b) Modulation of the grip aperture (i.e., the distance between the center of the markers placed on the nail side of the thumb and index finger, respectively) over time for the small (blue) and the large (red) objects. Bars represent standard error. (c) Graphical representation (heat map) of *F* scores of kinematics variables. Using kinematic parameters at different time points as features for classification analysis, this approach allowed us to classify movements depending on the size of the target object. (d) Temporal evolution of the classification accuracy when all the kinematics variables taken as a group were used as predictive features for classification algorithm. The dotted line represents a chance level of classification accuracy. Adapted from Ansuini et al. (2015).

Action Observation Phase

Having assessed the timing at which diverse kinematic parameters evolve, in the action observation phase, we set out to determine the time course of advance information pickup from the observation of grasping movements delivered from lateral and frontal viewpoints.

Participants

Nineteen participants (12 women; $M_{\text{age}} = 22$, age range = 19–29 years old) took part in the experiment. All had normal or corrected-to-normal vision and were naive as to the purpose of the experiment. All participants self-reported to be right-handed. The experimental procedures were approved by Ethical Committee of the University of Torino and were carried out in accordance with the principles of the revised Helsinki Declaration (World Medical Association General Assembly, 2008). Written consent was provided by each participant.

Materials

Stimuli selection. Movements to be included in the observation experiment were selected to capture the time course of hand kinematic scaling to object size, that is, the relationship between movement kinematics and object size. With this in mind, for each participant, we first calculated the average values of each kinematics feature for movements toward the small and the large object, respectively (please refer to Ansuini et al., 2015, for details). Then, for each participant and for each object size, we computed the Euclidian distance between the values of all kinematics parameters in each trial and the average value of the participant. Then, we selected the two trials that minimized this distance. While retaining between-subjects variability, this procedure allowed us to identify the two trials that, for each participant and each object size, better approximated the average results for the small and the large object, respectively. The final set of video stimuli consisted of 60 videos (two movements for each object size for 15 participants) representative of movements toward the small and the large objects.

Postprocessing. One hundred twenty unique video clips corresponding to the selected movements (60 frontal viewpoint: 30 small object, 30 large object; 60 lateral viewpoint: 30 small object, 30 large object) were edited using Adobe Premiere Pro CS6 (avi format, disabled audio, 25 frames/s). Lateral viewpoint stimuli displayed the right arm, forearm, and hand of the agent from a lateral viewpoint. Frontal viewpoint stimuli displayed the right arm, forearm, and hand and part of the torso (below the shoulders) of the agent from a frontal viewpoint (see the online supplemental videos). For both viewpoints, digital video editing

was used to produce spatial occlusion of the to-be-grasped object (i.e., grapefruit or hazelnut). This was obtained by superimposing a black mask (i.e., a semicircular disk) on the target object location (see Figure 2). The size and the position of this mask were kept constant across participants. Additionally, to define the timing of advance information pickup, reach-to-grasp movements were presented for eight levels of temporal occlusion, from 10% up to 80% of movement duration in steps of 10%. One potential problem with this procedure is that this can lead to spurious correlation (Whitwell & Goodale, 2013). When time varies with target size, and is used to standardize a dependent measure of spatial position (e.g., grip aperture), comparisons at equivalent points in standardized time are comparisons of spatial position at different points in raw time. A correlation between grip aperture and target size may therefore simply reflect the fact that the duration of the movement is itself correlated with target size (Whitwell & Goodale, 2013). To rule out this possibility, we compared movement durations for small and large objects for the selected videos. Movement duration for small objects (768 ± 16 ms) was not significantly different from movement duration for large objects (750 ± 20 ms), $t(29) = 1.464$, $p = .155$. This ensures that time normalization itself did not introduce a spurious relationship between dependent measures of spatial position and object size.

Each video was edited so to begin at the start of the movement. To ensure that movement sequences could be temporally attended, that is, to provide participants enough time to focus on movement start, +1, +3, or +5 static frames were randomly added at the beginning of all video clips. To equate stimulus duration, static frames were added at the end of the videos in a compensatory manner (from +1 to +6 static frames).

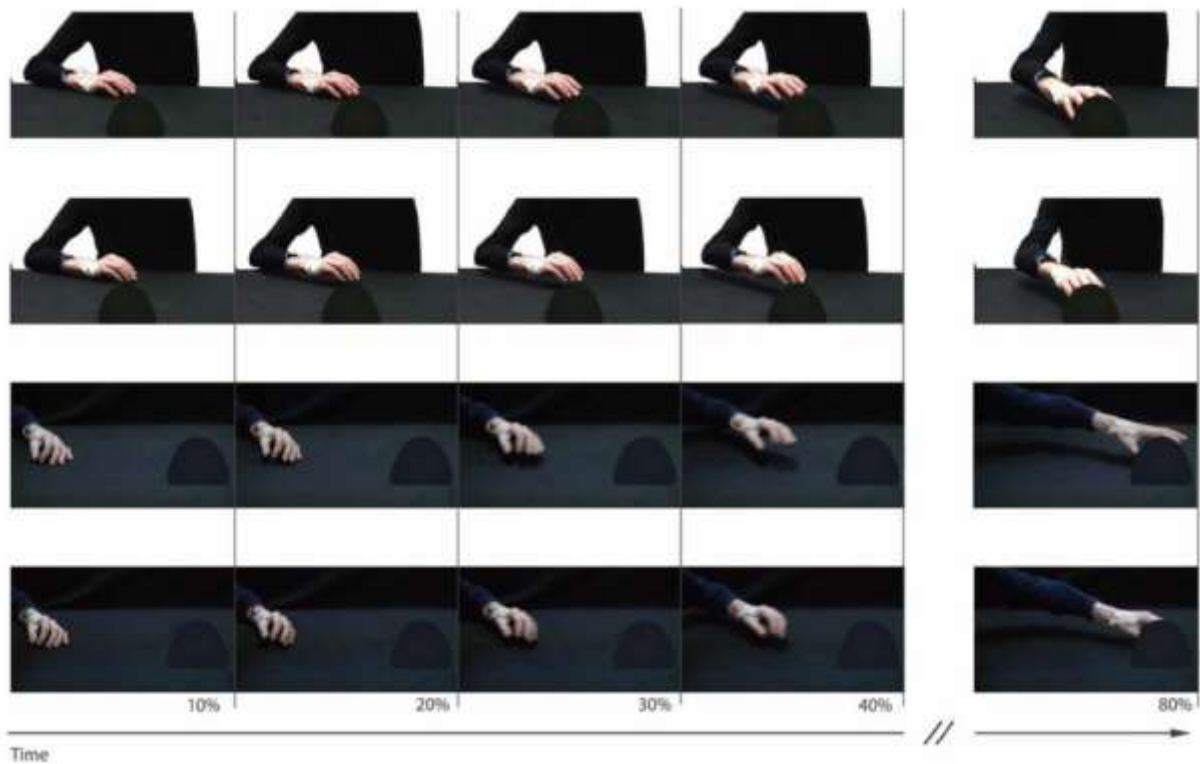


Figure 2. Snapshots of the videos used here as experimental stimuli. Reach-to-grasp movements toward the small and the large object in the frontal (upper rows) and the lateral (lower rows) viewpoint are displayed. Movements could be occluded at different time points. Here, we show the final frames corresponding to 10%, 20%, 30%, 40%, and 80% of movement duration for an exemplar video clip. The double slantwise lines indicate that three time intervals were omitted from representation (i.e., 50%, 60%, and 70% of normalized movement duration). Note that for display purposes, the frames in the lateral viewpoint are flipped over the vertical midline with respect to their original orientation. Videos are available in the online supplemental materials.

Procedure

The experiment was carried out in a dimly lit room. Participants sat in front of a 17-in. computer screen (1,280 X 800 pixels; refresh rate = 75 Hz; response rate = 8 ms) at a viewing distance of 48 cm. At movement onset, from the frontal viewpoint, the hand and the mask were presented at a visual angle of $3.9^\circ \times 8.4^\circ$ and $7.1^\circ \times 11.5^\circ$, respectively. From the lateral viewpoint, the visual angle subtended by the hand was $5.1^\circ \times 6.3^\circ$, whereas the angle subtended by the mask was $8.5^\circ \times 11.7^\circ$. Stimuli, timing, and randomization procedures were controlled using PsychToolbox script (Brainard, 1997; Pelli, 1997) running in MATLAB R2014a (MathWorks, Inc.). A two-alternative forced-choice paradigm was employed (see Figure 3). Each trial consisted of two intervals: a *target* interval (displaying a movement toward the target object, e.g., the small object) and a

nontarget interval (displaying a movement toward the nontarget object, e.g., the large object), with a 500-ms fixation cross (white against a black background) in between. Participants were asked to decide which of the two intervals displayed a movement toward the target object (e.g., “Where was the small object presented?”; Task 1; Figure 3). Responses were given by pressing with the index fingers one of two keys on a keyboard: a left key (i.e., “A”) when the target interval was presented as the first interval, and a right key (i.e., “L”) when the target interval was presented as the second interval. Participants were instructed to respond as accurately and as fast as possible. The maximum time allowed to respond was 2,000 ms. After this time had elapsed, participants were requested to rate confidence of their decision on a four-level scale by pressing a key (from 1 = *least confident* to 4 = *most confident*; Task 2; Figure 3). Participants were encouraged to use the entire confidence scale.

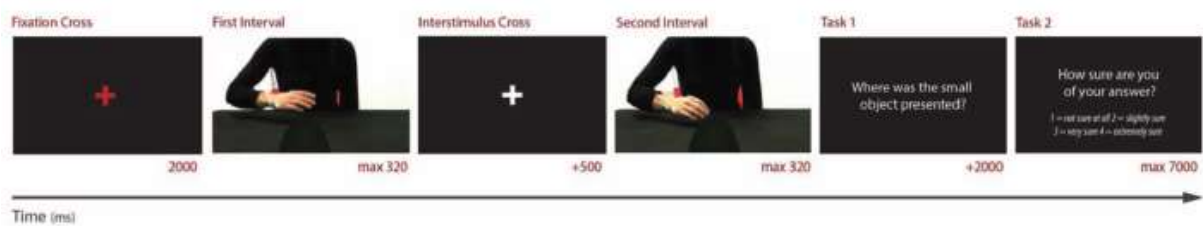


Figure 3. Schematic representation of the event sequence during a single experimental trial. Each trial began with a red fixation cross that lasted 2,000 ms. Subsequently, first and second intervals displaying reach-to-grasp movements were shown, with a 500-ms fixation in between. Note that the snapshots used here as examples of first and second intervals correspond to the last frame of videos displaying the 30% of reach-to-grasp movement toward the large and the small object, respectively. Participants were then asked to decide which of the two intervals displayed a movement toward the object signaled as “Target” (i.e., the small object in the reported exemplum) and then to rate the confidence of their decision on a four-level scale.

Stimuli displaying grasping movements from the lateral and frontal viewpoints were administered to participants in separate sessions on two consecutive days. In each session, participants completed eight blocks of 30 trials: four consecutive blocks in which the target interval contained a movement toward the small object, and four consecutive blocks in which the target interval contained a movement toward the large object. At the beginning of each block, participants were instructed about which object to consider as target, with the target interval appearing randomly either as the first interval or as the second interval. Each block included at least three repetitions of each time point. Half of the participants

completed the blocks in which the target was the small object first. Participants completed a total of 240 trials for each viewpoint, with eight repetitions of each time point. Feedback was provided at end of each block to encourage participants to maintain accurate responding (e.g., “Your mean accuracy in this block was 75%”).

To familiarize participants with the type of stimuli and the task, at the beginning of each experiment, we administered 10 practice trials. The practice trials were randomly selected from the main experimental videos. A performance feedback was provided at the end of the practice session. On each day, the experimental session lasted about 40 min. The order of the sessions was counterbalanced across participants.

Dependent Measures

No-response trials (less than 1%) were not analyzed. To compare participants’ discrimination performance from the frontal and the lateral viewpoints under the different levels of temporal occlusion, RTs, and detection sensitivity were analyzed on Task 1. Response times (RTs) were only analyzed for correct responses. The proportion of hits (arbitrarily defined as first interval responses when the target was in the first interval) and false alarms (arbitrarily defined as first interval responses when the target was in the second interval) was calculated for each participant for each level of temporal occlusion, separately for stimuli delivered from a lateral and frontal viewpoints. From hit and false alarm data, we then estimated d' , which provides a criterion-independent measure of detection performance (Type I sensitivity; Green & Swets, 1966; Macmillan & Creelman, 1991). In addition, confidence ratings were used to determine points on the receiver operating characteristic (ROC) curve, which plots the hit rate as a function of the false alarm rate. Because each response (first interval, second interval) had four ratings associated with it, there were eight possible responses for each trial (graded from the most confident first interval response to the most confident second interval response), resulting in seven points on the ROC curve. The area under the curve (AUC) is a measure of sensitivity unaffected by response bias and can be interpreted as the proportion of times participants would correctly identify the target, if target and nontarget were presented simultaneously (for a similar

approach, see Azzopardi & Cowey, 1997; Ricci & Chatterjee, 2004; Tamietto et al., 2015; Van den Stock et al., 2014). A diagonal ROC curve, connecting the lower left to upper right corners, which coincides with an AUC of 0.5, shows a chance level classification score. This curve would be interpreted as the observer having 50% probability of correctly discriminating movements toward small and large objects. On the contrary, a ROC curve on the left upper bound of the diagonal encompassing the entire unit square (so that the AUC is 1) indicates a perfect positive prediction with no false positives and an optimal decoding score. This curve would be interpreted as the observer having a 100% probability of correctly discriminating movements toward small and large objects.

To characterize how well confidence ratings discriminated between participants' own correct and incorrect responses, applying a similar set of principles, we also calculated *meta-d'*, a measure of Type II sensitivity (Maniscalco & Lau, 2012, 2014). Just as *d'* measures the observer's ability to discriminate between target and nontarget stimulus, *meta-d'* measures the observer's ability to discriminate between correct and incorrect stimulus classification. *Meta-d'* values are calculated by estimating Type I parameters that best fit the actual Type II data (generated by participants' confidence ratings), and represent the amount of signal available to a subject to perform Type II task. We computed *meta-d'* values by fitting each participant's behavioral data using MATLAB code provided by Maniscalco and Lau (2012).¹

Statistical Analyses

Separate analyses were conducted to evaluate participants' Type I and Type II sensitivity.

To evaluate the impact of viewpoint on the time course of size discrimination, *d'*, AUC values, and RTs were submitted to separate repeated measures ANOVAs, with Time (eight levels, from 10% to 80% in 10% intervals) and Viewpoint (two levels: frontal vs. lateral) as within-subjects factors. When a significant effect was found, post hoc comparisons were conducted applying Bonferroni's correction

¹ The MATLAB code is available at <http://www.columbia.edu/~bsm2105/type2sdt/>.

(alpha level = .05). *T* tests were also conducted on AUC values at each time interval (from 10% to 80%), for each viewpoint (i.e., frontal and lateral), to establish whether classification was above the chance/guessing level of 0.5. To control the risk of Type I Error, a Bonferroni's adjustment was used to reset the alpha level to .003 based on the number of *t*-test comparisons completed (.05/16 comparisons).

Additionally, to assess metacognitive ability to discriminate between one's own correct and incorrect responses, *meta-d'* values were also submitted to a repeated measures ANOVA, with Time (eight levels: from 10% to 80% in 10% interval) and Viewpoint (two levels: Frontal vs. Lateral) as within-subjects factors.

Results

Discrimination Between Stimulus Alternatives

Table 1
Behavioral Indices Obtained for Action Observation Phase for Movements Displayed From the Frontal and the Lateral View at Each Time Interval (From 10% to 80% of Movement Duration)

Time (%)	View	Dependent measure			
		<i>d'</i>	AUC	RT	IES
10	Frontal	<i>M</i> (\pm SE) = .069 \pm .101 95% CI [-.144, .282]	<i>M</i> (\pm SE) = .532 \pm .023 95% CI [.484, .579]	<i>M</i> (\pm SE) = 819 \pm 58 95% CI [697, 941]	<i>M</i> (\pm SE) = 1,642 \pm 118 95% CI [1,394, 1,891]
	Lateral	<i>M</i> (\pm SE) = .431 \pm .103 95% CI [.215, .647]	<i>M</i> (\pm SE) = .602 \pm .027 95% CI [.545, .660]	<i>M</i> (\pm SE) = 933 \pm 66 95% CI [794, 1,072]	<i>M</i> (\pm SE) = 1,656 \pm 134 95% CI [1,375, 1,937]
20	Frontal	<i>M</i> (\pm SE) = .421 \pm .107 95% CI [.197, .646]	<i>M</i> (\pm SE) = .616 \pm .021 95% CI [.571, .661]	<i>M</i> (\pm SE) = 821 \pm 53 95% CI [710, 931]	<i>M</i> (\pm SE) = 1,456 \pm 95 95% CI [1,256, 1,656]
	Lateral	<i>M</i> (\pm SE) = .874 \pm .167 95% CI [.522, 1.226]	<i>M</i> (\pm SE) = .697 \pm .026 95% CI [.642, .751]	<i>M</i> (\pm SE) = 921 \pm 61 95% CI [794, 1,049]	<i>M</i> (\pm SE) = 1,462 \pm 104 95% CI [1,244, 1,679]
30	Frontal	<i>M</i> (\pm SE) = 1.848 \pm .107 95% CI [1.623, 2.073]	<i>M</i> (\pm SE) = .886 \pm .015 95% CI [.854, .919]	<i>M</i> (\pm SE) = 701 \pm 38 95% CI [621, 781]	<i>M</i> (\pm SE) = 881 \pm 49 95% CI [777, 985]
	Lateral	<i>M</i> (\pm SE) = 1.675 \pm .156 95% CI [1.348, 2.003]	<i>M</i> (\pm SE) = .861 \pm .017 95% CI [.826, .896]	<i>M</i> (\pm SE) = 796 \pm 51 95% CI [688, 904]	<i>M</i> (\pm SE) = 1,016 \pm 65 95% CI [879, 1,152]
40	Frontal	<i>M</i> (\pm SE) = 3.140 \pm .110 95% CI [2.910, 3.370]	<i>M</i> (\pm SE) = .972 \pm .011 95% CI [.949, .995]	<i>M</i> (\pm SE) = 606 \pm 41 95% CI [521, 692]	<i>M</i> (\pm SE) = 643 \pm 42 95% CI [554, 732]
	Lateral	<i>M</i> (\pm SE) = 2.702 \pm .140 95% CI [2.407, 2.996]	<i>M</i> (\pm SE) = .958 \pm .013 95% CI [.929, .986]	<i>M</i> (\pm SE) = 679 \pm 43 95% CI [588, 770]	<i>M</i> (\pm SE) = 750 \pm 45 95% CI [656, 844]
50	Frontal	<i>M</i> (\pm SE) = 3.485 \pm .081 95% CI [3.315, 3.656]	<i>M</i> (\pm SE) = .988 \pm .007 95% CI [.973, 1.000]	<i>M</i> (\pm SE) = 482 \pm 33 95% CI [413, 552]	<i>M</i> (\pm SE) = 496 \pm 35 95% CI [422, 570]
	Lateral	<i>M</i> (\pm SE) = 3.149 \pm .110 95% CI [2.919, 3.380]	<i>M</i> (\pm SE) = .981 \pm .008 95% CI [.963, .999]	<i>M</i> (\pm SE) = 569 \pm 35 95% CI [494, 643]	<i>M</i> (\pm SE) = 603 \pm 39 95% CI [521, 686]
60	Frontal	<i>M</i> (\pm SE) = 3.405 \pm .091 95% CI [3.214, 3.596]	<i>M</i> (\pm SE) = .981 \pm .010 95% CI [.961, 1.000]	<i>M</i> (\pm SE) = 434 \pm 32 95% CI [368, 501]	<i>M</i> (\pm SE) = 448 \pm 32 95% CI [381, 515]
	Lateral	<i>M</i> (\pm SE) = 3.433 \pm .080 95% CI [3.264, 3.601]	<i>M</i> (\pm SE) = .987 \pm .006 95% CI [.974, .999]	<i>M</i> (\pm SE) = 505 \pm 33 95% CI [437, 574]	<i>M</i> (\pm SE) = 519 \pm 34 95% CI [447, 591]
70	Frontal	<i>M</i> (\pm SE) = 3.511 \pm .066 95% CI [3.372, 3.650]	<i>M</i> (\pm SE) = .983 \pm .009 95% CI [.965, 1.001]	<i>M</i> (\pm SE) = 430 \pm 30 95% CI [368, 492]	<i>M</i> (\pm SE) = 443 \pm 32 95% CI [377, 510]
	Lateral	<i>M</i> (\pm SE) = 3.395 \pm .091 95% CI [3.204, 3.586]	<i>M</i> (\pm SE) = .987 \pm .008 95% CI [.971, 1.004]	<i>M</i> (\pm SE) = 485 \pm 25 95% CI [432, 537]	<i>M</i> (\pm SE) = 501 \pm 27 95% CI [443, 559]
80	Frontal	<i>M</i> (\pm SE) = 3.480 \pm .058 95% CI [3.358, 3.602]	<i>M</i> (\pm SE) = .986 \pm .010 95% CI [.965, 1.007]	<i>M</i> (\pm SE) = 438 \pm 35 95% CI [364, 512]	<i>M</i> (\pm SE) = 459 \pm 38 95% CI [380, 538]
	Lateral	<i>M</i> (\pm SE) = 3.444 \pm .094 95% CI [3.246, 3.642]	<i>M</i> (\pm SE) = .982 \pm .008 95% CI [.965, .999]	<i>M</i> (\pm SE) = 478 \pm 22 95% CI [432, 523]	<i>M</i> (\pm SE) = 492 \pm 24 95% CI [443, 542]

Note. All values represent mean (*M*) \pm standard error (*SE*) of the mean and confidence intervals (*CI*s). AUC = area under the curve; RT = response time; IES = inverse efficiency score.

Table 1 summarizes the key measures for both the lateral and the frontal viewpoint at each time interval. The ANOVA on d' values yielded a significant main effect of time, $F(7, 126) = 406.143, p < .001, \eta_p^2 = .958$, with d' values increasing significantly from 10% to 50% of movement duration. From 60% to 80% of movement duration, as shown in Figure 4a and in Figure S1a of the online supplemental materials, sensitivity was close to ceiling. The effect of time was further qualified by a significant Time X Viewpoint interaction, $F(7, 126) = 5.842, p < .001, \eta_p^2 = .245$, indicating that the ability to discriminate object size across progressive occlusions was influenced by the viewpoint. Post hoc comparisons revealed that at 10% and 20% of movement duration, discrimination was significantly higher for the lateral than for the frontal viewpoint. At 40% and 50% of the movement duration, however, the pattern reversed, with d' being higher for the frontal than for the lateral viewpoint. No other differences reached statistical significance ($.162 < p < .728$). The main effect of viewpoint was not significant, $F(1, 18) = .247, p > .05, \eta_p^2 = .014$. This suggests that the specific display information available to participants from the lateral and the frontal two-dimensional stimuli varied over time. Specifically, it is plausible that at 10% and 20% of movement duration, early discriminating features such as grip aperture were partially occluded in the frontal viewpoint, resulting in an advantage for the lateral viewpoint. This advantage for the lateral viewpoint reversed about midway into the transport phase, when grip aperture was primarily expressed in terms of variations in the fronto-parallel plane, and was therefore best visible from a frontal view.

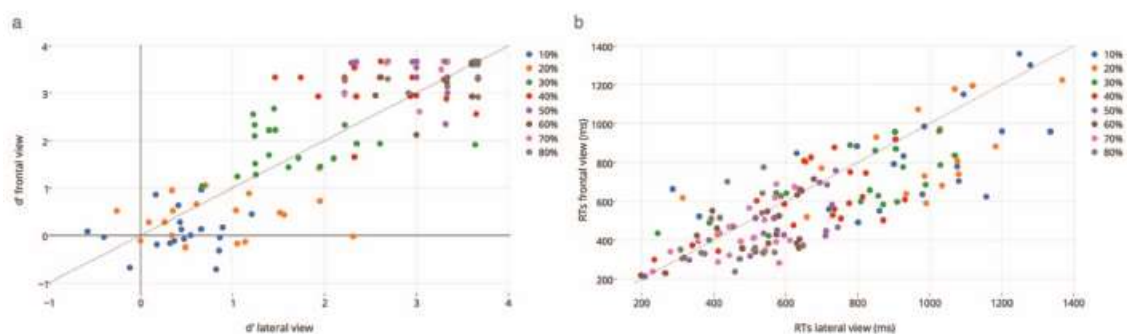


Figure 4. Results for Type I discrimination sensitivity (d') and response times (RTs) in the frontal and the lateral viewpoints. (a) d' , and (b) RTs in the frontal viewpoint (y-axis) plotted versus d' and RTs obtained in the lateral viewpoint (x-axis). Data points to all participants in the sample.

A similar pattern of results was obtained from the analysis of AUC values. Although the main effect of viewpoint was not significant, $F(1, 18) = 1.589$, $p > .05$, $\eta_p^2 = .081$, the ANOVA on AUC yielded a main effect of time, $F(7, 126) = 282.073$, $p < .001$, $\eta_p^2 = .940$, and significant interaction Time X Viewpoint, $F(7, 126) = 5.201$, $p < .001$, $\eta_p^2 = .224$. As for d' , post hoc comparisons indicated an advantage for the lateral over the frontal view at 10% and 20% of movement duration ($ps < .05$). This advantage disappeared during movement unfolding ($.278 < ps < .771$). Figure 5 depicts AUCs across successive time periods for stimuli in the lateral and frontal viewpoints. Critically, AUC values were significantly above chance threshold (0.5) at each time interval, for both viewpoints (t_{18} values ranging from 3.732 to 81.853, $ps < .003$), except for the 10% interval for stimuli delivered in the frontal view ($t_{18} = 1.391$, $p > .05$). This indicates that participants were able to predict object size from the earliest phases of the movement.

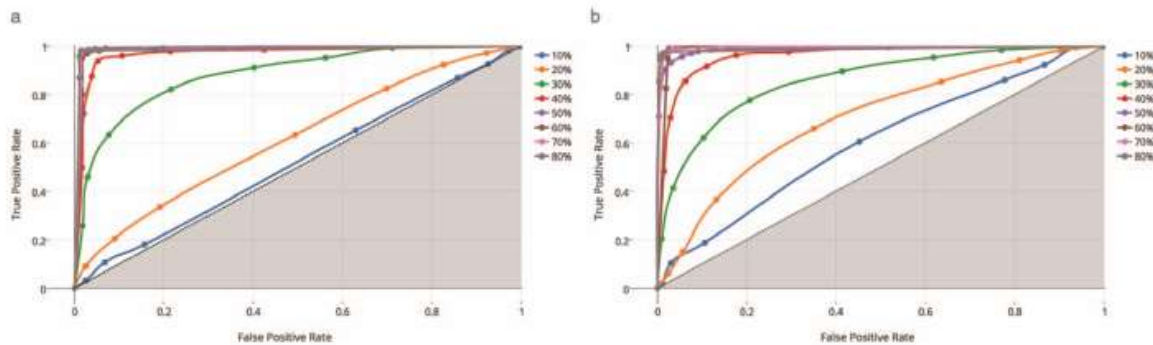


Figure 5. Results for the area under receiving operator characteristic (ROC) curve in the frontal and the lateral viewpoints. (a) 7-point ROC curves derived from participants' ratings showing probability of true positive rate (hit) versus (false-positive rate) false alarm rates for reach-to-grasp movements presented in the lateral, and (b) frontal viewpoint as a function of occlusion time point (from 10% up to 80% of reach-to-grasp movement). The discrimination ability increases as the ROC curve moves from the diagonal (dashed line corresponding to 0.5 _ random guess performance) toward the left top boundary of the graph (1.0 = perfect performance).

The ANOVA on RTs revealed a significant main effect of time, $F(7, 126) = 98.275$, $p < .001$, $\eta_p^2 = .845$, with RTs becoming increasingly faster from 20% to 50% of movement duration (see Figure 4b and Figure S1b of the online supplemental materials). The main effect of viewpoint was also significant, $F(1, 18) = 5.941$, $p < .03$, $\eta_p^2 = .248$, indicating that, regardless of time interval, participants responded faster to movements delivered in the frontal view ($M = 591.33$ ms, SE

= 37.12) than for those delivered in the lateral view ($M = 671$ ms, $SE = 39.04$). This is apparent in Figure 4b, showing that, at each time epoch, the majority of data falls below the diagonal line. No significant Time X Viewpoint interaction was found, $F(7, 126) = 1.344$, $p > .05$, $\eta_p^2 = .069$.

To control for a speed–accuracy trade-off effect, we computed an inverse efficiency score (IES) by dividing, for each condition and each participant, the mean correct RTs by the percentage of directionally correct responses, obtaining an index of overall performance (e.g., Rach, Diederich, & Colonius, 2011; Thorne, 2006; see also Townsend & Ashby, 1983). The ANOVA on IES revealed a main effect of time, $F(7, 126) = 152.090$, $p < .001$, $\eta_p^2 = .894$. In particular, it was found that from 20% up to 50% of reaching duration, the overall efficiency of the discrimination performance improved under the progressive display conditions. Neither the main effect of viewpoint, $F(1, 18) = 2.157$, $p > .05$, $\eta_p^2 = .107$, nor the interaction Time X Viewpoint, however, approached significance, $F(7, 126) = .756$, $p > .05$, $\eta_p^2 = .040$. A speed–accuracy trade-off explanation of the present results can therefore be excluded.

Discrimination Between One's Own Correct and Incorrect Responses

This analysis was devoted to determining how well participants discriminated between their correct and incorrect responses (*meta-d'*). The results of the analysis are shown in Figure 6. The ANOVA revealed a significant main effect of time, $F(7, 126) = 48.927$, $p < .001$, $\eta_p^2 = 0.731$. In line with predictions, this indicates that the ability to discriminate between correct and incorrect responses increased across progressive occlusions. Neither the main effect of viewpoint, $F(1, 18) = 0.135$, $p > .05$, $\eta_p^2 = 0.0075$, nor the Time X Viewpoint interaction, $F(7, 126) = 1.665$, $p > .05$, $\eta_p^2 = 0.085$, were significant.

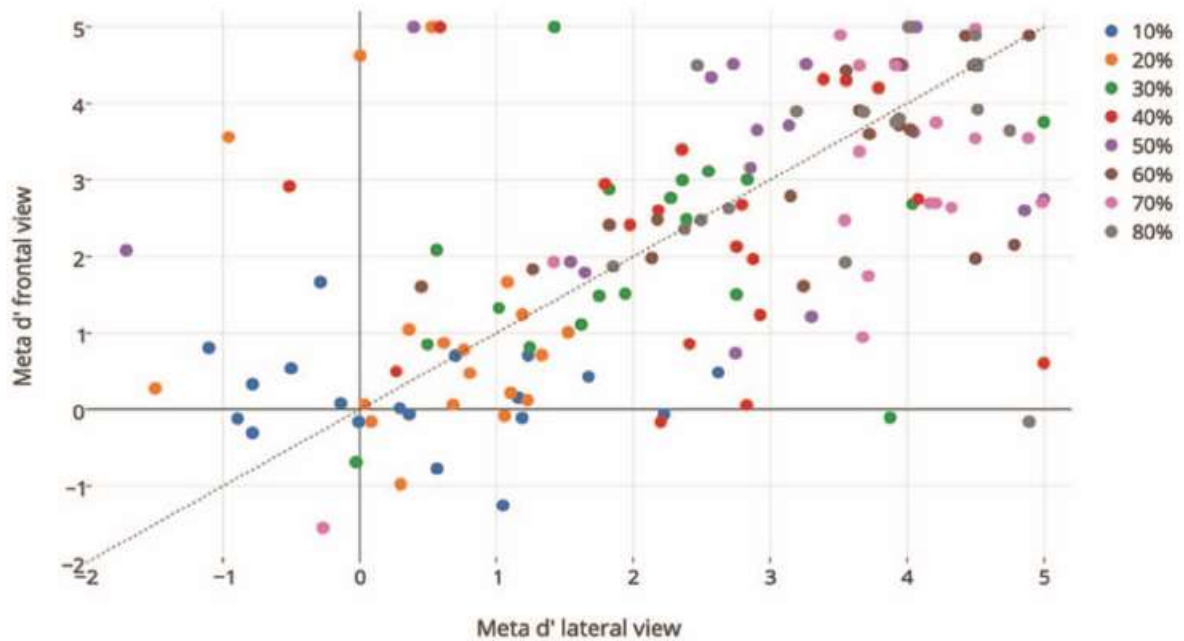


Figure 6. Results for Type II discrimination sensitivity (*meta d'*) in the frontal and lateral viewpoints. *Meta d'* in the frontal viewpoint (y-axis) plotted versus *meta d'* in the lateral viewpoint (x-axis). Data refer to all participants in the sample.

Discussion

Grasping kinematics provides information useful for judging the size of the to-be-grasped object (e.g., Ambrosini et al., 2011; Campanella et al., 2011). During reaching, however, the hand is molded only gradually to the contours of the object to be grasped, and, therefore, it remains unclear whether, during action observation, information gleaned from early events in the observed movement pattern is sufficient for visual size discrimination. In previous studies, a major barrier to the investigation of the time course of size discrimination has been the lack of execution data. Ambrosini et al. (2015), for example, report that information transmitted by hand shape about object size is only taken into account at 40% of movement duration. As the discriminatory capacity of kinematic features over time was not assessed, it is not possible to determine whether this

was because of the unavailability of size information early on in the movement or, rather, because of the lack of attunement to discriminative cues.

To quantitatively measure the timing of information pickup in absence of such confounds, in the present study, we first assessed how hand kinematics of reach-to-grasp movements toward differently sized objects evolves throughout movement duration. Next, we probed observers' ability to discriminate object size under eight different levels of temporal occlusion. We found clear evidence of advance information pickup from earlier kinematic events not assessed in previous studies. As early as approximately 80 ms after movement onset, participants were able to discriminate correctly the size of the to-be-grasped object for movements delivered in the lateral view. Inference from kinematics data suggests that this is attributable to the specification of grip aperture and wrist velocity, which, for the displayed reach-to-grasp movements, were already quite distinct at 10% of movement time. From 10% to 50% of movement duration, the capability of kinematic features to predict object size increased monotonically, achieving approximately 100% of accuracy at 60% of the movement. As shown in Figures 4a and 6, both Type I sensitivity and Type II sensitivity followed a similar temporal evolution, achieving an almost perfect level of performance at about 60% of the movement.

Although the metacognitive ability to discriminate one's own correct and incorrect responses was not influenced by viewpoint, we found a significant interaction of Time X Viewpoint for the ability to discriminate object size. This suggests that the specific display information available to participants from the lateral and the frontal view varied over time, with a reverse of the initial advantage for the lateral view about midway into the transport phase, when discriminative features such as grip aperture were most visible from a frontal viewpoint.

Regardless of this viewpoint modulation of Type I sensitivity, however, participants were always faster to respond for stimuli delivered in the frontal view than in the lateral view, with no evidence for a speed-accuracy trade-off. Several possibilities might be considered to account for this finding.

Actions directed toward oneself (displayed from a second-person perspective) actively engage the observer and invite an element of motor response not implicated by other-directed actions (consistent with stimuli delivered in a third-

person perspective; Schilbach et al., 2013). For example, Kourtis, Sebanz, and Knoblich (2010) demonstrate that the contingent negative variation, a marker of motor preparation reflecting supplementary motor area and primary motor cortex activity, was higher when participants expected an interaction partner to perform a specific action than when they anticipated that the same action would be performed by a third person they did not interact with. In a magnetoencephalography study of relevance to our stimuli, Kilner, Marchant, and Frith (2006) found that attenuation of α -oscillation during action observation, interpreted as evidence of a mirror neuron system in humans, was modulated when the actor was facing toward the participant, but not when the actor was facing away. These findings suggest that when we are personally addressed by others, the perception of their behavior relies on tight action-perception coupling with movement processing feeding into and promoting the preparation of an appropriate motor response (Schilbach et al., 2013). In line with this, faster responses for stimuli delivered in the frontal view compared with the lateral view might reflect enhanced connections between visual and motor areas, leading to higher responsiveness to observed actions displayed from a second-person perspective.

Activation in the parietal and premotor cortices is higher when the hand and finger movements are the only relevant information for inferring which object is going to be grasped (Thioux & Keysers, 2015). Interestingly, recent data show that mirror neurons in the ventral premotor cortex (vPMC) become active as quickly as 60 ms after the onset of the observed movement (Maranesi et al., 2013). The speed of this activation is substantially faster than the known temporal profile of biological motion neurons in the posterior part of superior temporal sulcus (~100 to 150 ms; Barraclough, Xiao, Baker, Oram, & Perrett, 2005). This has been taken to suggest that rapid activation of mirror neurons in the vPMC may reflect an initial guess about the specific action being perceived (Urgen & Miller, 2015), with the input for this initial guess originating in early visual cortex mediated by thalamocortical projections connecting the medial pulvinar with the vPMC (Cappe, Morel, Barone, & Rouiller, 2009). It is tempting to speculate that the exploitation of kinematic sources as early as 80 ms after movement onset may reflect this rapid activation of the mirror neuron system.

Conclusions and Future Directions

In everyday life, others' actions are often partly obscured from view. Yet even when the target object is not visible, observers are still able to understand others' actions and predict their goals. In the present study, we asked how rapidly observers are able to discriminate the size of an object occluded from view. We found that as early as 80 ms after movement onset, observers were able to predict whether the movement was direct toward a small or large object. Participants' information pickup closely matched the evolution of the hand kinematics, reaching an almost perfect performance well before the fingers made contact with the object (~60% of movement duration). These findings provide a notable demonstration of the ability to extract object information from early kinematic sources. It is a matter of future studies to investigate what kind of object representation seeing a grasp evokes in the observers. At least three research strategies seem useful in this regard. First, asking participants to watch grasping movements and to make perceptual estimates of a continuous range of object sizes may help to clarify whether and how metric properties of to-be-grasped objects are extracted from movement kinematics. Extant data indicate that the resolution of grasping execution is well below perceptually determined thresholds (e.g., size difference equal to 1%; Ganel et al., 2012). It will be important for future work to determine the resolution of grasping observation (i.e., how accurately object size can be estimated from grasping movements). Second, it will be beneficial to elucidate the robustness of this ability with respect to action variability. Showing that object size estimates track the variability across individual movements would provide compelling evidence that metric properties are extracted from the kinematics. Finally, it will also be important to uncover whether object properties can be decoded from brain regions contributing to action perception. Viewing a grasping has been shown to generate expectation of what should be seen next through a backward stream of information from the parietal and premotor nodes of the mirror neuron system to the anterior intraparietal sulcus and the lateral occipital cortex (Thioux & Keysers, 2015). Pattern recognition techniques could be used to investigate whether expected object size can be decoded from activity within this set of brain areas.

Supplemental materials: <http://dx.doi.org/10.1037/xhp0000169.supp>

References

- Abernethy, B., Zawi, K., & Jackson, R. C. (2008). Expertise and attunement to kinematic constraints. *Perception*, 37, 931–948.
- Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neurophysiology*, 106, 1437–1442.
- Ambrosini, E., Pezzulo, G., & Costantini, M. (2015). The eye in hand: Predicting others' behavior by integrating multiple sources of information. *Journal of Neurophysiology*, 113, 2271–2279.
- Ansuini, C., Cavallo, A., Koul, A., Jacono, M., Yang, Y., & Becchio, C. (2015). Predicting object size from hand kinematics: A temporal perspective. *PLoS ONE*, 10, e0120432.
- Azzopardi, P., & Cowey, A. (1997). Is blindsight like normal, nearthreshold vision? *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 94, 14190–14194.
- Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience*, 17, 377–391.
- Becchio, C., Sartori, L., & Castiello, U. (2010). Toward you: The social side of actions. *Current Directions in Psychological Science*, 19, 183–188.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Campanella, F., Sandini, G., & Morrone, M. C. (2011). Visual information gleaned by observing grasping movement in allocentric and egocentric perspectives. *Proceedings. Biological Sciences/The Royal Society*, 278, 2142–2149.
- Cappe, C., Morel, A., Barone, P., & Rouiller, E. M. (2009). The thalamocortical projection systems in primate: An anatomical support for multisensory and sensorimotor interplay. *Cerebral Cortex*, 19, 2025–2037.
- Carpinella, I., Jonsdottir, J., & Ferrarin, M. (2011). Multi-finger coordination in healthy subjects and stroke patients: A mathematical modelling approach. *Journal of Neuroengineering and Rehabilitation*, 8, 19.

- Carpinella, I., Mazzoleni, P., Rabuffetti, M., Thorsen, R., & Ferrarin, M. (2006). Experimental protocol for the kinematic analysis of the hand: Definition and repeatability. *Gait & Posture*, 23, 445–454.
- Danckert, J. A., Sharif, N., Haffenden, A. M., Schiff, K. C., & Goodale, M. A. (2002). A temporal analysis of grasping in the Ebbinghaus illusion: Planning versus online control. *Experimental Brain Research*, 144, 275–280.
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367, 1338–1349.
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, 329, 1541–1543.
- Ganel, T., Freud, E., Chajut, E., & Algom, D. (2012). Accurate visuomotor control below the perceptual threshold of size discrimination. *PLoS ONE*, 7, e36253.
- Glover, S., & Dixon, P. (2002). Dynamic effects of the Ebbinghaus illusion in grasping: Support for a planning/control model of action. *Perception & Psychophysics*, 64, 266–278.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.
- Holmes, S. A., Mulla, A., Binsted, G., & Heath, M. (2011). Visually and memory-guided grasping: Aperture shaping exhibits a time-dependent scaling to Weber's law. *Vision Research*, 51, 1941–1948.
- Izaute, M., & Bacon, E. (2005). Specific effects of an amnesic drug: Effect of lorazepam on study time allocation and on judgment of learning. *Neuropsychopharmacology*, 30, 196–204.
- Jeannerod, M. (1984). The timing of natural prehension movements. *Journal of Motor Behavior*, 16, 235–254.
- Kilner, J. M., Marchant, J. L., & Frith, C. D. (2006). Modulation of the mirror system by social relevance. *Social Cognitive and Affective Neuroscience*, 1, 143–148.
- Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: Social interaction modulates action simulation. *Biology Letters*, 6, 758–761.

- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Psychology Press.
- Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition: An International Journal*, 21, 422–430.
- Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of Type 1 and Type 2 data: Meta- d' , response-specific meta- d' , and the unequal variance SDT model. In S. M. Fleming & C. D. Frith (Eds.), *The cognitive neuroscience of metacognition* (pp. 25–66). Berlin, Germany: Springer.
- Maranesi, M., Ugolotti Serventi, F., Bruni, S., Bimbi, M., Fogassi, L., & Bonini, L. (2013). Monkey gaze behaviour during action observation and its relationship to mirror neuron activity. *European Journal of Neuroscience*, 38, 3721–3730.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Rach, S., Diederich, A., & Colonius, H. (2011). On quantifying multisensory interaction effects in reaction time and detection rate. *Psychological Research*, 75, 77–94.
- Ricci, R., & Chatterjee, A. (2004). Sensory and response contributions to visual awareness in extinction. *Experimental Brain Research*, 157, 85–93.
- Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of Neurophysiology*, 96, 1358–1369.
- Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience*, 1, 165–175.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112, 585–615.
- Santello, M., & Soechting, J. F. (1998). Gradual molding of the hand to object contours. *Journal of Neurophysiology*, 79, 1307–1320.

- Schilbach, L. (2010). A second-person approach to other minds. *Nature Reviews Neuroscience*, 11, 449.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36, 393–414.
- Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., & Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Consciousness and Cognition: An International Journal*, 20, 1787–1792.
- Southgate, V., Johnson, M. H., El Karoui, I., & Csibra, G. (2010). Motor system activation reveals infants' on-line prediction of others' goals. *Psychological Science*, 21, 355–359.
- Tamietto, M., Cauda, F., Celeghin, A., Diano, M., Costa, T., Cossa, F. M., ... de Gelder, B. (2015). Once you feel it, you see it: Insula and sensory-motor contribution to visual awareness for fearful bodies in parietal neglect. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 62, 56–72.
- Thioux, M., & Keysers, C. (2015). Object visibility alters the relative contribution of ventral visual stream and mirror neuron system to goal anticipation during action observation. *NeuroImage*, 105, 380–394.
- Thorne, D. R. (2006). Throughput: A simple performance index with desirable characteristics. *Behavior Research Methods*, 38, 569–573.
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 31, 155–165.
- Urgen, B. A., & Miller, L. E. (2015). Towards an empirically grounded predictive coding account of action understanding. *The Journal of Neuroscience* 35, 4789– 4791.
- Van den Stock, J., Tamietto, M., Zhan, M., Heinecke, A., Hervais- Adelman, A., Legrand, L. B., ... de Gelder, B. (2014). Neural correlates of body and face

- perception following bilateral destruction of the primary visual cortices. *Frontiers in Behavioral Neuroscience*, 8, 30.
- Whitwell, R. L., & Goodale, M. A. (2013). Grasping without vision: Time normalizing grip aperture profiles yields spurious grip scaling to target size. *Neuropsychologia*, 51, 1878 –1887.
- World Medical Association General Assembly (2008). Declaration of Helsinki. Ethical principles for medical research involving human subjects. *World Medical Journal*, 54, 122–125.
- Yang, Y., Chevallier, S., Wiart, J., & Bloch, I. (2014). Time-frequency optimization for discrimination between imagination of right and left hand movements based on two bipolar electroencephalography channels. *EURASIP Journal on Advances in Signal Processing*, 1, 38.