Cognition 168 (2017) 140–145

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Short Communication

The heaviness of invisible objects: Predictive weight judgments from observed real and pantomimed grasps

Jessica Podda [a], Caterina Ansuini [a], Roberta Vastano [b], Andrea Cavallo [c,a], Cristina Becchio [a,c,*]

[a] Cognition, Motion and Neuroscience, Fondazione Istituto Italiano di Tecnologia, Genova, Italy
[b] Robotics, Brain and Cognitive Sciences, Fondazione Istituto Italiano di Tecnologia, Genova, Italy
[c] Department of Psychology, University of Torino, Torino, Italy

ARTICLE INFO

Article history:
Received 15 February 2017
Revised 21 June 2017
Accepted 22 June 2017
Available online 1 July 2017

Keywords:
Action observation
Shared experience
Object representation
Hand kinematics
Prediction

ABSTRACT

Observation of others' actions has been proposed to provide a *shared experience* of the properties of objects acted upon. We report results that suggest a similar form of shared experience may be gleaned from the observation of pantomimed grasps, i.e., grasps aimed at pretended objects. In a weight judgment task, participants were asked to observe a hand reaching towards and grasping either a real or imagined glass, and to predictively judge its weight. Results indicate that participants were able to discriminate whether the to-be-grasped glass was empty, and thus light, or full, and thus heavy. Worthy of further investigation, this finding suggests that by observing others' movements we can make predictions, and form expectations about the characteristics of objects that exist only in others' minds.

1. Introduction

The behavior of others supplies a rich source of information about the world around us. The ability to process this information is key for learning about the properties of objects acted upon, as well as to read others' intentions and expectations (Cavallo, Koul, Ansuini, Capozzi, & Becchio, 2016; for review, see Ansuini, Cavallo, Bertone, & Becchio, 2015). By observing another person grasping and lifting a cup, for example, we can immediately deduce whether the cup is full or empty, even when we cannot see inside the cup (Bingham, 1987; Hamilton, Wolpert, & Frith, 2004; Maguinness, Setti, Roudaia, & Kenny, 2013). Through this, we may also perceive whether the other person had a correct or false expectation about the weight of the cup (Finisguerra, Amoruso, Makris, & Urgesi, in press; Runeson & Frykholm, 1983), and use this information to reduce 'surprise effects' in our own interactions with the environment (Meulenbroek, Bosga, Hulstijn, & Miedl, 2007).

As such, observing other people acting upon objects involves a form of *experience sharing* (Brown & Brüne, 2012; Limanowski & Blankenburg, 2013): we can learn about the properties of a given object through others' action, without needing to have first hand experience. In this way, supposedly hidden, internal properties of objects, such as weight, become available for perception (Runeson, 1985).

The question addressed in the present study is whether a similar form of shared experience may be gleaned from the observation of pantomimed actions, i.e., actions aimed at imagined, rather than real objects. Put simply: can we share through others' action the characteristics of an object that is not there?

The hypothesis that pantomimed actions contribute to shared experience of imagined objects is motivated, in part, by studies investigating the kinematics of pantomimed grasping actions. When pretending to pick up imagined objects, we move and shape our hands quite differently from when we grasp real objects (Cavina-Pratesi, Kuhn, Ietswaart, & Milner, 2011; Goodale, Jakobson, & Keillor, 1994). Still, pantomimed actions demonstrate at least some perceptual features of the pretended object. For example, during pantomimed grasping, grip width depicts the width of the imaginary object (Goldenberg, Hartmann, & Schlott, 2003). Moreover, there is evidence that, early on in the movement, the kinematics of both real and pantomimed movements is scaled to the weight of the object to be grasped (e.g., Ansuini et al., 2016; Eastough & Edwards, 2007). This raises the possibility that, even before contact, observers can take advantage of kinematic information in order to form a shared representation of the object acted upon – be it real or imagined. The present study aimed to test this hypothesis by asking participants to make predictive weight

* Corresponding author at: C'MON Unit, Istituto Italiano di Tecnologia, Center for Human Technologies, Via E. Melen 83, 16152, Italy.
E-mail address: cristina.becchio@iit.it (C. Becchio).

http://dx.doi.org/10.1016/j.cognition.2017.06.023
0010-0277/© 2017 The Author(s). Published by Elsevier B.V.

judgments from the observation of real and pantomimed reach-to-grasp movements.

2. Material and methods

2.1. Participants

Twenty-four participants took part in the experiment (12 females; M age = 24; age range = 19–30 years old). The sample size was determined in advance by power analysis using effect sizes observed in a pilot study. A sample size of 24 was calculated to detect a Cohen's d of 0.70 with alpha set at 0.05 (one-sided), and power set at 0.90. All participants were right handed, with normal or corrected-to-normal vision, and with no history of either psychiatric or neurological disorders. The research was approved by the local ethical committee (ASL 3 Genovese), and was carried out in accordance with the principles of the revised Helsinki Declaration (World Medical Association General Assembly, 2008). Written informed consent was obtained from each participant.

2.2. Experimental stimuli: video capturing, selection and editing procedure

To create the stimuli to be used in the main experiment, we filmed 15 agents (10 females; M age = 28.8; age range = 24–32 years old) performing real and pantomimed reach-to-grasp movements. For *real reach-to-grasp movements*, participants were requested to reach towards, and grasp, either an empty glass (139 g) or a glass filled with iron screws (838 g), placed at a distance of 48 cm from the participant's body midline. For *pantomimed reach-to-grasp movements*, the glass, either empty or filled, was positioned at a displaced location (12 cm away from the target position). Participants were instructed to imagine that an identical glass was positioned at the target position, and were asked to pretend to perform the very same action sequence towards the imagined glass.

Reach-to-grasp movements were filmed from a lateral viewpoint using a digital video camera (Sony Handycam 3D, 25 frames/s; Sony Corporation, Tokyo, Japan). Simultaneously, hand movement kinematics were recorded using a near-infrared camera motion capture system (frame rate: 100 Hz; Vicon Motion Systems Ltd, Oxford, UK). To assess the availability of weight information over time, a set of kinematic variables was calculated using a custom Matlab (MathWorks, Natick, MA, USA) script (see Table S1 for a detailed description of the kinematic variables). All variables were computed only considering the reach-to-grasp phase of the movement, i.e., from '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) to 'reach offset' (i.e., the time at which the wrist velocity dropped below a 20 mm/s threshold) at an interval of 10% of the normalized movement time (see Ansuini, Cavallo, Campus, et al., 2016; Ansuini et al., 2016 for further details).

With respect to the stimulus selection, we proceeded as follows: first, we submitted the computed kinematic variables of real and pantomimed reach-to-grasp movements to separate linear discriminant analyses (LDAs) to find the linear combinations of features that, for each type of movement, separated between heavy and light objects. Kinematic data from one participant were discarded due to technical problems with video recording. Discriminant function analyses using a leave-one-out cross validation method (Efron, 1982) revealed that classification of object weight was significantly above chance level (i.e., 50%) for both real and pantomimed reach-to-grasp movements (see Table 1 for details). This conclusion was supported by the results of permutation tests (1000 simulations for each LDA model) (all p values < 0.001).

The kinematic variables that contributed the most to weight classification were grip aperture, wrist velocity and thumb/index finger vertical displacement for real reach-to-grasp movements, wrist velocity and thumb vertical displacement for pantomimed reach-to-grasp movements. Fig. S1 provides a visual summary of how each kinematic variable contributed to the classification of object weight over time for real and pantomimed movements.

With the new space defined via the LDA, we next selected, for each type of reach-to-grasp movements (real, pantomime) and for each weight (light, heavy), the 50 movements that minimized the within-weight distance, i.e., the distance from the mean variate score of heavy versus light objects. This procedure allowed us to identify a final set of 200 representative movements (50 real reach-to-grasp/light; 50 real reach-to-grasp/heavy; 50 pantomimed reach-to-grasp/light; 50 pantomimed reach-to-grasp/heavy).

The 200 unique video clips corresponding to the selected movements were edited using Adobe Premiere Pro CS6 (.avi format, disabled audio, 25 frames/s; Adobe Systems Software Ltd, Dublin, Ireland). To produce spatial occlusion of the to-be-grasped object, a grey rectangular mask (height = 51.5 mm; length = 31.1 mm) was superimposed onto the target object location. The size and the position of this mask were kept constant across participants. Each video was edited so as to begin at reach onset and to end at reach offset (see Video S1). Movement durations (from reach onset to reach offset) did not differ significantly between light and heavy objects, both for real (Light object: M = 869.20 ms, 1SE = 25.76; Heavy object: M = 927.40 ms, 1SE = 27.02) ($t(98) = -1.56$, $p = 0.122$, $d = 0.31$, 95% CI [−132.28, 15.88]) and for pantomimed reach-to-grasp movements (Light object: M = 933.80 ms, 1SE = 29.36, Heavy object: M = 923.20 ms, 1SE = 30.96) ($t(98) = 0.25$, $p > 0.250$, $d = 0.05$, 95% CI [−74.07, 95.27]).

2.3. Procedure and measures

The experiment was carried out in a dimly lit room. Participants sat in front of a 17-in. computer screen (resolution: 1280×800; frame rate: 75 Hz) at a viewing distance of 50 cm. They were presented with video clips of the reach-to-grasp phase of the selected movements (see 'Experimental stimuli: video capturing, selection and editing procedure' section). A one-interval discrimination design was employed (see Fig. 1).

After each video, participants were asked to judge as accurately and as quickly as possible the weight of the object towards which the movement was directed (i.e., light versus heavy object). Responses were given by pressing one of two keys on a keyboard. For half of the participants, the Italian word 'leggero' (light) on the left prompted a button press with the index finger on the left button of a wireless keyboard touchpad, while the word 'pesante' (heavy) on the right prompted a button press with the middle finger on the touchpad right button. The position of the two words was counterbalanced within and across participants. Participants were instructed to respond either during the video, or within a maximum of 3000 ms after the video ended. To ensure that movement sequences could be temporally attended, that is, to provide participants enough time to focus on movement start and prevent anticipation, +13 up to +28 static frames in step of +1 were added at the beginning of all video clips. To equate stimulus duration within each type of reach-to-grasp movement (i.e., real and pantomimed), static frames were also added at the end of the videos in a compensatory manner (+14 up to +29 in step of +1). In this way, each real movement clip lasted exactly 2520 ms and each pantomimed movement clip lasted exactly 2600 ms. After indicating a response, participants were requested to rate the confidence