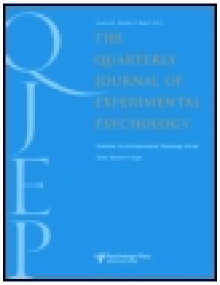

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Observation of another's action but not eye gaze triggers allocentric visual perspective

Elisabetta Mazzarella [a], Antonia Hamilton [b], Luigi Trojano [a], Bianca Mastromauro [a] & Massimiliano Conson [a]

[a] Neuropsychology Laboratory, Department of Psychology, Second University of Naples, Caserta, Italy

[b] School of Psychology, University of Nottingham, Nottingham, UK

Accepted author version posted online: 28 May 2012. Version of record first published: 17 Aug 2012

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, i*First*, 1–14

Psychology Press
Taylor & Francis Group

Observation of another's action but not eye gaze triggers allocentric visual perspective

Elisabetta Mazzarella[1], Antonia Hamilton[2], Luigi Trojano[1], Bianca Mastromauro[1], and Massimiliano Conson[1]

[1]Neuropsychology Laboratory, Department of Psychology, Second University of Naples, Caserta, Italy
[2]School of Psychology, University of Nottingham, Nottingham, UK

In the present paper, we investigated whether observation of bodily cues—that is, hand action and eye gaze—can modulate the onlooker's visual perspective taking. Participants were presented with scenes of an actor gazing at an object (or straight ahead) and grasping an object (or not) in a 2 × 2 factorial design and a control condition with no actor in the scene. In Experiment 1, two groups of subjects were explicitly required to judge the left/right location of the target from their own (egocentric group) or the actor's (allocentric group) point of view, whereas in Experiment 2 participants did not receive any instruction on the point of view to assume. In both experiments, allocentric coding (i.e., the actor's point of view) was triggered when the actor grasped the target, but not when he gazed towards it, or when he adopted a neutral posture. In Experiment 3, we demonstrate that the actor's gaze but not action affected participants' attention orienting. The different effects of others' grasping and eye gaze on observers' behaviour demonstrated that specific bodily cues convey distinctive information about other people's intentions.

Keywords: Visual perspective; Action observation; Eye gaze; Social attention.

Visual perspective taking is the ability to understand the visual experience of another agent. There are two levels in perspective taking (Flavell, Everett, Croft, & Flavell, 1981; Michelon & Zacks, 2006; Piaget & Inhelder, 1967): Level 1 consists of the ability to judge what another agent can see or not see from his/her own point of view, while Level 2 consists of the ability to understand how another person perceives a given object from his/her viewpoint. In this context, perception and interpretation of a visual scene from one's own point of view is termed egocentric (first person) perspective, whereas perception and interpretation of a visual scene from the other person's point of view is called allocentric (third person) perspective. It is worth noting here that although the term "allocentric" is often used to refer to the spatial relationship between two external points, in social cognition studies it refers to "being concerned with another person as opposed to oneself" (Frischen, Loach, & Tipper, 2009, p. 213; see also U. Frith & de Vignemont, 2005).

Correspondence should be addressed to Massimiliano Conson or Luigi Trojano, Neuropsychology Laboratory, Department of Psychology, Second University of Naples, via Vivaldi 43, 81100 Caserta, Italy. E-mail: massimiliano.conson@unina2.it or luigi.trojano@unina2.it

http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2012.697905

The egocentric perspective has primacy both in children and in adults (Epley, Morewedge, & Keysar, 2004), but it has been recently demonstrated that, in some circumstances, one cannot easily ignore what other people see. In a Level 1 perspective-taking task, Samson, Apperly, Braithwaite, Andrews, and Bodley Scott (2010) presented healthy subjects with pictures of a room in which a human avatar faced one of the walls, where red discs were placed. In one experimental condition, both the participant and the avatar could see the same number of discs (consistent perspective), whereas in another condition the participant and the avatar saw a different number of discs (inconsistent perspective). Participants had to judge how many discs could be seen, either from their own perspective or from the avatar's perspective. Results showed that participants made slower self-perspective judgements in the inconsistent condition than in the consistent condition; moreover, in the consistent condition, participants were even quicker in responding from the avatar's perspective than from their own perspective. Taken together, such findings demonstrated that another's point of view exerts a strong influence on one's own visual perspective.

In a negative priming task, Frischen and colleagues (2009) required participants to reach for a target stimulus whilst ignoring a distractor. Distractors located close to a participant's hand were inhibited strongly, fitting with an egocentric frame of reference; when participants performed the same reaching task in front of another agent, locations close to the hand of the observed agent (and far away from the participant's hand) were also inhibited strongly. Thus, the presence of another actor led participants to allocate selective attention to objects in a manner compatible with the actor's instead of their own point of view (allocentric perspective).

In a Level 2 perspective-taking task, Tversky and Hard (2009) required healthy participants to code (left or right) spatial location of a target object in a visual scene where no actor was present (control condition) or where an actor was present and either only looked at the target or looked at and reached for the target. Results showed that presence of the actor induced subjects to code the target location from the actor's point of view, without differences between the looking and looking/reaching conditions. However, Tversky and Hard's experiments did not include a condition where the actor did not gaze at the target object or act upon it and did not disentangle the effect of eye gaze and action. From the data reviewed above, some issues emerge that need to be clarified. These include the effect of the mere presence of the observed actor and possible differences between the effects of the actor's action and gaze on the onlooker's visual perspective taking.

Action and eye gaze are fundamental components of social interactions and can be considered relevant "allocentric bodily cues" (Chartrand & Bargh, 1999; Decety & Grezes, 1999; Gallese, Keysers, & Rizzolatti, 2004; Sommerville & Decety, 2006; Tipper, 2010). Eye gaze is a well-known social cue (Baron-Cohen, Campbell, Kamiloff-Smith, Grant, & Walker, 1995) capable of capturing the onlooker's attention: When an observer sees a person looking (or turning his head) to one side, his/her attention is drawn to the same side (e.g., Driver et al., 1999; Ricciardelli, Baylis, & Driver, 2000). This congruency effect is usually referred to as joint attention and might prepare the observer to perform efficient responses in complex situations (Fischer, Prinz, & Lotz, 2008). However, joint attention studies revealed that observation of pointing actions can also affect attention orienting. Crostella, Carducci, and Aglioti (2009) showed that distracting gaze stimuli selectively interfered with responses made with saccadic movements, whereas distracting pointing hand stimuli specifically interfered with pointing responses. Fischer and Szymkowiak (2004) demonstrated that pointing postures facilitated encoding of target locations via attentional orienting whereas grasping postures did not modulate observer's attention. These data suggest that specific hand postures can activate attentional social processes in a highly specific manner (Crostella et al., 2009; Fischer et al., 2008; Fischer & Szymkowiak, 2004; see also, Burton, Bindemann, Langton, Schweinberger, & Jenkins, 2009). This is consistent with neurofunctional evidence showing how observation of bodily movements recruits different brain regions depending on the effector (e.g., eyes, head, or hands; Castiello, 2003; Grèzes & Decety, 2001; Haxby, Hoffman, & Gobbini, 2002; for a review see Allison, Puce, & McCarthy, 2000).

On these bases, we hypothesized that another's action and eye gaze differently affect visual perspective taking and social attention. In particular, we predicted that others' grasping actions (but not eye gaze) can modulate participants' visual perspective taking, consistent with the idea of its motor underpinnings (Frischen et al., 2009; Tversky & Hard, 2009). Conversely, on the basis of the so-called gaze cueing effect (for a review, see Frischen, Bayliss, & Tipper, 2007), we expected that another's gaze can orient participants' attention in a visual target detection task, but also expected that others' grasping would not have the same effect (see Fischer & Szymkowiak, 2004). We tested these predictions by verifying the differential effect of another's bodily cues—that is, grasping movements or eye gaze—on Level 2 perspective taking (Experiments 1 and 2) and on joint attention (Experiment 3).

EXPERIMENT 1

The specific aim of this experiment was to test whether an onlooker's perspective taking was differentially affected by an actor's gaze, grasping, or a combination of the two, or by the "mere presence" of the actor in a neutral position (i.e., neither grasping nor gazing). For this purpose, in a Level 2 visual perspective task, two groups of subjects were *explicitly* required to code target location from an egocentric (their own point of view) or an allocentric (the actor's point of view) perspective.

We employed an explicit task to directly assess the effect of bodily cues on the two kinds of perspective taking. This contrasts with Tversky and Hard (2009), who adopted an implicit task in which participants were not committed to one specific perspective and inferred egocentric or allocentric coding from left or right subjects' responses. However, if one considers that egocentric perspective is more natural and has the primacy in spatial coding (Epley et al., 2004), it would be possible to posit that participants always code target location with respect to their own point of view, and that the so-called allocentric responses just represent errors. This means that an implicit task does not allow us to clearly distinguish whether the effects of the experimental conditions are related to an increase in error rate or to a specific influence on perspective taking. By contrast, a general increase in error rate in an explicit task would produce a pattern (higher number of egocentric responses in the allocentric group and vice versa) clearly discernible from the expected specific facilitation of allocentric perspective taking.

Method

Participants

Sixty right-handed healthy subjects (30 females; age range 22–30 years) participated in the experiment. The subjects were randomly assigned to one of two groups, balanced for age and gender: The egocentric group had to code target location from one's own perspective, whereas the allocentric group had to code target location from the actor's perspective. All subjects were naïve with respect to the aims and the hypothesis of the experiment. The study was conducted in accordance with the ethical standards of the Helsinki Declaration; written informed consent was obtained from all participants.

Stimuli and procedure

Participants were presented with scenes representing a human model (an actor) at a table on which one target object (a bottle or a glass) was positioned. Four "actor scenes" were devised (Figure 1). In the first scene, the actor had a straight gaze and did not grasp the target (no-gaze/no-action). In the second scene, the actor had a straight gaze but grasped the target (no-gaze/yes-action). In the third scene, the actor gazed towards the target but did not grasp it (yes-gaze/no-action), whereas in the fourth scene, the actor both gazed towards and grasped the target (yes-gaze/yes-action); in both yes-gaze conditions, the actor looked precisely at the point in which his hand (would) come into contact with the target object. In a control scene, no actor was

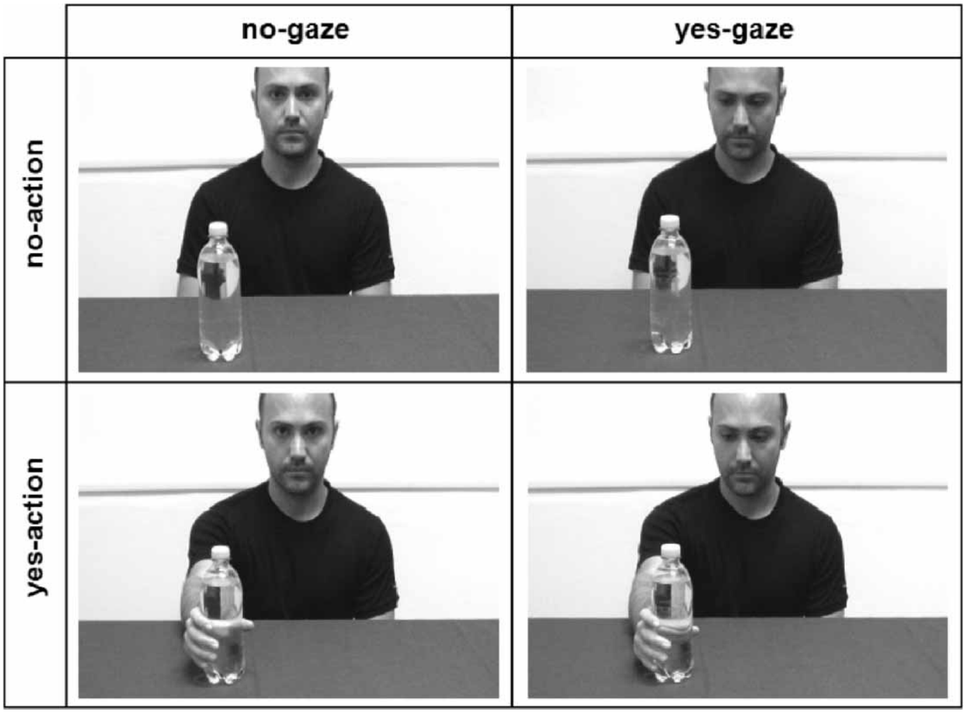

Figure 1. *Experiment 1. Schematic representation of the experimental design depicting the four experimental conditions involving the human model. In this display, the no-actor condition (the same as the other conditions but without the human model) is not represented.*

present (no-actor condition). Each scene was enclosed in a rectangular frame in a 700 × 500-pixel array. In each trial, a fixation point (500 ms) was followed by a visual scene that remained on the screen until subjects gave their response.

The five scenes were presented 12 times in a randomized order, for a total of 60 trials. Before starting the task, participants of the two groups were presented with task instructions specifying the reference frame (one's own or the actor's point of view) that they had to adopt for target coding. Thus, task instructions for the egocentric group were as follows: "Where is the bottle/glass? On the left or on the right with respect to your own point of view?" Instructions for participants of the allocentric group, instead, were: "Where is the bottle/glass? On the left or on the right with respect to the actor's point of view?" Participants responded by pressing one of two buttons on the computer keyboard ("B" for left and "H" for right on the QWERTY keyboard) with their right dominant hand. Before the task, several practice trials were given and were discarded from statistical analysis; both accuracy and response speed were recorded.

The subjects' responses were transformed according to a binary code: (left or right) responses consistent with an egocentric perspective were scored as 0, whereas (left or right) responses consistent with an allocentric perspective were scored as 1. Values deriving from this binary coding were arcsine transformed for performing parametrical statistical analysis (Hogg & Craig, 1995).

Results and comment

Mean proportion of allocentric responses in all the experimental conditions is reported separately for the two groups in Figure 2. Arcsine-transformed responses underwent a three-way mixed analysis of variance (ANOVA) with eye gaze (no-gaze or yes-gaze) and grasping (no-action or yes-action)

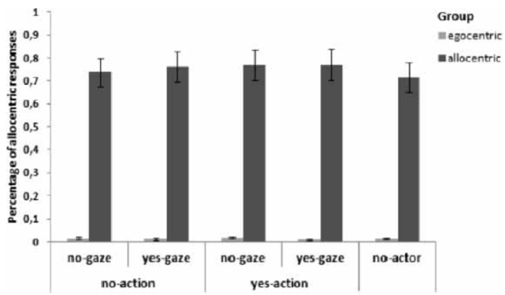

Figure 2. *Experiment 1. Mean proportion of allocentric responses (bars are* SEM*) in all the five experimental conditions, separately for the egocentric and the allocentric group.*

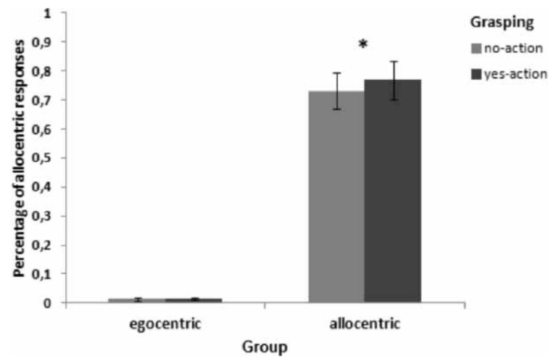

Figure 3. *Experiment 1. Mean proportion of allocentric responses (bars are* SEM*) in the no-action and yes-action conditions, separately for the egocentric and the allocentric group (i.e., significant interaction between grasping and group). *Significant at* p *= .028.*

as within-subject factors and with group (egocentric or allocentric) as a between-subjects factor.

Results showed significant main effects of grasping, $F(1, 58) = 5.366$, $p = .024$, $\eta^2_p = .085$, with a higher number of allocentric responses when the actor grasped the target object (mean = .39, *SEM* = .03) than when he did not grasp it (mean = .31, *SEM* = .02), and of group, $F(1, 58) = 112.657$, $p = .0001$, $\eta^2_p = .660$, with allocentric responses prevailing in the allocentric group (mean = .76, *SEM* = .05; since allocentric responses are correct responses in this group, .76 corresponded to mean accuracy) and virtually absent in the egocentric group (mean = .01, *SEM* = .05; since allocentric responses are wrong in this group, mean accuracy was .99). The main effect of eye gaze was not significant, $F(1, 58) = 1.459$, $p = .232$, $\eta^2_p = .025$ (no-gaze condition: mean = .38, *SEM* = .03; yes-gaze condition: mean = .39, *SEM* = .02). There was a significant interaction between grasping and group, $F(1, 58) = 5.175$, $p = .027$, $\eta^2_p = .082$: The egocentric group provided very few allocentric responses in both no-action (mean = .01, *SEM* = .03) and yes-action conditions (mean = .01, *SEM* = .04), whereas the allocentric group provided a higher number of allocentric responses in the yes-action (mean = .78, *SEM* = .05) than in the no-action condition (mean = .72, *SEM* = .04). No other interaction was significant: Eye Gaze × Group: $F(1, 58) = 2.347$, $p = .131$, $\eta^2_p = .039$; Eye Gaze × Grasping: $F(1, 58) = 2.937$, $p = .092$, $\eta^2_p = .048$; Eye Gaze × Grasping × Group: $F(1, 58) = 2.465$, $p = .122$, $\eta^2_p = .041$.

Post hoc comparisons (unpaired *t* tests) on the interaction between grasping and group showed that in the egocentric group the number of allocentric responses did not differ between the yes-action and the no-action condition ($t = -0.167$, $p = .869$), whereas in the allocentric group the number of allocentric responses was significantly higher in the yes-action than in the no-action condition ($t = -2.314$, $p = .028$; see Figure 3).

To verify whether the mere presence of an actor in the scene was sufficient to automatically elicit an allocentric perspective (Tversky & Hard, 2009), we performed planned contrasts (paired *t* tests) to compare the no-actor condition with the no-gaze/no-action condition and with the two conditions in which only one cue was present (no-gaze/yes-action and yes-gaze/no-action), taking into account both the egocentric and the allocentric group. Results showed that the no-actor condition (mean = .31, *SEM* = .05) did not differ from both the no-gaze/no-action condition (mean = .32, *SEM* = .05; $t = 1.199$; $p = .235$) and the yes-gaze/no-action condition (mean = .34, *SEM* = .05; $t = 1.620$; $p = .111$). On the contrary, the number of allocentric responses was significantly higher in the no-gaze/yes-action (mean = .38, *SEM* = .05) than in the no-actor condition ($t = 2.423$; $p = .018$) and in the no-gaze/no action condition ($t = -2.673$, $p = .010$).

Reaction times (RTs) for correct responses underwent a three-way mixed ANOVA with eye gaze (no-gaze or yes-gaze) and grasping (no-action or yes-action) as within-subject factors and with group (egocentric or allocentric) as a between-subjects factor. Four subjects from the allocentric group were excluded from the data set since they provided no correct responses in one of the four experimental conditions. Results showed a significant main effect of group, $F(1, 55) = 26.667$, $p = .0001$, $\eta^2_p = .327$, with faster responses in the egocentric (mean = 580.05, $SEM = 75.94$) than in the allocentric group (mean = 1,160.65, $SEM = 82.92$). The main effects of eye gaze, $F(1, 55) = 1.605$, $p = .211$, $\eta^2_p = .028$, and grasping, $F(1, 58) = 2.737$, $p = .104$, $\eta^2_p = .047$, were not significant. Moreover, no interaction was statistically significant: Eye Gaze × Group: $F(1, 55) = 0.532$, $p = .469$, $\eta^2_p = .010$; Grasping × Group: $F(1, 55) = 2.737$, $p = .104$, $\eta^2_p = .047$; Eye Gaze × Grasping: $F(1, 58) = 0.438$, $p = .511$, $\eta^2_p = .008$; Eye Gaze × Grasping × Group: $F(1, 58) = 1.314$, $p = .257$, $\eta^2_p = .023$.

In synthesis, in this experiment participants who were explicitly required to judge the target location from their own point of view (egocentric group) showed accuracy at ceiling and RTs significantly faster than those of participants of the allocentric group; moreover, their performance was not modulated by any experimental condition. Instead, when participants had to judge the target location from the actor's point of view (allocentric group), the actor's action strongly favoured participants' allocentric perspective, whereas eye gaze or the actor's mere presence had no effect.

Since we adopted an explicit task in which egocentric and allocentric perspective taking could be analysed independently, without any bias related to error rate, the present data demonstrated that observation of actor's grasping specifically facilitated allocentric perspective taking. However, our results seem to be at odds with Tversky and Hard's (2009) finding that observation of eye gaze alone is sufficient to trigger allocentric perspective. One important difference between the present and Tversky and Hard's study lies in the implicit versus explicit nature of the perspective-taking task. In fact, recent data demonstrated that taking another's point of view can be as natural and automatic as taking one's own perspective, thus suggesting that both the onlooker's and the actor's perspective can be processed at an implicit level (Samson et al., 2010).

EXPERIMENT 2

The aim of Experiment 2 was to verify whether the specific effect of action observation on the onlooker's visual perspective can be replicated in the same experimental set-up, when an implicit processing of one's own or the actor's point of view is required from participants.

Method

Participants

Thirty right-handed healthy subjects (15 females; age range 21–29 years) participated in the experiment. All participants were unaware of purposes and predictions of the experiment at the time of testing. The study was conducted in accordance with the ethical standards of Helsinki Declaration; written informed consent was obtained from all participants.

Stimuli and procedure

Stimuli and procedure were the same as those in the first experiment, with the exception that here task instructions did not make any mention on the perspective the participants had to assume for coding target location. Actually, the following instructions were displayed on the computer screen: "Where is the bottle/glass? On the left or on the right?"; no further information was provided to subjects. Before the task, several practice trials were given and were discarded from statistical analysis; both accuracy and response speed were recorded. Participants' responses were coded and transformed following the same procedure as that in Experiment 1.

Results and comment

Mean proportion of allocentric responses in all the experimental conditions is reported in Figure 4. A two-way repeated measure ANOVA with eye gaze (no-gaze or yes-gaze) and grasping (no-action or yes-action) as within-subject factors showed a significant main effect of grasping, $F(1, 29) = 7.453$, $p = .011$, $\eta^2_p = .204$, with a higher number of allocentric responses when the actor grasped the target object (mean = .13, *SEM* = .05) than when he did not grasp it (mean = .19, *SEM* = .06). The main effect of eye gaze, $F(1, 29) = 0.718$, $p = .404$, $\eta^2_p = .024$, was not significant (no-gaze: mean = .16, *SEM* = .05; yes-gaze: mean = .17, *SEM* = .06). Moreover, the interaction between eye gaze and grasping was not significant, $F(1, 29) = 2.029$, $p = .159$, $\eta^2_p = .067$.

As in Experiment 1, to verify whether the mere presence of an actor in the scene was sufficient to automatically elicit an allocentric perspective, we performed planned contrasts (paired *t* tests) to compare the no-actor condition with the no-gaze/no-action condition and with the two conditions in which only one cue was present (no-gaze/yes-action and yes-gaze/no-action). Results showed that the no-actor condition (mean = .11, *SEM* = .04) did not differ from both the no-gaze/no-action condition (mean = .13, *SEM* = .05; $t = 1.455$; $p = .156$) and the yes-gaze/no-action condition (mean = .15, *SEM* = .05; $t = 1.401$; $p = .172$). On the contrary, the number of allocentric responses was significantly higher in the no-gaze/yes-action (mean = .19, *SEM* = .05) than in the no-actor condition ($t = 2.162$; $p = .039$) and in no-gaze/no-action condition ($t = -2.301$; $p = .029$).

To analyse RT data, we planned a two-way repeated measure ANOVA with eye gaze (no-gaze or yes-gaze) and grasping (no-action or yes-action) as within-subject factors to be performed separately on egocentric and allocentric responses. However, only 8/30 subjects provided allocentric responses for all combinations of the two independent factors, thus precluding reliable statistical analyses on this measure. As regards egocentric responses, we could perform the ANOVA since sufficient data were gathered in 27/30 participants. Results did not show significant main effects of grasping, $F(1, 26) = 1.941$, $p = .175$, $\eta^2_p = .069$, and eye gaze, $F(1, 26) = 2.616$, $p = .118$, $\eta^2_p = .091$. Also the interaction between eye gaze and grasping was not significant, $F(1, 29) = 1.511$, $p = .230$, $\eta^2_p = .055$.

In synthesis, the present results confirmed that allocentric responses were more frequent when participants observed the actor grasping the target object than when they observed the actor gazing the target or the actor in a neutral position. Such findings replicated those from Experiment 1, thus demonstrating that the specific effect of action

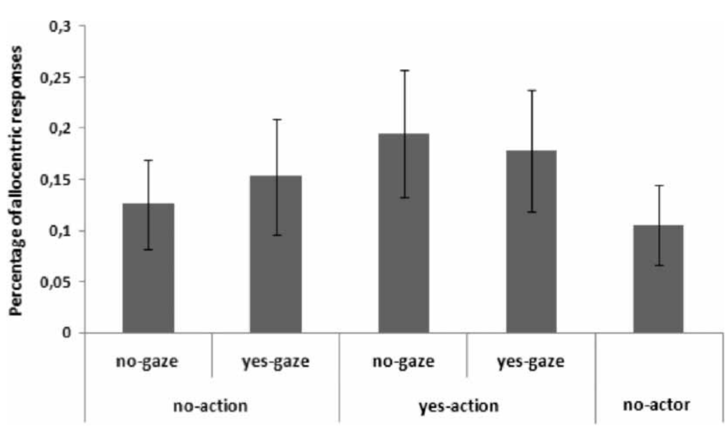

Figure 4. *Experiment 2. Mean proportion of allocentric responses (bars are* SEM*) in all the five experimental conditions.*

observation on allocentric perspective taking was not related to the explicit nature of the task adopted in Experiment 1. The divergences between the present and Tversky and Hard's (2009) findings as regards the effect of the looking condition alone could be accounted for by differences in stimuli, as is argued in the General Discussion.

EXPERIMENT 3

In both Experiments 1 and 2 we did not find any effect of actor's gaze on the onlooker's visual perspective taking, which was, instead, influenced by actor's grasping. This pattern stands in contrast with evidence gathered from studies on joint attention suggesting that other people's gaze influences the observer's attentional orienting (Driver et al., 1999; Frischen et al., 2007; Ricciardelli et al., 2000), whereas grasping actions do not (Fischer & Szymkowiak, 2004). On this basis, one could expect that in a target detection task with the experimental conditions employed in Experiments 1 and 2, we should found a significant effect of eye gaze but not of action observation on participants' performance. This prediction was tested in Experiment 3.

Method

Participants

Twenty-eight right-handed healthy subjects (14 females; age range 22–28 years) participated in the experiment. All subjects were naïve with respect to the aims and the hypothesis of the study. The study was conducted in accordance with the ethical standards of the Helsinki Declaration, and informed consent was obtained from all participants.

Stimuli and procedure

Visual scenes were photos depicting an actor at a table on top of which no object was present. The actor could perform a grasping movement and/or gaze towards the right or the left side of the table, whereas in other conditions he did not perform any grasping or gazing (straight gaze). Thus nine scenes were derived from the combination of the actor's grasp and gaze.

In each trial, after a fixation point (500 ms), one of these scenes appeared on the computer screen, followed by the same image together with the target object (i.e., a glass) appearing unpredictably and equally often on the left or on the right side of the visual scene (Figure 5). Subjects were asked to make speeded target detections by pressing a central button on the keyboard ("B" on the QWERTY keyboard).

The nine conditions were repeated twice for the presentation of the target (left or right) and multiplied 6 times for two cue duration (300 and 600 ms) for a total of 216 trials. Seventy-two catch trials were also included (36 for 300 ms and 36 for 600 ms), in which no target appeared after the cue. The catch trial remained on the screen for 1,500 ms, and the subjects had to withdraw from responding and to wait for presentation of the next trial.

The experimental design allowed us to manipulate two kinds of congruency between actor's grasping/eye gaze and target location: action congruency (congruent, incongruent, and neutral–no action) and eye gaze congruency (congruent, incongruent, and neutral–straight gaze).

Results

Mean RTs in all the experimental conditions are reported in Figure 6. A three-way repeated measure ANOVA with eye gaze (left, right, or no-gaze), grasping (left, right, or no-grasp), and target location (left or right) as within-subject factors performed on mean RTs showed a significant main effect of eye gaze, $F(2, 58) = 7.445$, $p = .001$, $\eta^2_p = .204$, but not of grasping, $F(2, 58) = 0.060$, $p = .942$, $\eta^2_p = .002$, or target location, $F(1, 29) = 0.534$, $p = .471$, $\eta^2_p = .018$. There was a significant interaction between eye gaze and target location, $F(2, 58) = 5.338$, $p = .007$, $\eta^2_p = .155$, whereas all the other interactions were not significant: Eye Gaze × Grasping: $F(4, 116) = 0.129$, $p = .972$, $\eta^2_p = .004$; Grasping × Target Location, $F(2, 58) = 0.411$, $p = .665$, $\eta^2_p = .014$; Eye Gaze ×

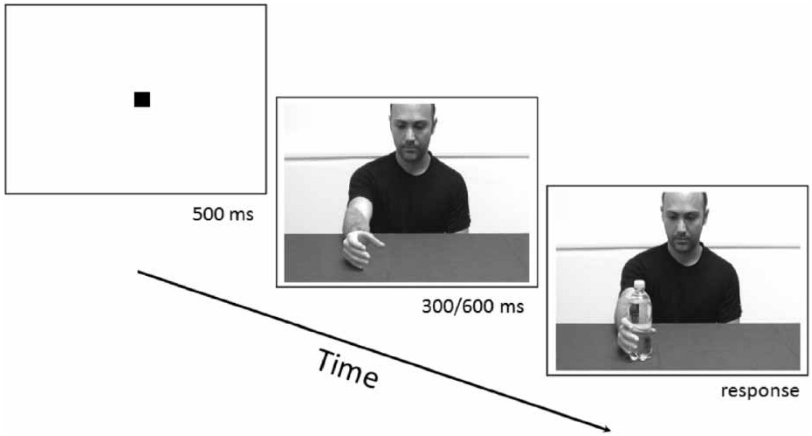

Figure 5. *Experiment 3. Example of a trial sequence depicting the yes-gaze/yes-action condition.*

Grasping × Target Location: $F(4, 116) = 0.026$, $p = .999$, $\eta_p^2 = .001$.

Post hoc comparisons (paired *t* tests) on the main effect of eye gaze showed that subjects' RTs were significantly slower when the target was preceded by a straight gaze (no-gaze condition: mean = 464.03 ms, *SEM* = 11.36) with respect to rightward (mean = 439.97, *SEM* = 14.86; $t = 2.746$, $p = .010$) or a leftward (mean = 437.26, *SEM* = 11.43; $t = 3.982$, $p = .0001$) gaze, whereas there were no differences between the two averted gaze conditions ($t = 0.371$, $p = .713$).

Post hoc comparisons (paired *t* tests) on the interaction between eye gaze and target location

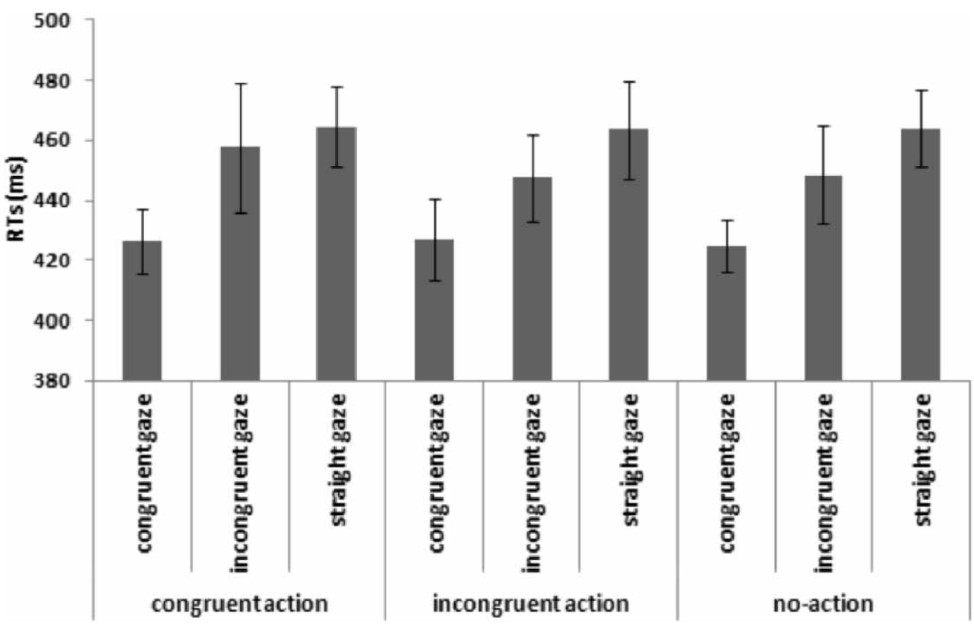

Figure 6. *Experiment 3. Mean reaction times (RTs; bars are* SEM*) in all the nine experimental conditions derived from combination of the agent's grasping, eye gaze, and target location.*

(Figure 7) showed that right-sided targets were detected significantly faster when preceded by a congruent rightward gaze than an opposite ($t = -2.688$, $p = .012$) or a straight gaze ($t = 3.665$, $p = .001$), whereas opposite and straight gaze cues did not differ ($t = 1.508$, $p = .142$). Left-sided targets were detected significantly faster when preceded by a congruent leftward gaze than an opposite ($t = 2.120$, $p = .048$) or a straight gaze ($t = 3.913$, $p = .001$), whereas there was no significant difference between opposite and straight gaze cues ($t = 0.901$, $p = .375$).

In synthesis, results showed faster detection times when the target was in a location congruent with gaze direction, consistent with the so-called gaze-cueing effect (Frischen et al., 2007). Instead, we did not find evidence of cueing effect related to the actor's grasping. This finding is remarkable when considering that using exactly the same stimuli in Experiments 1 and 2, actor's action but not eye gaze strongly affected participants' visual perspective taking. Such data provided a strong argument to discuss the differential effects of grasping and gaze on perspective taking and attentional orienting, respectively.

GENERAL DISCUSSION

We investigated the influence of different social cues—that is, another's action and eye gaze—on participants' performance in Level 2 visual perspective taking and joint attention tasks. Results showed that allocentric perspective was activated in the onlooker when a human actor grasped the target object, but not when the actor gazed towards the target. Moreover, the mere presence of the actor was not sufficient to induce participants to adopt an allocentric perspective. Such an action-related shifting from egocentric to allocentric visual perspective was a reliable finding not due to a bias in the participants' responses. By contrast, the direction of actor's gaze selectively affected participants' joint attention.

Tversky and Hard (2009) recently suggested that coding an object location in a scene where no person is present may be simply accomplished by defining spatial relations between the objects with respect to one's own body; however, if the scene includes a person, coding of object location may be preceded by an attempt to interpret the entire scene focusing on the role of that person. Since

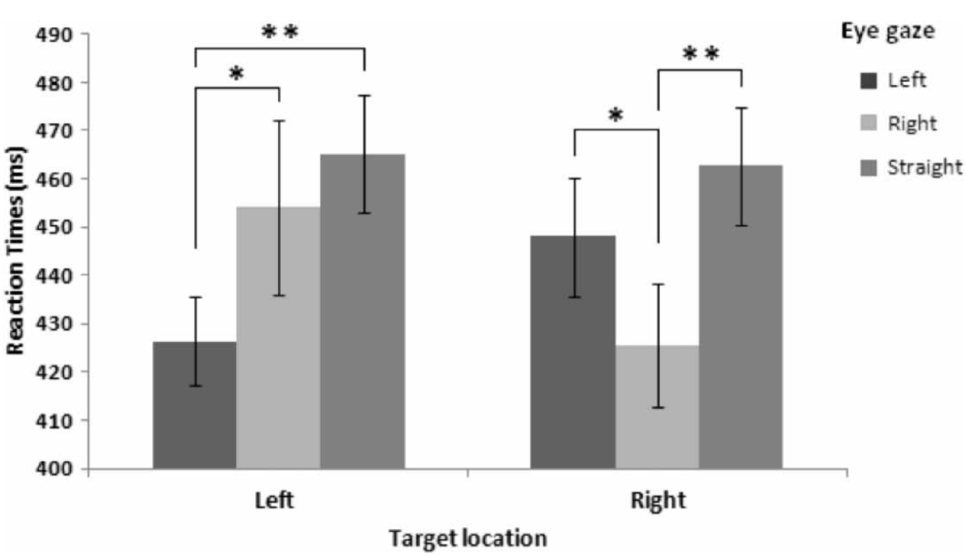

Figure 7. *Experiment 3. Mean reaction times (RTs) in left, right, and straight (no-gaze) conditions (bars are* SEM*) plotted against left and right target locations (i.e., significant interaction between eye gaze and target location revealing the gaze-cueing effect). *Significant at* $p < .048$. ***Significant at* $p < .001$.

these authors found that the mere presence of an actor in the scene encouraged many participants to take the actor's point of view, they suggested that the relevance of the actor in the participant's interpretation would be increased if the actor is perceived by the onlooker as "potentially interacting" with the target object. Similarly, understanding the actor's intentions while acting on the target would be facilitated by taking that person's perspective. Tversky and Hard did not find differences between the effect of the looking versus the looking/acting condition and suggested that other people's action and gaze are equivalent in their "potential interaction meaning". Such findings are in apparent contrast with the present data, but the significant effect of the actor's gaze in Tversky and Hard's looking condition might be ascribed to the fact that the actor both gazed and turned his head towards the target, since head orientation might enhance the "potential interaction" of the looking condition. This explanation could be consistent with Langton and colleagues' studies (Langton, 2000; Langton & Bruce, 1999; see also Langton & Bruce, 2000) demonstrating a strong facilitation effect on the onlooker's attentional orienting when the directions of eye gaze, head, and body were congruent. However, Hietanen (1999) manipulated both gaze direction and head orientation in a cueing paradigm and found that gaze affected reflexive orienting of attention only when eye and head direction did not coincide. Analogous results were obtained when body orientation was manipulated too (Hietanen, 2002). Although we are aware that the dynamic interactions among different bodily cues need to be further explored, this issue was outside the main aims of the present study. Here we tried to assess the contribution of the actor's gaze to visual perspective taking and social attention independently from clear directional cues provided by head orientation. Results demonstrated that eye gaze alone had no specific effect on perspective taking (Experiments 1 and 2), whereas the same eye gaze stimuli could effectively orient participants' joint attention (Frischen et al., 2007), as revealed Experiment 3. In this experiment, the gaze-cueing effect contrasted with the lack of significant influence exerted by the actor's grasping. Fischer and Szymkowiak (2004) compared the effect of observed pointing and grasping postures on joint attention and found that only pointing but not grasping led to attention shifts towards the potential target of an action. The authors suggested that while pointing is a deictic action signalling a target object of interest, grasping means that an action toward the target has already been performed, and the observer would not have to attend to that location any longer. Accordingly, Crostella and colleagues (2009) demonstrated that a (centrally presented) pointing hand can orient subjects' attention towards peripheral targets if subjects have to respond with congruent hand pointing movements. Crostella and colleagues' study differed from our joint-attention experiment for the kind of cue (isolated pointing hand versus actor's grasping) and for the response modality (pointing movement versus button press), and this might explain the divergences between findings.

Different behavioural data have highlighted the role of others' action potential in social attentional processes and visual perspective taking (Fischer et al., 2008; Fischer & Szymkowiak, 2004; Sartori, Becchio, & Castiello, 2011; Tversky & Hard, 2009). When interacting with other people it is important to understand what they are currently doing in order to predict what they will do next (C. D. Frith & Frith, 2006). We can make inferences about other persons' intentions by watching their movements; inferences start when the movement begins and are verified by predicting how the movement will continue (Wolpert, Doya, & Kawato, 2003). Such predictive skills strongly imply the ability to take another person's visual perspective and judge the actions of others from their own point of view (C. D. Frith & Frith, 2006). Therefore, shifting from one's own (egocentric) to other person's (allocentric) perspective could be triggered by the need for the onlooker to infer the other's intentions from their movements. This could further be related to selective attention processes: When participants interact with another person, they would select an object of interest by taking the agent's point of view and ignoring stimuli that the other person ignores (Frischen et al., 2009).

The human brain has developed specialized mechanisms to predict what others are about to do on the basis of eye, head, or body information (Fischer et al., 2008). In particular, another's eye gaze, head turns, pointing or grasping actions convey different information about the actions another person might perform next. Eye gaze would not be so relevant to understand what the actor is currently doing, since eye gaze would communicate that the gazed object is the target of the actor's attention, desires, or preferences (Bayliss, Paul, Cannon, & Tipper, 2006; Frischen et al., 2007), but without implying an ongoing interaction as the grasping implies. Actually, one could act upon the gazed objects in the immediate future or in the far future, or one could decide not to act upon it. Eye gaze is not critically involved, as grasping is, in communicating the actor's immediate intentions, and therefore it might not be effective in triggering allocentric perspective. The difference we found here between the effects of another's grasping and gaze on visual perspective taking could reflect the relevance of the information conveyed by these social cues on current (action) and future (eye gaze) intentions of another person.

Sartori and colleagues (2011) required participants to observe an actor gazing towards and grasping a wooden block with the intent to cooperate with a partner, compete against an opponent, or perform an individual action. Results showed that participants could discriminate across different intentions using information gained during viewing of the initial phase of the actor's action. Moreover, the authors found that the actor's arm and face conveyed information on different intentions: Seeing the arm was more helpful than seeing the face for judging whether the movement was performed at natural or fast speed (what the actor is currently doing), whereas seeing the face was more helpful than seeing the arm for discriminating the future intent of the action. Borrowing Sartori and colleagues' framework, we can suggest that arm cues are more effective in communicating what the actor is currently doing whereas eye gaze is more effective in communicating what the actor will do next.

Neurofunctional studies have demonstrated that an observer can attribute intentionality to others' actions in a different way depending on the combination of hand/limb movements, gaze direction, head orientation, and body posture. For instance, neuroimaging data on healthy humans comparing passive viewing of eyes, mouth, and hand movements demonstrated activations in different portions of the superior temporal sulcus, depending on the effector involved in action (Pelphrey, Morris, Michelich, Allison, & McCarthy, 2005). Different components of the intraparietal sulcus are also differentially engaged when watching an actor orient towards an object with different effectors. Goal-directed hand actions engage anterior intraparietal sulcus (Hamilton & Grafton, 2006), while gaze towards a particular object engages posterior intraparietal sulcus (Ramsey, Cross, & Hamilton, 2011). Evidence of separate neural mechanisms involved in processing another's action and gaze would support the distinction we showed here between the effect of another's action and eye gaze on visual perspective taking and joint attention. However, specific neuroimaging studies employing the present behavioural paradigm would be necessary to define the neural structures involved in processing different bodily cues when adopting the other's person point of view.

To conclude, the present paper provides novel behavioural evidence that observing another person's hand actions leads to increased visual perspective taking, while observing another person's eye gaze leads to changes in visual attention. These results demonstrate effector-specific influence of action observation and novel links between the perception of action and the ability to take another person's perspective in space.

Original manuscript received 7 September 2011
Accepted revision received 8 May 2012
First published online 20 August 2012

REFERENCES

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13,* 379–398.

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13,* 1061–1066.

Burton, A. M., Bindemann, M., Langton, S. R., Schweinberger, S. R., & Jenkins, R. (2009). Gaze perception requires focused attention: Evidence from an interference task. *Journal of Experimental Psychology: Human Perception and Performance*, *35,* 108–118.

Castiello, U. (2003). Understanding other people's actions: Intention and attention. *Journal of Experimental Psychology: Human Perception and Performance*, *29,* 416–430.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76,* 893–910.

Crostella, F., Carducci, F., & Aglioti, S. M. (2009). Reflexive social attention is mapped according to effector-specific reference systems. *Experimental Brain Research*, *197,* 143–151.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, *3,* 172–178.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Shared attention and the social brain: Gaze perception triggers automatic visuospatial orienting in adults. *Visual Cognition*, *6,* 509–540.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40,* 760–768.

Fischer, M. H., Prinz, J., & Lotz, K. (2008). Grasp cueing shows obligatory attention to action goals. *Quarterly Journal of Experimental Psychology*, *61,* 860–868.

Fischer, M. H., & Szymkowiak, A. (2004). Joint attention for pointing but not grasping postures. *Cortex*, *40,* 168–170.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology*, *17,* 99–103.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133,* 694–724.

Frischen, A., Loach, D., & Tipper, S. P. (2009). Seeing the world through another person's eyes: Simulating selective attention via action observation. *Cognition*, *111,* 212–218.

Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Research*, *1079,* 36–46.

Frith, U., & de Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Consciousness and Cognition*, *14,* 719–738.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8,* 396–403.

Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12,* 1–19.

Hamilton, A. F., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience*, *26,* 1133–1137.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51,* 59–67.

Hietanen, J. K. (1999). Does your gaze direction and head orientation shift my visual attention? *Neuroreport*, *10,* 3443–3447.

Hietanen, J. K. (2002). Social attention orienting integrates visual information from head and body orientation. *Psychological Research*, *66,* 174–179.

Hogg, R. V., & Craig, A. T. (1995). *Introduction to mathematical statistics* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

Langton, S. R. H. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology*, *53A,* 825–845.

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6,* 541–567.

Langton, S. R. H., & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26,* 747–757.

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & Psychophysics*, *68,* 327–337.

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, *15,* 1866–1876.

Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York, NY: Norton.

Ramsey, R., Cross, E. S., & Hamilton, A. F. (2011). Eye can see what you want: Posterior intraparietal sulcus encodes the object of an actor's gaze. *Journal of Cognitive Neuroscience*, *23,* 3400–3409.

Ricciardelli, P., Baylis, G., & Driver, J. (2000). The positive and negative of human expertise in gaze perception. *Cognition*, *77,* 1–14.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36,* 1255–1266.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, *119,* 242–252.

Sommerville, J. A., & Decety, J. (2006). Weaving the fabric of social interaction: Articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. *Psychonomic Bulletin and Reviews*, *13,* 179–200.

Tipper, S. P. (2010). From observation to action simulation: The role of attention, eye-gaze, emotion, and body state. *Quarterly Journal of Experimental Psychology*, *63,* 2081–2105.

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110,* 124–129.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *358,* 593–602.