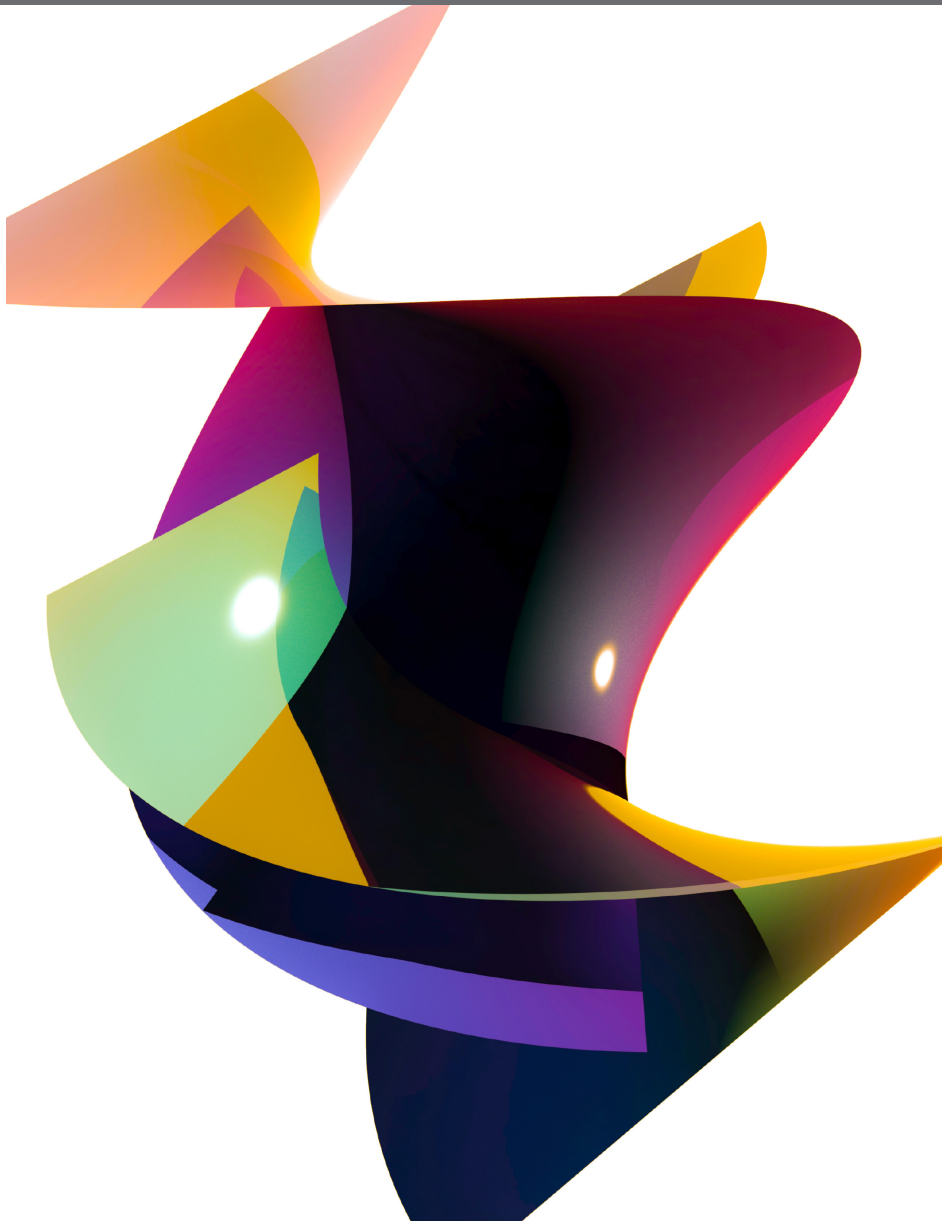


REACHING TO GRASP COGNITION: ANALYZING MOTOR BEHAVIOR TO INVESTIGATE SOCIAL INTERACTIONS

EDITED BY: Claudia Gianelli and Maurizio Gentilucci
PUBLISHED IN: Frontiers in Psychology





frontiers

Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-600-0

DOI 10.3389/978-2-88945-600-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

REACHING TO GRASP COGNITION: ANALYZING MOTOR BEHAVIOR TO INVESTIGATE SOCIAL INTERACTIONS

Topic Editors:

Claudia Gianelli, University of Potsdam, Germany

Maurizio Gentilucci, University of Parma, Italy

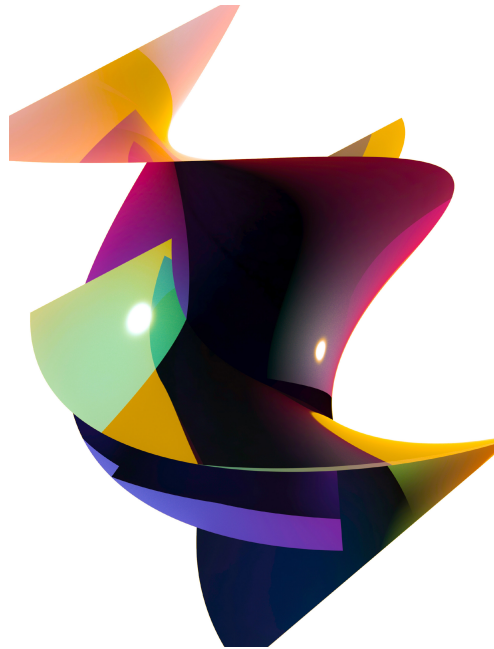


Image: andrey_l/Shutterstock.com

Citation: Gianelli, C., Gentilucci, M., eds. (2018). Reaching to Grasp Cognition: Analyzing Motor Behavior to Investigate Social Interactions. Lausanne: Frontiers Media.
doi: 10.3389/978-2-88945-600-0

Table of Contents

- 05 Editorial: Reaching to Grasp Cognition: Analyzing Motor Behavior to Investigate Social Interactions**

Claudia Gianelli and Maurizio Gentilucci

SECTION 1

SHARED REPRESENTATIONS BETWEEN ACTION AND PERCEPTION

- 07 Motor Interference in Interactive Contexts**

Eris Chinellato, Umberto Castiello and Luisa Sartori

- 16 Different but Complementary Roles of Action and Gaze in Action Observation Priming: Insights From Eye- and Motion-Tracking Measures**

Clément Letesson, Stéphane Grade and Martin G. Edwards

INDIVIDUAL, LINGUISTIC AND CULTURAL DIFFERENCES

- 28 Individual Differences in Reading Social Intentions From Motor Deviants**

Daniel Lewkowicz, Francois Quesque, Yann Coello and Yvonne N. Delevoye-Turrell

- 40 Factors Affecting Athletes' Motor Behavior After the Observation of Scenes of Cooperation and Competition in Competitive Sport: The Effect of Sport Attitude**

Elisa De Stefani, Doriana De Marco and Maurizio Gentilucci

- 49 Grasping the Agent's Perspective: A Kinematics Investigation of Linguistic Perspective in Italian and German**

Claudia Gianelli, Michele Marzocchi and Anna M. Borghi

- 62 The Multilingual CID-5: A New Tool to Study the Perception of Communicative Interactions in Different Languages**

Valeria Manera, Francesco Iani, Jérémy Bourgeois, Maciej Haman, Łukasz P. Okruszek, Susan M. Rivera, Philippe Robert, Leonhard Schilbach, Emily Sievers, Karl Verfaillie, Kai Vogeley, Tabea von der Lühne, Sam Willems and Cristina Becchio

- 71 Fifteen-Month-Old Infants use Velocity Information to Predict Others' Action Targets**

Janny C. Stapel, Sabine Hunnius and Harold Bekkering

SECTION 2

EXPERIMENTAL SETTINGS AND STIMULI

- 81 Automatic Imitation in a Rich Social Context With Virtual Characters**

Xueni Pan and Antonia F. de C. Hamilton

- 91 Video Stimuli Reduce Object-Directed Imitation Accuracy: A Novel Two-Person Motion-Tracking Approach**

Arran T. Reader and Nicholas P. Holmes

SECTION 3

SHARED GOALS, COMPLEMENTARY ACTIONS

105 *Social Cues to Joint Actions: The Role of Shared Goals*

Lucia M. Sacheli, Salvatore M. Aglioti and Matteo Candidi

112 *Complementary Actions*

Luisa Sartori and Sonia Betti

118 *Grasping Actions and Social Interaction: Neural Bases and Anatomical Circuitry in the Monkey*

Stefano Rozzi and Gino Coudé



Editorial: Reaching to Grasp Cognition: Analyzing Motor Behavior to Investigate Social Interactions

Claudia Gianelli^{1*} and Maurizio Gentilucci²

¹ Universität Potsdam, Potsdam, Germany, ² Università degli Studi di Parma, Parma, Italy

Keywords: kinematics, social cognition, action observation, imitation, joint action, complementary actions, cooperation and competition, embodied cognition

Editorial on the Research Topic

Reaching to Grasp Cognition: Analyzing Motor Behavior to Investigate Social Interactions

INTRODUCTION

Action planning and execution have always been fascinating topics for neuroscience and psychology. In particular, kinematics studies have contributed to shed light on how very basic actions (e.g., reaching-grasping) are affected by manipulating target properties, visually or linguistically presented stimuli and contextual information. Interestingly, recent studies have also shown how the social context in which actions take place and their relevance for human interactions can also affect action execution.

This Research Topic brought together researchers studying socially relevant aspects of cognition (e.g., action observation, imitation, joint and complementary actions) with a wide range of methodologies and theoretical points of view.

Altogether, their contributions carefully represent the status of the field and foresee future developments.

SHARED REPRESENTATIONS BETWEEN ACTION AND PERCEPTION

The topic of shared representations between perceived and executed actions is largely present throughout this Research Topic.

Chinellato et al. compared the effects of object-oriented hand actions on motor responses in interactive and non-interactive conditions. Their results showed that a socially relevant condition is quickly taken into account by the motor system, producing an overall slowdown of the motor responses (interference effect). Interestingly, this suggests that the emergence of an interference effect is affected not only by the motor properties of the observed action but also by the available social information.

Letesson et al. approached action observation from the point of view of action priming, i.e., facilitation of motor responses following observation.

Eye- and motion-tracking measures showed that agent's gaze and action kinematics are both contributing to the representation of the action goal, and they do so in a complementary manner. Interestingly, this suggests that—while capable of eliciting motor representations independently—combined gaze and action information produces a more refined action representation.

OPEN ACCESS

Edited and reviewed by:

Bernhard Hommel,
Leiden University, Netherlands

*Correspondence:

Claudia Gianelli
isotopia@gmail.com

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 15 May 2018

Accepted: 27 June 2018

Published: 14 August 2018

Citation:

Gianelli C and Gentilucci M (2018)
Editorial: Reaching to Grasp
Cognition: Analyzing Motor Behavior
to Investigate Social Interactions.
Front. Psychol. 9:1236.
doi: 10.3389/fpsyg.2018.01236

INDIVIDUAL, LINGUISTIC AND CULTURAL DIFFERENCES

The use of experimental paradigms including basic action stimuli and simple motor responses allowed for several possible extensions to diverse samples.

Lewkowicz et al. investigated how individual differences affect the successful detection of social intentions. Individual scores in standard questionnaires of imagery and social cognition were correlated with participants' ability to distinguish between actions with the same motor goal but different intentions (personal vs. social). Data showed that the ability to successfully distinguish personal and social actions by recognizing motor deviants is highly correlated with social cognition scores.

De Stefani et al. incorporated expert athletes' reported "attitude" (competitive vs. cooperative) into a classic action observation design which used cooperative and competitive actions. Their data showed that the expected facilitation of motor responses arose when there was a match between attitude and intended action, but only in the case of cooperative participants. On the contrary, competitive participants were overall faster but without showing significant differences between conditions.

Gianelli et al. extended the investigation of action observation with the use of linguistically described actions and reported the results from experiments in two languages. This cross-linguistic approach showed different motor effects in the two languages. While this study did not include any form of social conditions, a similar paradigm (linguist stimuli + motor responses) could be usefully extended to more interactive contexts.

Manera et al. took a cross-cultural approach by presenting a multilingual database to investigate non-conventional communicative gestures across different languages. The authors created a large set of point-light displays reproducing communicative interactions between two agents and single-agent non-communicative actions. Results from testing this set across seven languages, show that the proposed stimuli were correctly recognized as communicative or individual based on the information available in point-light displays.

The paper by Stapel et al. provided further evidence of the possibilities of extending action the investigation to diverse samples. In this case, they tested infants of various ages (9-, 12-, 15-months old) and adults in an eye-tracking experiment investigating the observer's ability to predict the target of an action based on the velocity of natural object-directed actions. Indeed, the authors showed that, as soon as 15-months of age (but not 9 and 12), infants start using velocity information to predict the outcome of observed actions, similarly to what adults do.

EXPERIMENTAL SETTINGS AND STIMULI

A crucial aspect of investigating social interactions is the possibility to use controlled yet realistic settings.

Pan and Hamilton extended the investigation of action observation to the use of Virtual Characters (VCs) and sequential hand-arm actions. Comparing the response to VCs actions with the same actions sequence indicated by virtual balls, they showed

that participants automatically imitated the actions of the VCs, but not those implied by the virtual balls. VCs might thus be a powerful tool for the study of imitation, providing a richer social context compared to the use of isolated pictures or videos.

Reader and Holmes also questioned the use of video-stimuli when measuring action imitation.

By using a novel motion-tracking paradigm, the authors showed that accuracy was worsened in the case of video stimuli, as compared to face-to-face interaction. Based on these results, the authors suggest that previously reported effects might have been biased by the use of these stimuli and the potential limitations of video stimuli should be further investigated.

SHARED GOALS, COMPLEMENTARY ACTIONS

Sacheli et al. reviewed how realistic contexts could integrate a manipulation of the interpersonal cognitive/emotional dimension to investigate the impact of these factors on interacting behaviors. The authors presented and discussed the possibilities of a novel joint-grasping task, which allow for disentangling individual and shared goals in controlled yet naturalist contexts.

Sartori and Betti took a similar approach while reviewing existing evidence on complementary actions, i.e., forms of interaction where agents have to adapt their individual actions to a shared goal, e.g., one individual's action complete the one of the other in order to achieve a common aim.

Finally, Rozzi and Coudé provided an extensive discussion of the neural bases of these processes reporting current anatomic-functional evidence regarding how contextual information could be integrated into an extended motor network. This network might subserve basic social functions, such as allowing two individuals to perform complementary actions toward a shared goal.

AUTHOR CONTRIBUTIONS

CG and MG discussed and agreed upon the content of this Editorial. CG drafted the manuscript and approved its final version.

ACKNOWLEDGMENTS

MG, a pioneer in motor control studies and also in their extension to the social domain, left us before our joint editorial work was completed. This Research Topic is dedicated to his memory.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Gianelli and Gentilucci. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Motor interference in interactive contexts

Eris Chinellato¹, Umberto Castiello^{2,3,4*} and Luisa Sartori^{2,3*}

¹ School of Computing, Faculty of Engineering, University of Leeds, Leeds, UK, ² Dipartimento di Psicologia Generale, Università di Padova, Padova, Italy, ³ Cognitive Neuroscience Center, University of Padova, Padova, Italy, ⁴ Centro Beniamino Segre, Accademia Nazionale dei Lincei, Rome, Italy

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Inge Volman,
University College London, UK
Francois Quesque,
University of Lille, France

*Correspondence:

Umberto Castiello and Luisa Sartori,
Dipartimento di Psicologia Generale,
Università di Padova, Via Venezia 8,
35131 Padova, Italy
umberto.castiello@unipd.it;
luisa.sartori@unipd.it

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 19 March 2015

Accepted: 26 May 2015

Published: 11 June 2015

Citation:

Chinellato E, Castiello U and Sartori L
(2015) Motor interference
in interactive contexts.
Front. Psychol. 6:791.
doi: 10.3389/fpsyg.2015.00791

Action observation and execution share overlapping neural substrates, so that simultaneous activation by observation and execution modulates motor performance. Previous literature on simple prehension tasks has revealed that motor influence can be two-sided: facilitation for observed and performed congruent actions and interference for incongruent actions. But little is known of the specific modulations of motor performance in complex forms of interaction. Is it possible that the very same observed movement can lead either to interference or facilitation effects on a temporally overlapping congruent executed action, depending on the context? To answer this question participants were asked to perform a reach-to-grasp movement adopting a precision grip (PG) while: (i) observing a fixation cross, (ii) observing an actor performing a PG with interactive purposes, (iii) observing an actor performing a PG without interactive purposes. In particular, in the interactive condition the actor was shown trying to pour some sugar on a large cup located out of her reach but close to the participant watching the video, thus eliciting in reaction a complementary whole-hand grasp. Notably, fine-grained kinematic analysis for this condition revealed a specific delay in the grasping and reaching components and an increased trajectory deviation despite the observed and executed movement's congruency. Moreover, early peaks of trajectory deviation seem to indicate that socially relevant stimuli are acknowledged by the motor system very early. These data suggest that interactive contexts can determine a prompt modulation of stimulus–response compatibility effects.

Keywords: action observation, interference effect, movement kinematics, complementary actions

Introduction

Human beings spend most of their time interacting with others. But despite interest, relevance, and theoretical development on how people represent their own and other person's actions, there is still a considerable lack of understanding of the precise cognitive mechanisms governing interactive performance. At least part of this remarkable gap is due to the fact that several paradigms have typically relied on single individuals passively observing or imitating other individuals. In contrast, when engaging in interactive contexts, individuals are often required to perform *complementary* parts of a given action, i.e., completing each other's movement in a balanced manner rather than acting in the same manner (Sebanz et al., 2006; Sartori et al., 2013a). How and when one's own action execution is influenced by other's actions execution during social interactions is just beginning to be understood. A large amount of behavioral

(e.g., Pfister et al., 2014) as well as neurophysiological studies (e.g., Baess and Prinz, 2015) is providing consistent evidence for the existence of shared representations between action and perception – within and between individuals (Adolphs, 2003; Sebanz et al., 2005). According to the action co-representation account, human agents represent their coactor's task, and this can *facilitate* action prediction and coordination with others (for a review, see Knoblich et al., 2011).

Others have argued that since one's own action and the actions of another person are represented in the same way (Hommel, 2009, 2011), actively representing our own and another person's actions can create a *conflict* between concurrently activated representations (Schubö et al., 2001; Dolk et al., 2014). Concurrent activation during action selection would then produce a discrimination problem, leading participants to emphasize the features that best distinguish selected responses. This implies that more similarity between observed and executed action would put more emphasis on the discriminating features, leading to increased reaction times (RTs) with every extra feature dimension that event-coding processes consider.

The most prominent cognitive paradigms that have been adopted to test these hypotheses in single and joint settings are based on the principle of Stimulus–Response Compatibility (SRC). The term SRC commonly refers to the finding that a compatible mapping of stimulus and response position is associated with shorter RTs as compared to longer RTs due to incompatible mapping (Fitts and Seeger, 1953).

In kinematic terms as well, it has been shown that observing the movements of others can either facilitate or interfere with concurrent movement execution, depending on observed and executed movement congruency (Brass et al., 2000; Castiello et al., 2002; Edwards et al., 2003; Stürmer and Leuthold, 2003; Bertenthal et al., 2006; Liepelt et al., 2010; Hardwick and Edwards, 2012; for reviews, see Blakemore and Frith, 2005; Wilson and Knoblich, 2005). In other words, observing a movement primes the execution of that movement, thereby interfering with the execution of another movement (*motor priming*). Behavioral research on motor priming has shown that responses to human hand movement stimuli (e.g., a hand opening) are faster and more accurate when they involve execution of the same movement (e.g., hand opening) than when they involve execution of an alternative movement (e.g., hand closing; Sturmer et al., 2000). Similarly, if the subjects are instructed to perform a finger tapping in response to a finger tapping (compatible) or lifting (incompatible), the RT to initiate the prepared movement significantly slows down when the stimulus is incompatible (Brass et al., 2001). This effect is thought to be an index of perceptual-motor matching and has been replicated featuring diverse stimulus displays (e.g., grasping, pointing, hand, and arm movements; Kilner et al., 2003) and a variety of stimulus-response arrangements, emphasizing not only the role of perception on concurrent action, but also the influence of movement production over motion perception (Müsseler et al., 2000; Craighero et al., 2002; Hamilton et al., 2004; Schubö et al., 2004; Zwickel et al., 2010; Christensen et al., 2011).

Most importantly for the issue at stake here, Liepelt et al. (2010) found a *reversed compatibility* effect when observing a

human extending the right hand for a handshake. When viewing a right-handed shake-hands image, participants responded faster with their own right hand, instead of mirroring the stimulus hand. Notably, we usually shake an extended right hand with our right hand, leading to spatial incompatibility of the relative position of the hand (see also Flach et al., 2010). This reversal of the classic compatibility effect is not surprising in the light of recent finding emphasizing the idea of a complementary action system (Sebanz et al., 2006; Newman-Norlund et al., 2007a,b, 2008; van Schie et al., 2008; Sartori et al., 2013a). It strongly indicates that the overlearned response to extend the right hand when observing a right hand is able to modulate the motor priming effect: when a specific behavior is contingent on a non-matching behavior, an incongruent association is formed (Catmur et al., 2009), so that social response preparation can overwhelm the automatic response (Hamilton, 2013).

The purpose of the present study was to further this line of investigation by exploring how the context specifically modulates actions under 'complementary' conditions. The following experiment addressed this issue by adopting ecologically valid stimuli: (i) requiring a specific complementary response (i.e., functionally related to the observed action), (ii) temporally overlapping with the participants' ongoing action, and (iii) depicting familiar object-oriented hand actions, given that motor familiarity with the observed action is thought to be positively related to the mapping between observed and executed actions (e.g., Calvo-Merino et al., 2005; Cross et al., 2006). We capitalized on an established paradigm for inducing complementary activations in the observers' muscles (Sartori et al., 2011b, 2012, 2013b,c). In one of these studies (Sartori et al., 2013c), participants watched videos of action sequences showing an actor pouring sugar with a tablespoon, grasped with a precision grip (PG), into a set of cups. At the start of the videos, participants showed a small activation in the little finger muscle, consistent with the actor's actions. The key manipulation came when the actor stretched the arm toward the last cup, which was placed close to the participant. The socially appropriate response would require to pick up the cup by using a whole hand grip, and offer it to the actor. At this point, the observers' muscular activations changed, with a large response of the little finger muscle even though the actor in the video maintained the same grip and the participant did not perform any actual response. In the present study, we had participants observing the video of an actor grasping a tablespoon and then stretching toward a cup which was placed close to the participant (interactive action). In another video, the same actor was shown pouring some sugar and then simply coming back to the starting point (non-interactive action). Crucially, the task was to simultaneously observe these perceptual events and perform a congruent prehension (i.e., a PG). Observed and executed action features were thereby maintained compatible across both conditions. By introducing the complementary request by the actor, we expected nonetheless to find an increase of variance in movement trajectory while *planning* an incompatible movement, in line with previous studies demonstrating that trajectory deviations increase when an object is grasped with the intention to interact with a human

agent (Becchio et al., 2008a,b; Quesque et al., 2013; Quesque and Coello, 2014). A control condition was also set, in which participants simply observed a fixation cross while performing the task.

Interestingly, previous TMS studies showed *early* changes in observers' cortico-spinal excitability during observation of hand actions leading to a complementary request (Sartori et al., 2013b,c). More specifically, the changes in cortico-spinal excitability were modulated by early kinematic changes in the observed movement signaling the start of the social request.

Accordingly, in the present experiment we synchronized the 'go' signal with the start of the social request (i.e., the time at which the actor finished pouring sugar into the close cup, just before she began stretching her arm toward the out-of-reach cup). We reasoned that if the observer can easily predict the future course of the observed action from the actor's kinematics, then an *early* motor interference effect should occur on his/her action. In particular, we expected to find a prompt response for the interactive condition in terms of arm trajectory deviation. This would confirm results from a previous kinematic study indicating that socially relevant stimuli are acknowledged by the motor system very early (Sartori et al., 2009). Moreover, since response competition involves inhibition, here we expected to see increased inhibition in the interactive condition, regardless of the fact that the same type of grasp was observed.

Materials and Methods

Participants

Fifteen volunteers (nine females and six males, between the ages of 21 and 27) with normal or corrected-to-normal vision participated in the experiment. All the participants were right-handed (Briggs and Nebes, 1975), reported normal or corrected-to-normal visual acuity, and were naive as to the experimental purpose of the study. A right-handed non-professional actor (female, 28 years-old) was also recruited for video clips recording. All participants gave their informed written consent to participate in the study. The experimental procedures were approved by the Institutional Review Board at the University of Padua and were in accordance with the Declaration of Helsinki (Sixth revision, 2008).

Stimuli

The stimuli were two digitally recorded video clips showing the actor: (i) pouring sugar with a tablespoon (PG) in a cup located nearby, and then stretching out her arm trying to pour some sugar on a large cup located out of her reach (interactive action; **Figure 1D**), (ii) pouring sugar in the same cup, and then coming back to the starting point (non-interactive action). Crucially, the out-of-reach cup was placed in the video foreground, closer to the participant watching the video, thus eliciting a complementary reaction with a whole-hand grip when the actor was trying to reach for it. All of the videos were taken from a frontal view and were equal in length. At the beginning of each video-clip the hand of the actor was shown in a prone position resting on the table. The model started her reach-to-grasp movement 1 s

later and her fingers made contact with the sugar spoon at 4.9 s. The model finished pouring sugar into a close cup 5.8 s after the onset of the video in the interactive condition and 5.9 s in the non-interactive condition. For the participants' prehension task we adopted a small plastic fork (130 mm length, the same size as the sugar spoon in the videos). An affixed colored dot was signaling which part of the object had to be grasped in order to perform a stable and consistent PG, congruent to the actor's movement. We choose a fork instead of a spoon to avoid eliciting in the participant the idea of pouring something into the cup – instead of grasping the cup – during the interactive condition. Since gaze is a crucial component of social interactions and could have biased the results, the face of the actor was not visible on the video clips. Eye gaze, in fact, may enhance observers' abilities in predicting and anticipating others' actions (e.g., Sartori et al., 2011a).

Procedure

The experimental set up is depicted in **Figure 1**. Each participant sat on a height-adjustable chair in front of a table (900 mm × 900 mm) with the elbow and wrist resting on the table surface and the right hand in the designated position. The hand was pronated with the palm resting on a starting platform (60 mm × 70 mm; 5 mm thick), which was shaped to allow for a comfortable and repeatable posture of all digits, i.e., slightly flexed at the metacarpal and proximal interphalangeal joints. The starting platform was attached 90 mm away from the edge of the table surface 50 mm away from the midsection. The fork was placed on a target platform (10 cm × 10 cm; 5 mm thick), located at a distance of 350 mm from the starting platform, with the handle pointing slightly rightward (i.e., with an angle of 30° with respect to the midsection) to allow for an accurate prehension. The participants had to execute a reach-to-grasp movement with a PG toward the fork placed on the table and to watch the video clips that were presented on a 19" monitor (resolution 1280 × 1024 pixels, refresh frequency 75 Hz, background luminance of 0.5 cd/m²) set at eye level (the eye-screen distance was 60 cm). The experiment included three different conditions:

- Execution-only (control condition): the participants performed their task while observing a fixation cross on the screen.
- Interactive condition: the participants performed their task while observing the video showing an actor performing a PG, and then stretching out her arm trying to pour some sugar on a large cup located close to them, thereby inviting them to grasp it.
- Non-interactive condition: the participants performed their task while observing the video showing an actor performing a PG, and then coming back to the starting point. Notably, in this video the out of reach cup was still visible in the video foreground, therefore controlling for possible affordance effects.

Each condition was presented 15 times in random order. In total, the experiment was composed of 45 trials, each lasting

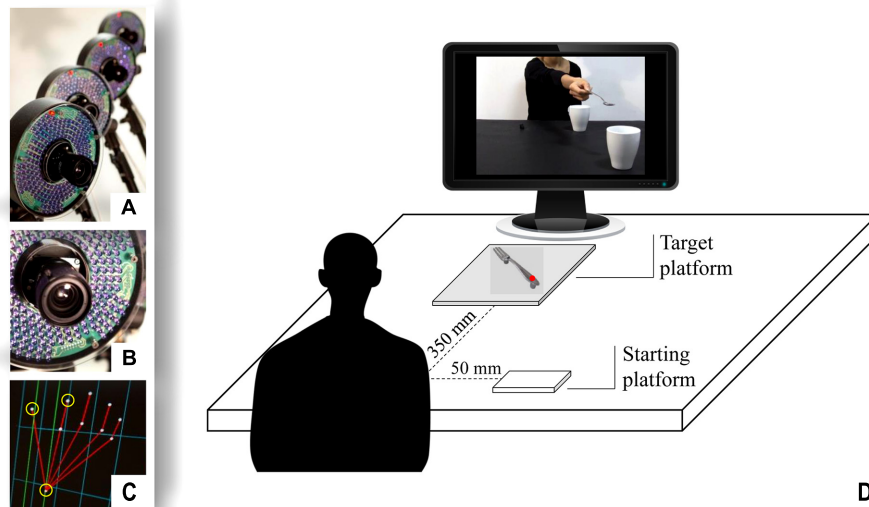


FIGURE 1 | Set up and procedure. A 3D-Optoelectronic SMART-D system was used to track the kinematics of the participant's right upper limb by means of six video cameras (A,B). Infrared reflective markers were taped to the following points: thumb, index finger, and wrist to measure the grasp and reach component of the movement (C). Participants sat in front of a monitor (D).

A starting platform was attached 50 mm away from the midsection and the object was placed on a target platform located at a distance of 350 mm. During the interactive condition, the video was showing an actor stretching out her arm trying to pour some sugar on a large cup located close to the participants, thereby inviting them to grasp it.

approximately 9 s. Participants were asked to look at the actor's hand throughout the trials and were instructed to begin their movements as soon as an acoustic 'go' signal switched on ('Go' instruction). The 'go' signal was released 5.8 s after the onset of each video (i.e., the time at which the actor finished pouring sugar into the closely located cup during the interactive condition). Since different attention effects due to different cognitive load in different conditions could have biased our data, we presented the 'go' signal when participants were observing the very same gesture (i.e., pouring sugar into the close cup) instead of synchronizing it with the end of the videos. No instruction was given concerning the speed of the movement, and participants were asked to perform their movement at their own pace.

Kinematics Recording

A 3D-Optoelectronic SMART-D system (Bioengineering Technology and Systems, B| T| S) was used to track the kinematics of the participant's right upper limb. Two lightweight infrared reflective markers (0.25 mm in diameter; B| T| S) were placed on each participant's hand to measure the grasping component of the action and one marker was placed on the wrist to measure the reaching component of the action (Figure 1C, yellow circles). In particular, the three infrared reflective markers were taped to the following points: (i) thumb (ulnar side of the nail); (ii) index finger (radial side of the nail); and (iii) wrist (dorsodistal aspect of the radial styloid process). Six video cameras (sampling rate 140 Hz) detecting the markers were placed in a semicircle at a distance of 1–1.2 m from the table (see Figure 1). The camera position, roll angle, zoom, focus, threshold, and brightness were calibrated and adjusted to optimize marker tracking before the trials were

begun. Static and dynamic calibration was then carried out. For the static calibration, a three-axes frame of five markers at known distances from each other was placed in the middle of the table. For the dynamic calibration, a three-marker wand was moved throughout the workspace of interest for 60 s. The spatial resolution of the recording system was 0.3 mm over the field of view. The SD of the reconstruction error was below 0.2 mm for the x, y, and z axes.

Data Processing

Following data collection, each trial was individually checked for correct marker identification and the SMART-D Tracker software package (B| T| S) was used to provide a 3-D reconstruction of the marker positions as a function of time. The data were then filtered using a finite impulse response linear filter (transition band = 1 Hz, sharpening variable = 2, cut-off frequency = 10 Hz; D'Amico and Ferrigno, 1990, 1992). The measurements were made along the three Cartesian axes [i.e. X (left–right), Y (up–down), and Z (anterior–posterior) axes] of the participants in an upright sitting position. Movement onset was defined as the time at which the tangential velocity of the wrist marker crossed a threshold (5 mm/s) and remained above it for longer than 500 ms. End of movement was defined as the time at which the hand made contact with the object and quantified as the time at which the hand opening velocity crossed a threshold (5 mm/s) after reaching its minimum value and remained above it for longer than 500 ms. The following kinematic parameters were extracted for each individual movement using a custom protocol run in Matlab, 2014b, (The 4 MathWorks, Natick, MA, USA): the time interval between movement onset and end of grasping (Movement Time), the time at which the tangential

velocity of the wrist was maximum from movement onset (Time to Peak Wrist Velocity), the time at which the distance between the 3D coordinates of the thumb and index finger was maximum between movement onset and hand contact time (Time to Maximum Grip Aperture), the time at which the tangential velocity of the 3D coordinates of the thumb and index finger was maximum from movement onset (Time to Maximum Grip Velocity), and the maximum distance reached by the 3D coordinates of the thumb and index finger (Maximum Grip Aperture). Grip aperture was calculated at 25, 50, and 75% of the movement to assess during which part of the movement possible interference may occur. In addition, wrist trajectories were computed for each condition, by normalizing each trial according to movement time, so that they were reduced to the same number of time-steps (420). We then considered a spatial trajectory measure that has been proved to be sensitive to variations in social context: the direction, amplitude, and time course of the distance of the trajectory path from a straight line linking the starting position and the object location (Trajectory Deviation; Sartori et al., 2009; Innocenti et al., 2012). For this measure, we gave a positive sign to right deviations and a negative sign to left deviations and we calculated values at 25, 50, 75, and 100% of the movement in both Cartesian distance and signed values. Moreover, temporal delay between the 'go' signal and movement onset was computed as RT.

Data Analysis

The mean value for each parameter of interest were determined for each participant and then entered into separate repeated-measures ANOVAs with Condition (execution-only, interactive, non-interactive) as within-subjects factor. Preliminary analyses were conducted to check for normality, sphericity (Mauchly test), univariate and multivariate outliers, with no serious violations noted. Main effects were used to explore the means of interest (*post hoc t-test*), and Bonferroni's corrections (alpha level of $p < 0.05$) were applied.

Results

All the results are displayed in **Table 1**. For the sake of clarity, only parameters differing with respect to interactive vs. non-interactive conditions will be reported. Notably, the fragment of

video clip displayed before the go signal (i.e., pouring sugar with a PG) was the same in both these conditions.

Reaction Time

A repeated-measures ANOVA revealed a significant effect of condition [$F_{(2,28)} = 5.80, p = 0.008, \eta_p^2 = 0.29$]. *Post hoc* analyses showed that the RT was shorter in the execution-only condition compared to the interactive ($p = 0.02, \eta_p^2 = 0.31$) and non-interactive conditions ($p = 0.04, \eta_p^2 = 0.27$). Moreover, it was more delayed in the interactive condition compared to the non-interactive condition ($p = 0.05, \eta_p^2 = 0.23$).

Movement Time

The ANOVA performed on movement time revealed a significant effect of condition [$F_{(2,28)} = 5.72, p = 0.008, \eta_p^2 = 0.29$]. Observing an interactive action did influence movement performance with respect to the execution-only condition, leading to an increase in the execution time ($p = 0.01, \eta_p^2 = 0.43$). This effect was significant also for the non-interactive condition ($p = 0.03, \eta_p^2 = 0.31$).

Time to Maximum Grip Velocity

The ANOVA performed on the time of maximum grip velocity revealed a significant main effect of condition [$F_{(2,28)} = 10.01, p = 0.001, \eta_p^2 = 0.42$]. *Post hoc* analyses showed that peak grip velocity was reached earlier in the execution-only condition compared to the interactive ($p = 0.001, \eta_p^2 = 0.54$) and to the non-interactive ($p = 0.02, \eta_p^2 = 0.34$) conditions. Moreover, peak grip velocity was reached later in the interactive than in the non-interactive conditions ($p = 0.04, \eta_p^2 = 0.26$). In normalized terms, a 3% delay of peak velocity for the interactive with respect to the execution-only condition was found, whereas a 2% delay of peak velocity for the non-interactive condition was revealed.

Trajectory Deviation

The ANOVA performed on the distance of the trajectory path from an ideal straight line linking the starting position and the object location indicates that it was specifically modulated as a function of the condition [$F_{(2,28)} = 5.32, p = 0.01, \eta_p^2 = 0.28$]. A significant increase in trajectory deviation was detected for the interactive compared to the non-interactive ($p = 0.02; \eta_p^2 = 0.36$) and to the execution-only condition ($p = 0.001; \eta_p^2 = 0.33$). Notably, when considering the direction and time

TABLE 1 | Statistically significant key kinematic parameters and reaction times (RTs) across conditions.

	Execution-only	Interactive	Non-interactive
Movement time (ms)	1287.11 (± 61.01)	1339.21 (± 64.78)	1336.89 (± 62.88)
Time to peak wrist velocity (ms)	535.17 (± 29.78)	569.02 (± 31.70)	569.24 (± 30.28)
Time to peak grip aperture (ms)	681.67 (± 46.38)	734.67 (± 52.88)	713.18 (± 46.12)
Time to peak grip aperture (%)	62.70 (± 2.10)	64.70 (± 2.00)	64.20 (± 1.60)
Time to peak grip velocity (ms)	485.71 (± 40.36)	536.86 (± 44.44)	510.25 (± 36.89)
Trajectory deviation (mm)	5.17 (± 1.00)	5.68 (± 1.45)	5.08 (± 0.91)
Reaction Times (ms)	5437.09 (± 305.88)	6425.68 (± 184.49)	6093.11 (± 173.76)

Mean and standard errors per condition.

course of this effect, a statistically significant leftward deviation was detected within the first 25% of the movement for the interactive compared to the non-interactive condition (-1.96 vs. -1.72 , respectively; $p = 0.04$, $\eta_p^2 = 0.26$; **Figure 2**). No effect was found for trajectory deviation at 50, 75, and 100% of the movement ($p_s > 0.05$).

Maximum Grip Aperture

The ANOVA performed on the maximum aperture did not revealed any significant effect of condition [$F_{(2,28)} = 2.19$, $p > 0.05$, $\eta_p^2 = 0.14$]. However, when considering the time course of grip aperture, a significant decrease was detected for the interactive compared to the execution-only condition at 50% of movement time (41.72 vs. 43.58, respectively; $p = 0.01$, $\eta_p^2 = 0.36$). No effect was found for grip aperture at 25 and 75% of the movement ($p_s > 0.05$).

Discussion

Many daily activities involve performing an action while simultaneously encoding other perceptual events. This is particularly interesting when others' actions elicit a complementary response which differ from our ongoing action. The aim of this study was to determine what critical process underlies such mismatching conditions and how they affect the precision and performance of executed movements. Participants were asked to perform a PG with their right hand while concurrently observing a similar action, but requiring (or not) a complementary incongruent response. Our main finding is that although observed and executed action features were maintained compatible across conditions, an increase in RTs, Movement Time, Time to Maximum Grip Velocity, and Trajectory Deviation occurred for the interactive compared to the non-interactive condition, in line with previous studies

demonstrating a general delay in the grasping and reaching components and an increased trajectory deviation when an object is grasped with the intention to interact with a human agent (Becchio et al., 2008a,b; Sartori et al., 2009; Quesque et al., 2013; Quesque and Coello, 2014). The very fact that we found a prompt response for the interactive condition (deviation peak at 25% of the movement) indicates that the socially relevant action was acknowledged very early by the motor system (Sartori et al., 2009).

Reversing Classic Interference Effects

The common coding theory states that perception of an action leads to simulative production of that action on the part of the observer (Brass et al., 2001). But if the central motor system is perfectly tuned during the execution and concurrent observation of a congruent action, what happens when we are required to make a qualitatively different (incongruent) gesture? In this case, the motor program (or representation) associated with the incongruent movement interferes with both the outgoing motor output and the observed movement. And motor interference arises as a general delay in the grasping and reaching components and as an increase of variance in movement trajectory. This result confirms and extends previous findings reporting interference effects when simply observing incongruent moving stimuli presented either face-to-face or in video (Kilner et al., 2003). Moreover, it generalizes the results of Sartori et al. (2009) and Liepelt et al. (2010) showing that planning a complementary, functionally related action has the power to elicit associated responses and reverse classic interference effects. Depending on its posture and context, an extended hand can lead to a handshake or other actions, and this suggests that in our everyday interactions the automatic and rapid decoding of social cues influences our intentional behavior, maximizing the efficiency of our responses. It is widely accepted that during action observation, the specific networks subserving that particular movement are already tuned for action (Fadiga et al., 2005). But the present results demonstrate that even observing congruent stimuli presented on a video display can have a measurable interference effect on simultaneously executed actions, depending on the context. The precise nature of this effect depends on the type of action presented in the video stimuli, with interference found for observation of a complementary request, and to a less degree for a non-interactive action. A possible explanation for our data comes from the hypothesis of a competition between different representations (Schubö et al., 2001; Dolk et al., 2014). According to these authors, the representations that underlie perceptual and motor activities, such as producing a movement while concurrently encoding an independent stimulus motion, must be "kept separate" so that the two activities can be carried out without interfering. In our study, we found a different degree of motor interference on the latency of Time to Peak Grip Velocity ranging from the non-interactive (2% delay) to the interactive (3% delay) conditions, despite the observed and executed movements were similar. Interestingly, the higher interference on grip aperture was connected to the planning of a complementary movement, thus suggesting a higher degree of competition between different representations.

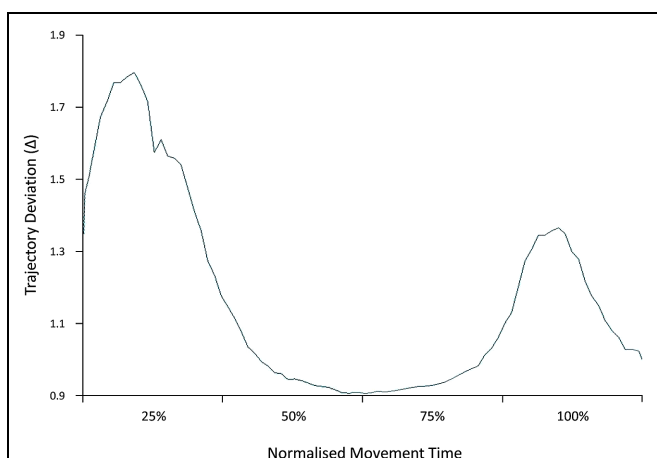


FIGURE 2 | Early trajectory deviation. Distance of the trajectory path for the interactive compared to the non-interactive condition is represented at 25, 50, 75, and 100% of the movement. A significant increase in trajectory deviation for the interactive condition was detected within the first 25% of the movement.

Response Competition and Inhibition Processes

In terms of grasping, an interference on the amplitude of Maximum Grip Aperture was specifically detected at 50% of movement execution. The direction of this effect (i.e., a decreased Grip Aperture) could be the byproduct of an automatic inhibition of representational features related to the complementary response (e.g., see Prinz and Hommel, 2002), in line with previous literature pointing to a bi-phasic pattern of interference of perception on ongoing action: initial assimilation followed by contrast (Dijkerman and Smit, 2007; Grosjean et al., 2009). Schubö et al. (2001) proposed that the representations that underlie distinct activities, such as producing a movement while concurrently encoding a perceptual event, must be “kept separate” so that the two activities can be carried out without interfering. The mechanism in question would involve a form of inhibition (Tipper et al., 1997) of the features shared by perception and action. Many models have thenceforth accounted for inhibition by referring to mechanisms associated with response competition (Swinnen, 2002; Duque et al., 2010, 2012; van den Berg et al., 2011; Verstynen and Ivry, 2011; Klein et al., 2012; Labruna et al., 2014). Since inhibition reflects the operation of a process involved in resolving response competition, here planning a whole-hand response had a repulsive effect on what was produced, decreasing the Maximum Grip Aperture and shifting the Trajectory path leftward (i.e., in the opposite direction with respect to the object requiring a whole-hand grasp).

Maximum Grip Aperture

In this study we did not expect any change in Maximum Grip Aperture, since the same movement (i.e., a PG) was always repeated within the task and grip amplitude is known to covary linearly with object size (Jeannerod et al., 1995). The results were in line with our expectations. *Post hoc* analyses revealed that grip aperture remained constant throughout the experiment and well calibrated to the object size. Interestingly, the interactive request did not result in greater uncertainty in the performance of the participant's grasping movement, since when subjects are uncertain during the grasp, they open their hand wider (Paulignan et al., 1991a,b). We may therefore assume that the preservation of maximal grip aperture across conditions is evidence that participants were confident in the movement to be executed.

Motor Facilitation

Observing a congruent movement did not facilitate movement execution (Kilner et al., 2003; Bouquet et al., 2007; Stanley et al., 2007). This is probably due to the fact that observation of a congruent grasping action during execution of a similar action facilitates precision of the grasp component only if the two events are *highly synchronized* (Ménoret et al., 2013). Here, participants were asked to perform their movement at their own pace and no instruction was given concerning the speed of the movement. Moreover, the movement observed in the non-interactive condition had an additional level of complexity due

to pouring the sugar in a cup as compared to the instructed movement of the participant. This could have also played a role in activating a partial competition between different representations.

The Social Associative Memory Hypothesis

Overall, this study provides evidence that online interference occurs when an observed movement requires an incongruent grasping with respect to the prehension simultaneously observed and executed. This result, together with recent TMS studies on cortico-spinal excitability (Sartori et al., 2011b, 2012, 2013b,c) and previous kinematic data (Sartori et al., 2009), suggests that observing an interactive gesture automatically generates an internal *representation* of the required movement. Such an internal representation can cause interference in the execution of the grasping movement, when active at the same time.

An accumulating body of evidence seems to suggest the existence of a human *motor vocabulary* (Rizzolatti et al., 2001) in which congruent – and incongruent (Sartori et al., 2013a) – motor representations would be activated automatically during the observation of motor actions. According to Chinellato et al. (2013), a social associative memory would be in charge of matching certain actions to their natural social response, irrespective of who is actually performing the action. If action B (e.g., take) usually follows action A (e.g., give), the observation of a partner executing A elicits the pre-planning of B by the observer. On the other hand, if the subject executes A, she expects to see the partner performing B in response. The same concept has been put forward by Butterfill and Sinigaglia (2014): “Two outcomes, A and B, match in a particular context just if, in that context, either the occurrence of A would normally constitute or cause, at least partially, the occurrence of B or vice versa” (see also Catmur et al., 2009).

As in the case of previous literature on social Simon effect (Guagnano et al., 2010; Humphreys and Bedford, 2011; Dittrich et al., 2013), it remains to be clarified whether the social context is a necessary prerequisite or not for this interference effect. In this respect, a previous experiment with similar stimuli and an arrow indicating the target object instead of the social gesture suggested that the mere presence of an arrow pointing toward the object had the ability to determine MEP activation. However, such activity was reduced with respect to when the context was characterized by a request gesture toward the object (Sartori et al., 2011b). Those findings corroborate the idea that it is the social nature of the observed gesture, along with the presence of the object, to determine the observed effect.

Acknowledgments

This work was supported by EU-FP7 grants 600623 (STRANDS project) to EC, and by a grant N. 287713 of FP7: REWIRE project and Progetto Strategico, Università di Padova (N. 2010XPMFW4) to UC.

References

- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat. Rev. Neurosci.* 4, 165–178. doi: 10.1038/nrn1056
- Baess, P., and Prinz, W. (2015). My partner is also on my mind: social context modulates the N1 response. *Exp. Brain Res.* 233, 105–113. doi: 10.1007/s00221-014-4092-9
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008a). Both your intention and mine are reflected in the kinematics of my reach to grasp movement. *Cognition* 106, 894–912. doi: 10.1016/j.cognition.2007.05.004
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008b). The case of Dr. Jekyll and Mr. Hyde: a kinematic study on social intention. *Conscious. Cogn.* 17, 557–564. doi: 10.1016/j.concog.2007.03.003
- Bertenthal, B. I., Longo, M. R., and Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 210–225. doi: 10.1037/0096-1523.32.2.210
- Blakemore, S. J., and Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia* 43, 260–267. doi: 10.1016/j.neuropsychologia.2004.11.012
- Bouquet, C. A., Gaurier, V., Shipley, T., Toussaint, L., and Blandin, Y. (2007). Influence of the perception of biological or non-biological motion on movement execution. *J. Sports Sci.* 25, 519–530. doi: 10.1080/02640410600946803
- Brass, M., Bekkering, H., and Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychol.* 106, 3–22. doi: 10.1016/S0001-6918(00)00024-X
- Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn.* 44, 124–143. doi: 10.1006/brcg.2000.1225
- Briggs, G. G., and Nebes, R. D. (1975). Patterns of hand preference in a student population. *Cortex* 11, 230–238. doi: 10.1016/S0010-9452(75)80005-0
- Butterfill, S. A., and Sinigaglia, C. (2014). Intention and motor representation in purposive action. *Philos. Phenom. Res.* 88, 119–145. doi: 10.1111/j.1933-1592.2012.00604.x
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007
- Castiello, U., Lusher, D., Mari, M., Edwards, M. G., and Humphreys, G. W. (2002). “Observing a human or a robotic hand grasping an object: differential motor priming effects,” in *Attention and Performance*, Vol. 19, eds W. Prinz and B. Hommel (Cambridge: MIT Press).
- Catmur, C., Walsh, V., and Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2369–2380. doi: 10.1098/rstb.2009.0048
- Chinellato, E., Ognibene, D., Sartori, L., and Demiris, Y. (2013). “Time to change: deciding when to switch action plans during a social interaction,” in *Biomimetic and Biohybrid Systems*, eds N. F. Lepora, A. Mura, H. G. Krapp, P. F. M. J. Verschure, and T. J. Prescott (Berlin: Springer), 47–58.
- Christensen, A., Ilg, W., and Giese, M. A. (2011). Spatiotemporal tuning of the facilitation of biological motion perception by concurrent motor execution. *J. Neurosci.* 31, 3493–3499. doi: 10.1523/JNEUROSCI.4277-10.2011
- Craighero, L., Bello, A., Fadiga, L., and Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia* 40, 492–502. doi: 10.1016/S0028-3932(01)00134-8
- Cross, E. S., Hamilton, A. F. C., and Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage* 31, 1257–1267. doi: 10.1016/j.neuroimage.2006.01.033
- D’Amico, M., and Ferrigno, G. (1990). Technique for the evaluation of derivatives from noisy biomechanical displacement data using a model based bandwidth-selection procedure. *Med. Biol. Eng. Comput.* 28, 407–415. doi: 10.1007/BF02441963
- D’Amico, M., and Ferrigno, G. (1992). Comparison between the more recent techniques for smoothing and derivative assessment in biomechanics. *Med. Biol. Eng. Comput.* 30, 193–204. doi: 10.1007/BF02446130
- Dijkerman, H. C., and Smit, M. C. (2007). Interference of grasping observation during prehension, a behavioural study. *Exp. Brain Res.* 176, 387–396. doi: 10.1007/s00221-006-0627-z
- Dittrich, K., Dolk, T., Rothe-Wulf, A., Klauer, K. C., and Prinz, W. (2013). Keys and seats: spatial response coding underlying the joint spatial compatibility effect. *Atten. Percept. Psychophys.* 75, 1725–1736. doi: 10.3758/s13414-013-0524-z
- Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., and Liepelt, R. (2014). The joint Simon effect: a review and theoretical integration. *Front. Psychol.* 5:974. doi: 10.3389/fpsyg.2014.00974
- Duque, J., Labruna, L., Verset, S., Olivier, E., and Ivry, R. B. (2012). Dissociating the role of prefrontal and premotor cortices in controlling inhibitory mechanisms during motor preparation. *J. Neurosci.* 32, 806–816. doi: 10.1523/JNEUROSCI.4299-12.2012
- Duque, J., Lew, D., Mazzocchio, R., Olivier, E., and Ivry, R. B. (2010). Evidence for two concurrent inhibitory mechanisms during response preparation. *J. Neurosci.* 30, 3793–3802. doi: 10.1523/JNEUROSCI.5722-09.2010
- Edwards, M. G., Humphreys, G. W., and Castiello, U. (2003). Motor facilitation following action observation: a behavioural study in prehensile action. *Brain Cogn.* 53, 495–502. doi: 10.1016/S0278-2626(03)00210-0
- Fadiga, L., Craighero, L., and Olivier, E. (2005). Human motor cortex excitability during the perception of others’ action. *Curr. Opin. Neurobiol.* 15, 213–218. doi: 10.1016/j.conb.2005.03.013
- Fitts, P. M., and Seeger, C. M. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *J. Exp. Psychol.* 46, 199–210. doi: 10.1037/h0062827
- Flach, R., Press, C., Badet, A., and Heyes, C. M. (2010). Shaking hands: priming by social effects. *Br. J. Psychol.* 101, 739–749. doi: 10.1348/000712609X484595
- Grosjean, M., Zwickel, J., and Prinz, W. (2009). Acting while perceiving: assimilation precedes contrast. *Psychol. Res.* 73, 3–13. doi: 10.1007/s00426-008-0146-6
- Guagnano, D., Rusconi, E., and Umiltà, C. A. (2010). Sharing a task or sharing space? On the effect of the confederate in action coding in a detection task. *Cognition* 114, 348–355. doi: 10.1016/j.cognition.2009.10.008
- Hamilton, A., Wolpert, D., and Frith, U. (2004). Your own action influences how you perceive another person’s action. *Curr. Biol.* 14, 493–498. doi: 10.1016/j.cub.2004.03.007
- Hamilton, A. F. (2013). The mirror neuron system contributes to social responding. *Cortex* 49, 2957–2959. doi: 10.1016/j.cortex.2013.08.012
- Hardwick, R. M., and Edwards, M. G. (2012). Motor interference and facilitation arising from observed movement kinematics. *Q. J. Exp. Psychol.* 65, 840–847. doi: 10.1080/17470218.2012.672995
- Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychol. Res.* 73, 512–526. doi: 10.1007/s00426-009-0234-2
- Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychol.* 136, 189–202. doi: 10.1016/j.actpsy.2010.04.011
- Humphreys, G. W., and Bedford, J. (2011). The relations between joint action and theory of mind: a neuropsychological analysis. *Exp. Brain Res.* 211, 357–369. doi: 10.1007/s00221-011-2643-x
- Innocenti, A., De Stefani, E., Bernardi, N. F., Campione, G. C., and Gentilucci, M. (2012). Gaze direction and request gesture in social interactions. *PLoS ONE* 7:e36390. doi: 10.1371/journal.pone.0036390
- Jeannerod, M., Arbib, M. A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320. doi: 10.1016/0166-2236(95)93921-J
- Kilner, J. M., Paulignan, Y., and Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Cur. Biol.* 13, 522–525. doi: 10.1016/S0960-9822(03)00165-9
- Klein, P.-A., Olivier, E., and Duque, J. (2012). Influence of reward on corticospinal excitability during movement preparation. *J. Neurosci.* 32, 18124–18136. doi: 10.1523/JNEUROSCI.1701-12.2012
- Knoblich, G., Butterfill, S., and Sebanz, N. (2011). “Psychological research on joint action: theory and data,” in *The Psychology of Learning and Motivation*, Vol. 54, ed. B. H. Ross (San Diego, CA: Academic Press), 59–101. doi: 10.1016/b978-0-12-385527-5.00003-6
- Labruna, L., Lebon, F., Duque, J., Klein, P.-A., Cazaes, C., and Ivry, R. B. (2014). Generic inhibition of the selected movement and constrained inhibition of

- nonselected movements during response preparation. *J. Cogn. Neurosci.* 26, 269–278. doi: 10.1162/jocn_a.00492
- Liepelt, R., Prinz, W., and Brass, M. (2010). When do we simulate non-human agents? Dissociating communicative and non-communicative actions. *Cognition* 115, 426–434. doi: 10.1016/j.cognition.2010.03.003
- Ménoret, M., Curie, A., Des Portes, V., Nazir, T. A., and Paulignan, Y. (2013). Simultaneous action execution and observation optimise grasping actions. *Exp. Brain Res.* 227, 407–419. doi: 10.1007/s00221-013-3523-3
- Müsseler, J., Wühr, P., and Prinz, W. (2000). Varying the response code in the blindness to response-compatible stimuli. *Vis. Cogn.* 7, 743–767. doi: 10.1080/13506280050144416
- Newman-Norlund, R. D., Bosga, J., Meulenbroek, R. G., and Bekkering, H. (2008). Anatomical substrates of cooperative joint-action in a continuous motor task: virtual lifting and balancing. *Neuroimage* 41, 169–177. doi: 10.1016/j.neuroimage.2008.02.026
- Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. G., and Bekkering, H. (2007a). Exploring the brain basis of joint action: co-ordination of actions, goals and intentions. *Soc. Neurosci.* 2, 48–65. doi: 10.1080/17470910701224623
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., and Bekkering, H. (2007b). The mirror neuron system is more activated during complementary compared with imitative action. *Nat. Neurosci.* 10, 817–818. doi: 10.1038/nn1911
- Paulignan, Y., Jeannerod, M., MacKenzie, C., and Marteniuk, R. (1991a). Selective perturbation of visual input during prehension movements. 2. The effects of changing object size. *Exp. Brain Res.* 87, 407–420. doi: 10.1007/BF00231858
- Paulignan, Y., MacKenzie, C., Marteniuk, R., and Jeannerod, M. (1991b). Selective perturbation of visual input during prehension movements. 1. The effects of changing object position. *Exp. Brain Res.* 83, 502–512. doi: 10.1007/BF00229827
- Pfister, R., Dolk, T., Prinz, W., and Kunde, W. (2014). Joint response–effect compatibility. *Psychon. Bull. Rev.* 21, 817–822. doi: 10.3758/s13423-013-0528-7
- Prinz, W., and Hommel, B. (eds). (2002). “Common mechanisms in perception and action,” in *Attention and Performance*, Vol. 19 (New York, NY: Oxford University Press).
- Quesque, F., and Coello, Y. (2014). For your eyes only: effect of confederate's eye level on reach-to-grasp action. *Front. Psychol.* 5:1407. doi: 10.3389/fpsyg.2014.01407
- Quesque, F., Lewkowicz, D., Delevoye-Turrell, Y., and Coello, Y. (2013). Effects of social intention on movement kinematics in cooperative actions. *Front. Neurobot.* 7:14. doi: 10.3389/fnbot.2013.00014
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670. doi: 10.1038/35090060
- Sartori, L., Becchio, C., Bulgheoni, M., and Castiello, U. (2009). Modulation of the action control system by social intention: unexpected social requests override preplanned action. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1490–1500. doi: 10.1037/a0015777
- Sartori, L., Becchio, C., and Castiello, U. (2011a). Cues to intention: the role of movement information. *Cognition* 119, 242–252. doi: 10.1016/j.cognition.2011.01.014
- Sartori, L., Cavallo, A., Buccioni, G., and Castiello, U. (2011b). Corticospinal excitability is specifically modulated by the social dimension of observed actions. *Exp. Brain Res.* 211, 557–568. doi: 10.1007/s00221-011-2650-y
- Sartori, L., Betti, S., and Castiello, U. (2013a). When mirroring is not enough: that is, when only a complementary action will do (the trick). *Neuroreport* 24, 601–604. doi: 10.1097/WNR.0b013e3283630a66
- Sartori, L., Betti, S., and Castiello, U. (2013b). Corticospinal excitability modulation during action observation. *J. Vis. Exp.* 82:e51001. doi: 10.3791/51001
- Sartori, L., Buccioni, G., and Castiello, U. (2013c). When emulation becomes reciprocity. *Soc. Cogn. Affect. Neurosci.* 8, 662–669. doi: 10.1093/scan/nss044
- Sartori, L., Cavallo, A., Buccioni, G., and Castiello, U. (2012). From simulation to reciprocity: the case of complementary actions. *Soc. Neurosci.* 7, 146–158. doi: 10.1080/17470919.2011.586579
- Schubö, A., Aschersleben, G., and Prinz, W. (2001). Interactions between perception and action in a reaction task with overlapping S–R assignments. *Psychol. Res.* 65, 145–157. doi: 10.1007/s004260100061
- Schubö, A., Prinz, W., and Aschersleben, G. (2004). Perceiving while acting: action affects perception. *Psychol. Res.* 68, 208–215. doi: 10.1007/s00426-003-0133-x
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76. doi: 10.1016/j.tics.2005.12.009
- Sebanz, N., Knoblich, G., and Prinz, W. (2005). How to share a task: co-representing Stimulus-Response Mapping. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 1234–1246. doi: 10.1037/0096-1523.31.6.1234
- Stanley, J., Gowen, E., and Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 915–926. doi: 10.1037/0096-1523.33.4.915
- Stürmer, B., Aschersleben, G., and Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1746–1759. doi: 10.1037/0096-1523.26.6.1746
- Stürmer, B., and Leuthold, H. (2003). Control over response priming in visuomotor processing: a lateralized event-related potential study. *Exp. Brain Res.* 153, 35–44. doi: 10.1007/s00221-003-1579-1
- Swinnen, S. P. (2002). Intermanual coordination: from behavioural principles to neural-network interactions. *Nat. Rev. Neurosci.* 3, 348–359. doi: 10.1038/nrn807
- Tipper, S. P., Howard, L. A., and Jackson, S. R. (1997). Selective reaching to grasp: evidence for distractor interference effects. *Vis. Cogn.* 4, 1–38. doi: 10.1080/713756749
- van den Berg, F. E., Swinnen, S. P., and Wenderoth, N. (2011). Excitability of the motor cortex ipsilateral to the moving body side depends on spatio-temporal task complexity and hemispheric specialization. *PLoS ONE* 6:e17742. doi: 10.1371/journal.pone.0017742
- van Schie, H. T., van Waterschoot, B. M., and Bekkering, H. (2008). Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1493–1500. doi: 10.1037/a0011750
- Verstynen, T., and Ivry, R. B. (2011). Network dynamics mediating ipsilateral motor cortex activity during unimanual actions. *J. Cogn. Neurosci.* 23, 2468–2480. doi: 10.1162/jocn.2011.21612
- Wilson, M., and Knoblich, G. (2005). The case of motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460
- Zwicker, J., Grosjean, M., and Prinz, W. (2010). On interference effects in concurrent perception and action. *Psychol. Res.* 74, 152–171. doi: 10.1007/s00426-009-0226-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Chinellato, Castiello and Sartori. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Different but complementary roles of action and gaze in action observation priming: Insights from eye- and motion-tracking measures

Clément Letesson^{1,2}, Stéphane Grade^{1,2} and Martin G. Edwards^{1,2*}

¹ Psy-NAPS Group, Institut de Recherches en Sciences Psychologiques, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, ² Institute of Neuroscience, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Luisa Sartori,
University of Padova, Italy
Matthias Hartmann,
University of Potsdam, Germany

*Correspondence:

Martin G. Edwards,
Psy-NAPS Group, Institut
de Recherches en Sciences
Psychologiques, Université
Catholique de Louvain, Place Cardinal
Mercier, 10 B-1348
Louvain-la-Neuve, Belgium
martin.edwards@uclouvain.be

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 05 February 2015

Accepted: 20 April 2015

Published: 05 May 2015

Citation:

Letesson C, Grade S
and Edwards MG (2015) Different but
complementary roles of action
and gaze in action observation
priming: Insights from eye-
and motion-tracking measures.
Front. Psychol. 6:569.
doi: 10.3389/fpsyg.2015.00569

Action priming following action observation is thought to be caused by the observed action kinematics being represented in the same brain areas as those used for action execution. But, action priming can also be explained by shared goal representations, with compatibility between observation of the agent's gaze and the intended action of the observer. To assess the contribution of action kinematics and eye-gaze cues in the prediction of an agent's action goal and action priming, participants observed actions where the availability of both cues was manipulated. Action observation was followed by action execution, and the congruency between the target of the agent's and observer's actions, and the congruency between the observed and executed action spatial location were manipulated. Eye movements were recorded during the observation phase, and the action priming was assessed using motion analysis. The results showed that the observation of gaze information influenced the observer's prediction speed to attend to the target, and that observation of action kinematic information influenced the accuracy of these predictions. Motion analysis results showed that observed action cues alone primed both spatial incongruent and object congruent actions, consistent with the idea that the prime effect was driven by similarity between goals and kinematics. The observation of action and eye-gaze cues together induced a prime effect complementarily sensitive to object and spatial congruency. While observation of the agent's action kinematics triggered an object-centered and kinematic-centered action representation, independently, the complementary observation of eye-gaze triggered a more fine-grained representation illustrating a specification of action kinematics toward the selected goal. Even though both cues differentially contributed to action priming, their complementary integration led to a more refined pattern of action priming.

Keywords: mirror neurons, action observation, eye gaze, action priming, action prediction

Introduction

Making sense of the behaviors of others and predicting the likely outcome of their actions is an essential component of interactive behavior (Wilson and Knoblich, 2005), thought to rely on common action observation and execution neural processes (e.g., the mirror neuron system, MNS; Rizzolatti and Craighero, 2004). These overlapping processes between observation and

execution allow for the observation of a specific action to activate the observer's motor system. One consequence of these shared processes is that the observation of action can moderate action execution for different action components, including action speed or timing (Edwards et al., 2003), action force (Salama et al., 2011) and action spatial trajectory (Hardwick and Edwards, 2011). This facilitation, known as the action priming effect, demonstrates that the shared neural processes between action observation and execution must encode precise information about the perceived action.

Recently, Wilson and Knoblich (2005) suggested that action observation neural processes incorporate predictive cognition. Observers use their own motor system to model (or represent) observed actions, allowing for the computation and prediction of an agent's behavior and unfolding action. In this sense, the neural processes are not simply activated in a bottom-up fashion by the mere observation of others' actions, but rather in anticipation to them. For example, Kilner et al. (2004) presented participants with video clips of either a stationary hand flanked by an object or a moving hand grasping an object. The color of the object indicated the type of video stimuli presented; either the hand would remain stationary, or whether it would move to grasp the object. The results showed that the object color associated with the moving grasp action condition caused predictive motor neural activity in anticipation of the initiation of the moving hand stimuli, whereas there was no such motor neural activity in anticipation of the static hand stimuli. A similar effect was reported by Umiltà et al. (2001). They reported results from a set of motor neurons that was active both for the observation of a fully visible reach and grasp action to an object, and also for the observation of a similar action where part of the reach and the grasp was occluded from vision by a screen (though the observer knew that there was a target object behind the screen). In this case, an early anticipation of the action goal must have been computed. The observer must have somehow relied on other visual information to extract relevant cues regarding the presence of the occluded action and target, and this caused predictive motor neural activation.

One might assume that the observation of the unfolding action could be sufficient for the observer to anticipate the action goal. However, Flanagan and Johansson (2003) showed that participants paid very little attention to the unfolding action, but instead, they implemented proactive eye movements similar to those used during actual action execution, where eye-gaze was anticipatively directed to the end-point or the goal of the action (see Rotman et al., 2006 for similar findings). This attentive pattern was explained as a procedure to provide visual feedback about the ongoing action execution relative to the target object or final action goal, and that any errors in the action trajectory could be anticipated and perceived to provide information for correction (Land and Furneaux, 1997; Gesierich et al., 2008). Indeed, during visually guided actions, there is little doubt that proactive gaze behaviors are essential for correct planning and coherent control of the executed motor program (Johansson et al., 2001).

The finding that observers use predictive eye-gaze patterns during action observation suggests that information from different visual sources must be obtained in order to infer the intended

action goal. These visual cues could emerge from an early analysis of the agent's behaviors before the observed action is fully executed. In this sense, Ambrosini et al. (2011) measured how fast and how accurately participants were able to anticipatively gaze at an agent's intended action target. Participants were asked to observe several types of reach-to-grasp actions while their eye movements were recorded. The observed actions could be directed to one of two different sized objects (small versus large), and the agent could either correctly pre-shape their hand to the target object (e.g., precision grip versus whole hand grip to the small and large sized objects respectively) or the agent showed no pre-shaping of their hand when acting to the objects (e.g., closed fist). The results showed that the hand pre-shaping condition caused participants to gaze at the correct target object quicker and more accurately than for the no hand pre-shaping condition. This suggests that hand information during the action observation provided a reliable cue to allow an early prediction of the intended target or action goal.

Hand-shape motoric cues are not the only source of information allowing for the prediction of the agent's action goal. In action execution, we normally first gaze toward an object that we intend to interact with, before actually acting upon the object (Johansson et al., 2001; Land and Hayhoe, 2001). The information that the observer could glean from the agent's gaze would constitute a reliable cue to predict object interaction intention (Becchio et al., 2008). Indeed, many studies have already examined the automatic tendency of the observer to orient their gaze to the same location as an agent's perceived gaze (Driver et al., 1999; Langton et al., 2000). Similarly, Castiello (2003) showed that the observation of another person's gaze toward a target object, as well as an actual action, reliably primed action execution. Further, Pierro et al. (2006) used functional magnetic resonance imaging (fMRI) to measure brain activity when participants observed video clips of a human model either reaching and grasping a target object or gazing at an object. The contrast between these two conditions and a control condition (in which the agent stood behind the object and performed no action or gaze) revealed similar profiles of brain activity. This suggests that the two types of information might be represented in a common motor code, and that either information could be sufficient to prime action execution. However, currently it remains unclear how action and gaze information interact during action observation, and whether the different types of information moderate the observer's executed gaze patterns and subsequent action responses.

The aim of the present study was twofold. First, we aimed to provide evidence for attentive and predictive eye-gaze behavior during action observation by measuring the observer's eye movements to different specified regions of interest (ROI), and by measuring the speed and accuracy of anticipatory gaze relative to manipulations of the agent's gaze and action kinematics to a target object. The interest of the latter analysis was to determine which of the gaze or action visual cue information would be selected when the two types of information were manipulated in the visual scene. We aimed to replicate previous studies that have investigated the predictive functioning of action observation processes (Rotman et al., 2006; Webb et al., 2010; Ambrosini

et al., 2011, 2012) and in addition, demonstrate that the agent's gaze provides early cues that indicate an intention to grasp a particular target object. Additionally, as observers have been shown to be efficient at extracting action intention information from both gaze cuing and observed actions (Sartori et al., 2011), we expected that, in the absence of the agent's gaze, the participants would orient their attention to the ongoing action as a secondary source of information. The second aim of the study was to better understand the different and complementary effects that gaze and action cues could have on the action priming effect. Recent studies showed that observed actions can be encoded in terms of their goal (Rizzolatti and Craighero, 2004), but also that goal representations of observed actions can be accompanied with more specific information regarding action kinematics, such as action trajectories (Griffiths and Tipper, 2009; Hardwick and Edwards, 2011). We therefore expected that the observation of both gaze-object and hand-object interactions would moderate subsequent action execution kinematics. However, in the current scientific literature, it remains unclear if these cue-induced priming effects are driven by a similarity of goals and/or trajectories between the observed and the executed actions. Therefore, we assessed the contribution of goal information by manipulating the congruency of the target objects during action observation and action execution, and further, we investigated the contribution of kinematics information through the manipulation of spatial congruency between the observed and executed actions.

Materials and Methods

Participants

We tested a total of 22 persons, though three participants were excluded because of corruptions in their data recording (i.e., recording failures causing unusable data) and were not analyzed any further. The mean age of the remaining 19 participants was 22.1 years (range: 2.3 years), all were right-handed (self-reported) and had normal or corrected-to-normal visual acuity. All participants gave their informed consent to take part in the study and they were remunerated for their participation. The Université Catholique de Louvain, Faculty of Psychology Ethics Commission approved the experiment.

Apparatus and Stimuli

To record participants' eye movements, we used the Eyelink 1000 desktop mounted eye tracker (SR Research, Canada; sampling rate of 1000 Hz; average accuracy range 0.25–0.5°, gaze tracking range of 32° horizontally and 25° vertically). Participants sat at a distance of 60 cm from the eye tracker camera and head movements were prevented by using a chin and forehead stabilizer. At the beginning of each trial block, a standard 9-point protocol was used to calibrate the participant's eye-gaze position to a display screen using standard Eye-Link software. This allowed the computation of the actual gaze position on the screen. To record participants' hand actions, we used the Polhemus Liberty electromagnetic 3D motion tracker (Polhemus Incorporated, Colchester, Vermont; sampling rate of 240 Hz, accuracy 0.076 cm

for position and 0.15° for orientation). Sensors were attached to two target objects, and the participant's wrist, thumb, and index finger using adhesive tape and a flexible wrist splint. The kinematic data were analyzed offline and the dependent measurements were extracted from the 3D XYZ coordinates.

The laboratory arrangement consisted of three wooden tables (120 × 80 cm) creating an L-shaped workspace (see **Figure 1**). Both the participant and the experimenter faced the same direction, with the participant to the left of the experimenter. The experimenter was positioned (offset) behind the participant allowing a view of the participants' workspace/screen without distracting them. We further shielded distractions by placing a wooden panel between the tables (occluding all of the computer equipment that we used for the eye-tracker and motion tracker recordings). On the participant's table, the chin and forehead stabilizer was placed centrally, 5 cm from the table edge. A computer screen (LCD; resolution 1080 × 1920; refresh rate 60 Hz) was placed 70 cm from the chin and forehead stabilizer, and was used to display visual stimuli for the experiment. The visual stimuli were presented using E-Prime (v2.0.8.90 PRO; Schneider et al., 2002).

The experiment involved the use of two types of stimuli (visual and physical). The visual stimuli consisted of video clips presented on the computer screen depicting reach-to-grasp actions (AVI format, 25 fps, 1920 × 1080 pixels). The video clips consisted of a male agent sitting at a wooden table and looking into the camera, with his right hand holding a reference object (Ø: 2 cm) in front of him. A small and a large object were also presented (Ø: 4 and 7 cm), 25 cm in front of the agent, one on the left and the other on the right of the agent's sagittal midline (25 cm apart, and their position counterbalanced). In the videos, we manipulated the availability of the agent's gaze and action (see **Figure 2**). In the gaze and action (FULL) condition, the video started by showing the agent positioned facing

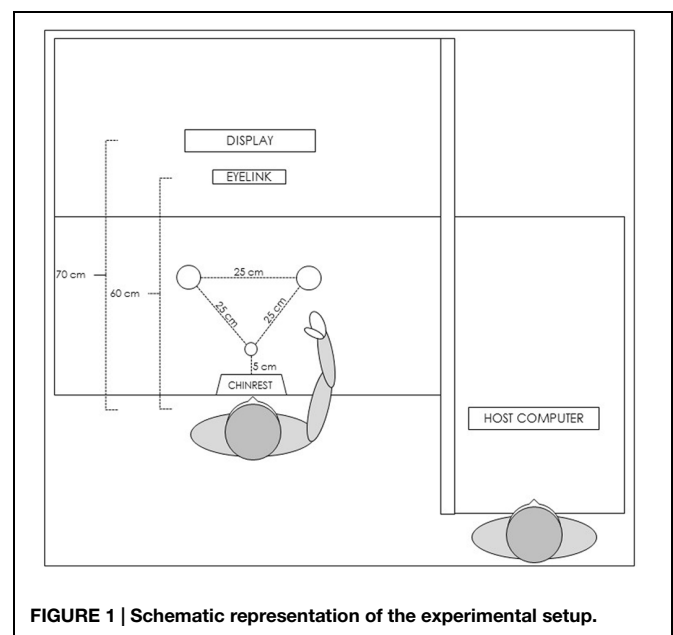


FIGURE 1 | Schematic representation of the experimental setup.



FIGURE 2 | Examples of observation video clips from each experimental condition. From top to bottom: conditions (A) FULL, (B) ACTION, (C) GAZE and (D) CONTROL. The videos were presented on a HD screen (1920*1080 pixel). The actor's eyes area

occupied a 3 cm screen area and participants viewed the screen at a distance of 70 cm (equivalent to $\sim 2.45^\circ$ of visual angle and equivalent to observing a real-sized human at a distance of ~ 220 cm).

the participant, looking into the camera lens, and a small and a large object presented on the table, symmetrical and of equal distance to the sagittal axis. Next, the agent directed his gaze toward either the small or the large object, and then he executed a reach-to-grasp action to the object that he gazed toward. In the gaze only condition (GAZE), the video again started by showing the agent looking into the camera lens (etc.). Next, he directed his gaze toward one of the two objects (as in the previous condition), but this time, no action was executed. Therefore, the agent's gaze direction was the only available cue indicating the target object. In the action only condition (ACTION), we placed a mask on the eyes area of the agent. The video started by showing the agent with the eye mask (preventing gaze cues; but with all other aspects of the video matched to the FULL condition). The only cue was of the agent executing a reach-to-grasp action to the object. The final condition was the no gaze and no action condition (CONTROL). In this condition, neither eye-gaze information nor action information was presented. The agent remained still throughout the video. For each of the experimental conditions (FULL, GAZE, ACTION), there were four different types of videos that balanced the size and position of the target objects (small left, small right, large left, large right). In the control condition (CONTROL), there were only two different videos (small object left and large object right versus large object left and small object right). The videos were matched in length (4500 ms) and each video was presented eight

times across two blocks of trials (with a total of 112 trials per participant).

The physical stimuli were presented in the participants' physical workspace. On the participant's sagittal axis, we placed a starting action reference object (\emptyset : 2 cm) positioned 5 cm further from the chin and forehead stabilizer. We also placed a small and a large round object (\emptyset : 4 and 7 cm; the same objects as those presented in the visual stimuli conditions), 25 cm from the reference object, and symmetrical to the participant's sagittal axis (with the edge of the objects 12.5 cm from the sagittal midline). The object positions were counterbalanced. At the beginning of each trial, the participant was asked to hold the starting action reference object with a right hand light grip (providing a common action origin point for the comparison between responses).

Design and Procedure

Each trial started with the presentation of a fixation cross in the center of the screen and the participant was instructed to look at the cross. This ensured that each participant would start observing the video sequences from the same origin point, allowing for a comparison between gaze paths. The fixation cross was also used as a drift check to verify and confirm the reliability of the eye-gaze calibration. The fixation cross was displayed until the experimenter manually confirmed that the participants' gaze was fixed to the cross position. As soon as participants' gaze

position was confirmed, one of the video clips was randomly presented. The participant was instructed to observe and attend to the video carefully, as they would be required to make an action in subsequent part of the trial. At the end of the video, a sound was presented that indicated the size of the object that the participant had to grasp during the final execution part of the trial. A low-tone indicated that the participant would have to grasp the large physical object and a high-tone indicated that the participant would have to grasp the small physical object. The participants were instructed that when they heard the sound, they had to reach, grasp and lift the target object in a natural manner (“as if you were reaching and grasping your cup of tea”), each time, initiating the action from the action reference object.

Data Analyses

The results were analyzed for both eye-tracking and motion-tracking measures. For the post-tests analyses, a Bonferroni correction was applied. For the data analyses, we separated the visual scene in the video clips into five ROI that were slightly larger than the part of interest in the visual scene (compensating for any variance in the eye-tracking data). The regions selected corresponded to the fixation cross, the agent’s head, the agent’s hand and the two target objects (the left object and the right object). Participant’s eye-gaze to each ROI was considered in the analyses and any eye-gaze outside of the ROI was not included in the data analyses. The grasped or gazed object was denominated as “TARGET” and the non-target object as “NON-TARGET” (irrespective of the object size or location).

We used three dependent variables to analyze the eye-tracking data. The first was the proportion of total fixation time spent in each ROI (with ROI added as an independent variable). The aim of this analysis was to investigate how manipulations to gaze and action cues moderated the participants’ attention to the manipulated bodily cues. Therefore, we analyzed whether the observations conditions induced different attentional profiles to specific ROIs. For this analysis, we only included trials where the participant started the trial by fixating the cross (98% valid trials). The second dependent variable was prediction speed; derived from the time elapsed between the participant’s first fixation to the target object ROI and either (i) the time when the agent’s gaze was directed to the target object (gaze-based index) or (ii) the time when the agent’s hand reached the target object (action-based index). The prediction speed indexes were specific to the conditions that manipulated these cues; with planned *t*-test contrasts for the gaze-based index comparing GAZE versus FULL conditions, and the action-based index comparing ACTION versus FULL conditions; no comparisons were possible between the CONTROL or between the GAZE and ACTION conditions. Therefore, the gaze-based index and the action-based index allowed us to measure respectively the contribution of action and gaze information to prediction speed. Here, only trials where participants correctly oriented their gaze toward the target of the agent’s attention or action were included. We excluded trials in which participants oriented their gaze toward a target before any gaze or action cues were presented in the video clip (84.5% valid trials). The third

dependent variable was prediction accuracy that was defined as the proportion of trials where the participant correctly oriented their eye (attention) toward the target of the agent’s attention or action. This variable measured the efficiency of our experimental manipulations in producing correct predictions. The same contrasts as those used for prediction speed were applied to this variable.

For the motion-tracking analyses, we determined two levels of congruency between the observation and the execution conditions: (i) ‘object congruency’ irrespective of spatial location (congruent: observation of action to the same sized object as that grasped in the execution condition; small–small or large–large versus incongruent: observation of action to a different sized object as that grasped in the execution condition; small–large or large–small); and (ii) ‘spatial congruency’ irrespective of object size (congruent: the same egocentric spatial location for the agent and observer in the observed and execution conditions; agent reaching to their left and participant reaching to their right or agent reaching to their right and participant reaching to their left versus incongruent: different egocentric spatial action location for the agent and observer in the observed and execution conditions; agent reaching to their left and participant reaching to their left or agent reaching to their right and participant reaching to their right). The prime effect was measured with three dependent variables: reaction time (ms; with action initiation being defined as the time when the hand velocity reached 50 mm/s for two successive frames), time to peak velocity (ms), and time to peak grasp aperture (ms). As the aim of the kinematics analysis was to understand how the cue-induced priming effects were differentially sensitive to object and spatial congruency, we defined a priori analyses to check how these two variables would interact throughout each video condition. Three-way interactions were decomposed by evaluating how object congruency and spatial congruency interacted in each video condition independently. If second order two-way interactions were significant, we performed multiple comparisons between object and spatial congruency. The rationale behind the present statistical approach is similar to that described by Howell and Lacroix (2012) for decomposing three-way interactions.

Results

Eye-Tracking Analyses

The repeated measures ANOVA for the proportion of total fixation time spent in the ROIs compared within-participant factors of video conditions (FULL, GAZE, ACTION, CONTROL) and ROIs (fixation, the agent’s head, the agent’s hand, the target and the non-target). We found significant main effects for the video conditions [$F(3,54) = 3.98, p < 0.05, \eta_p^2 = 0.18$] and ROIs [$F(4,72) = 282.18, p < 0.001, \eta_p^2 = 0.94$], and a significant interaction [$F(12,216) = 106.17, p < 0.001, \eta_p^2 = 0.85$]. The pairwise comparisons of the main effects showed that the proportion of total fixation time within the five ROIs was significantly lower for the ACTION than for the FULL and GAZE video conditions. Also, the proportion of total fixation time

was significantly different for each ROI, except for the target object and the fixation areas. The agent's head was fixated significantly longer than any other ROI, and the target object was fixated longer than the non-target object. We decomposed the interaction by evaluating each ROI separately. This showed significant observation condition effects for the ROIs of the agent's head [$F(3,54) = 90.13$, $p < 0.001$, $\eta_p^2 = 0.83$; with GAZE and CONTROL conditions eliciting longer fixation time compared to ACTION or FULL conditions], the agent's hand [$F(3,54) = 29.25$, $p < 0.001$, $\eta_p^2 = 0.62$; with ACTION significantly different from the other conditions], and the target [$F(3,54) = 167.32$, $p < 0.001$, $\eta_p^2 = 0.9$; with both CONTROL and GAZE being different from each other and the other two conditions]. There was a significant effect for the non-target ROI [$F(3,54) = 3.19$, $p < 0.05$, $\eta_p^2 = 0.15$; though the pairwise comparisons did not highlight any significant differences]. See **Figure 3A**.

The analysis of prediction speed and accuracy dependent variables compared the FULL condition with each of the manipulated GAZE or ACTION conditions using planned t -test and evaluated a gaze-based index and an action-based index (time between participant's target fixation and agent's eye-gaze or time between participant's target fixation and the agent's hand action to the target). Analysis of the gaze-based index showed no significant effect for prediction speed [$t(18) = 0.85$, $p = 0.4$, $d = 0.19$], but there was a significant effect of prediction accuracy [$t(18) = -4.5$, $p < 0.001$, $d = 1.08$]. The participant fixated to the correct target more frequently when both the agent's eye-gaze and action cue information was presented compared to the videos with the gaze cue alone, indicating that the presence of action information contributed to the correct orientation of participant's attention to the target object. Analysis of the action-based index showed a significant effect of prediction speed [$t(18) = -8.7$, $p < 0.001$, $d = 2$], but no effect of prediction accuracy [$t(18) = -1.16$, $p = 0.26$, $d = 0.26$]. The participant fixated to the target faster when both the agent's eye-gaze and action cue information was presented compared to the videos with the action cue alone, indicating that the processing of gaze information contributed to prediction speed (see **Figure 3**).

Motion-Tracking Analyses¹

To assess the action priming effect, we tested the independent variables of video conditions (FULL, GAZE, ACTION), object congruency, and spatial congruency using repeated measures ANOVAs on the three dependent variables of reaction time, time to peak velocity and time to peak grip aperture. CONTROL condition could not be included in the model as it did not vary for the spatial and object congruency independent variables. All of the results for this section are presented in **Figure 4**.

The analysis of the participants' reaction time showed a significant main effect of video condition [$F(2,36) = 47.37$, $p < 0.001$, $\eta_p^2 = 0.72$], with pairwise comparisons showing that the actions executed after the observation of the ACTION or FULL video conditions were initiated significantly faster than after the observation of the GAZE video condition. There was also a significant

main effect of spatial congruency [$F(1,18) = 5.87$, $p < 0.05$, $\eta_p^2 = 0.25$], with spatial incongruent actions being initiated faster than spatial congruent actions. Finally, the results showed a significant interaction between video condition and object congruency [$F(2,36) = 5.87$, $p < 0.01$, $\eta_p^2 = 0.25$], showing that object congruent trials were initiated faster than object incongruent trials in the ACTION condition only [$t(18) = 2.86$, $p < 0.05$, $d = 0.66$].

The analysis of time to peak velocity showed a significant main effect of spatial congruency [$F(1,18) = 5.16$, $p < 0.05$, $\eta_p^2 = 0.22$], showing that the observation of spatial incongruent action caused a faster time to peak velocity compared to spatial congruent actions. There was also a significant three-way interaction [$F(2,36) = 3.36$, $p < 0.05$, $\eta_p^2 = 0.16$] that was decomposed with an ANOVA for each video condition. These analyses only showed a significant spatial congruency effect for the ACTION video condition [$F(1,18) = 5.78$, $p < 0.05$, $\eta_p^2 = 0.24$], with spatial incongruent trials reaching their peak velocity quicker than the congruent ones. None of the other contrasts were significant once corrected.

There was a significant effect for time to peak grasp aperture for spatial congruency [$F(1,18) = 5.43$, $p < 0.05$, $\eta_p^2 = 0.23$], with the peak grip aperture being reached quicker for spatial incongruent than congruent actions. There was a significant interaction between video conditions and spatial congruency [$F(2,36) = 3.27$, $p < 0.05$, $\eta_p^2 = 0.15$], showing a faster time to peak grasp aperture in the spatial incongruent compared to congruent actions in the FULL video condition only [$t(18) = 2.95$, $p < 0.01$, $d = 0.67$]. There was also an interaction between video conditions and object congruency [$F(2,36) = 5.7$, $p < 0.01$, $\eta_p^2 = 0.24$] showing a significantly faster time to peak grasp aperture for the object congruent than incongruent trials in the ACTION condition [$t(18) = 3.87$, $p < 0.001$, $d = 0.88$]. Finally, there was a three-way interaction between all of the independent variables [$F(2,36) = 7.7$, $p < 0.01$, $\eta_p^2 = 0.30$] that was analyzed using ANOVAs for each video condition separately. Significant two-way interactions between spatial and object congruency were found for the FULL and GAZE video conditions [$F(1,18) = 5.27$, $p < 0.05$, $\eta_p^2 = 0.23$; $F(1,18) = 7.23$, $p < 0.05$, $\eta_p^2 = 0.29$; respectively]. Based on our hypothesis, paired-samples t -tests defined *a priori* contrasted spatial and object congruent versus incongruent conditions. For the FULL video condition, there was only a significant effect of spatial congruency when the objects were congruent [$t(18) = 3.28$, $p < 0.01$, $d = 0.75$], showing that spatial incongruent actions reached their peak grasp aperture faster than the spatial congruent actions. For the GAZE condition, none of the pairwise comparisons reached significance.

Discussion

The aim of the present study was firstly to evaluate foveal attention and predictive eye-gaze behaviors during action observation and, secondly, to better understand the complementary effects that eye-gaze and action cues have on the action observation priming. To evaluate natural foveal attention during the observation conditions, we determined how manipulations to

¹Summary tables of all of the results are given in the supplementary materials.

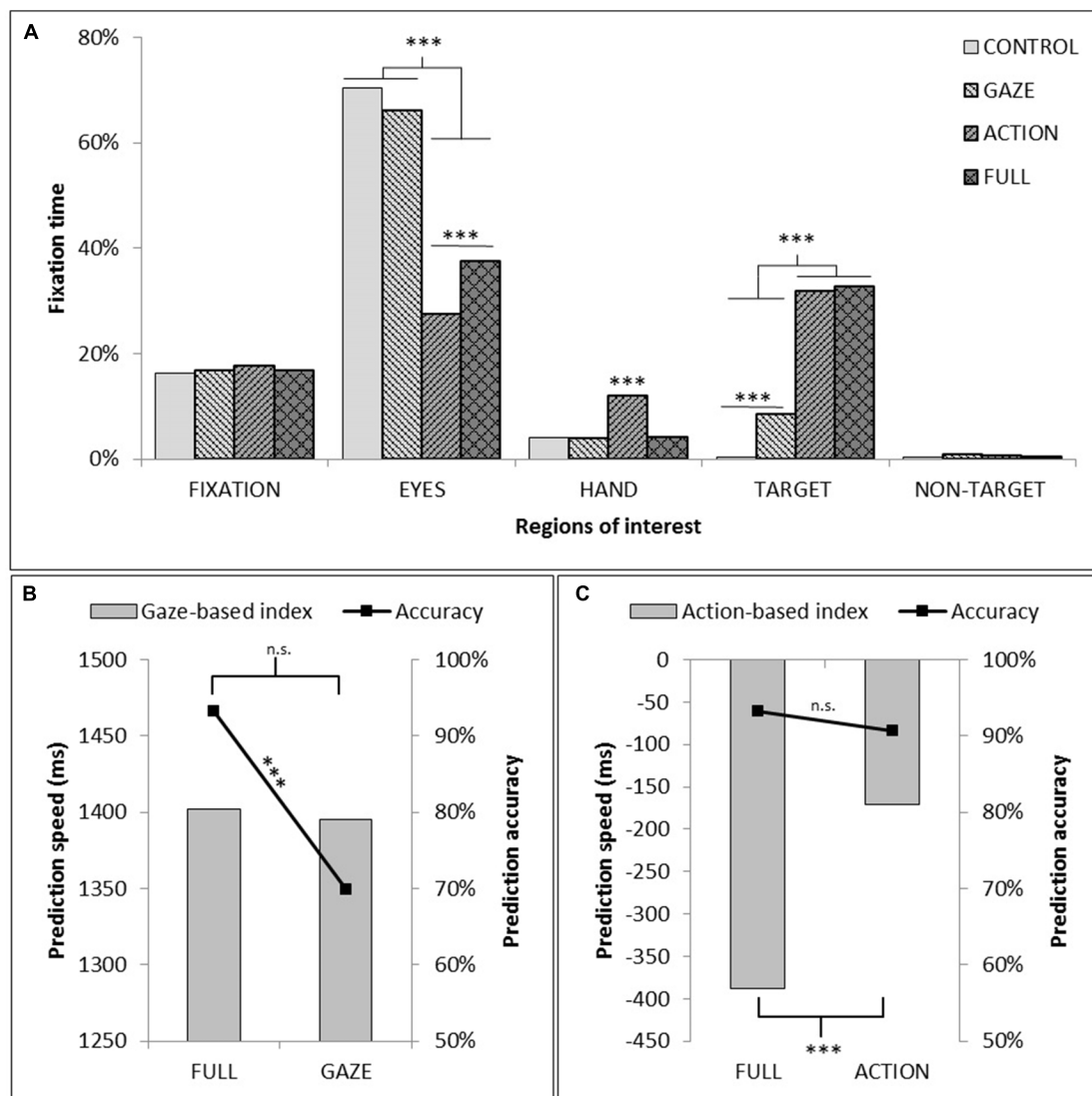


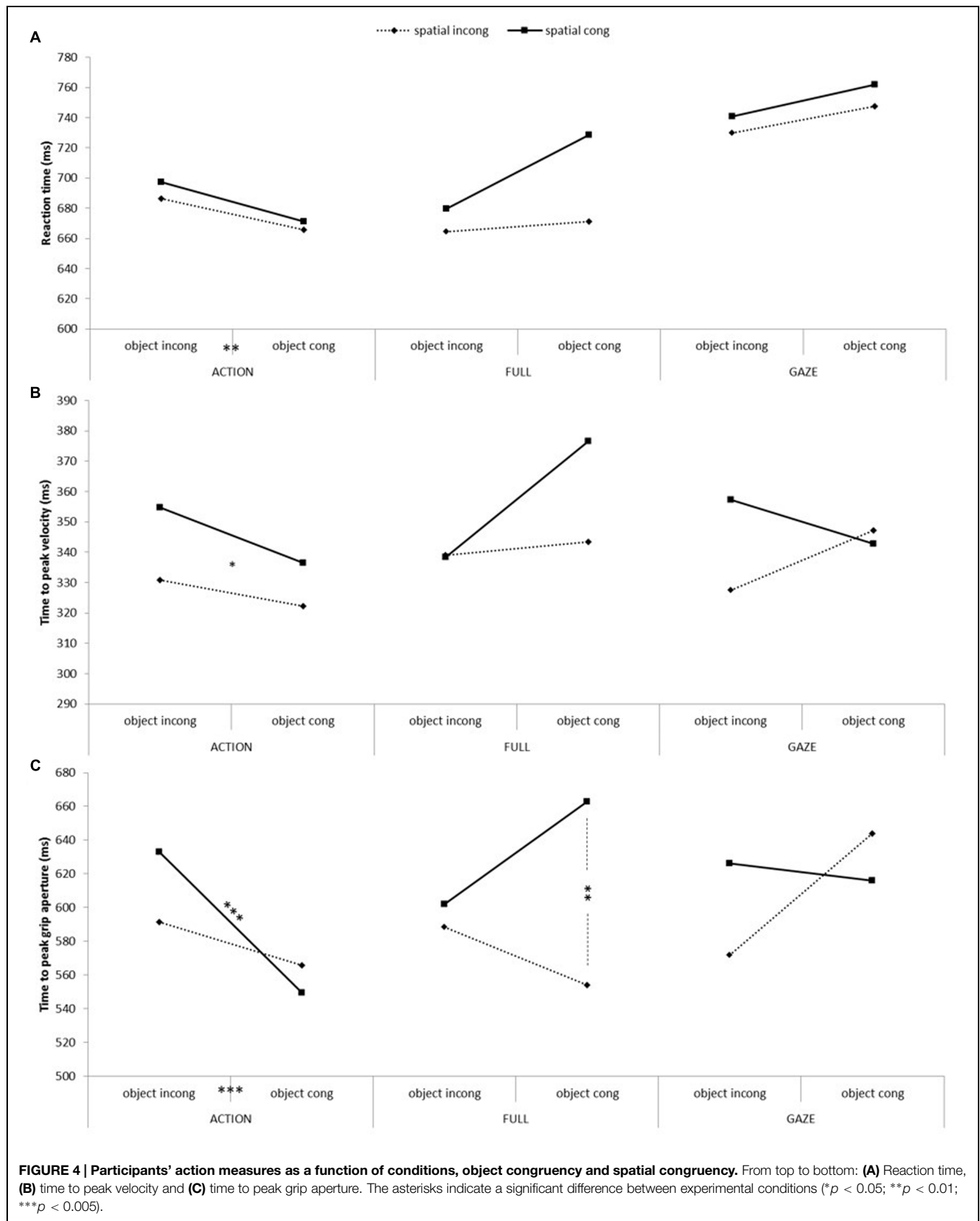
FIGURE 3 | (A) Proportion of total fixation time within each of the five ROIs as a function of observation conditions (In the CONTROL condition, as no specific object could be reported as 'TARGET' or 'NON-TARGET,' the average fixation time on both objects was plotted). **(B)** Measure of the contribution of action information to prediction speed and prediction accuracy through the

comparison of GAZE and FULL conditions. **(C)** Measure of the contribution of gaze information to prediction speed and prediction accuracy through the comparison of ACTION and FULL conditions. The asterisks indicate a significant difference between experimental conditions (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$).

gaze and action cues moderated the participants' attention to specific ROIs. Overall, this showed that participants spent more time fixating the head area of the agent in all of the observation conditions, although it was fixated more in the GAZE and CONTROL conditions compared to the ACTION and FULL observation conditions. In addition, the hand area of the agent was fixated more when only ACTION information was presented compared to the other observation conditions, and participants looked at the target for longer in the FULL and ACTION observation conditions relative to the GAZE and CONTROL conditions.

We also measured the differential contribution of the agent's gaze and action information cues on the participant's action prediction. We proposed that particular cues might facilitate the

participant's prediction to selectively attend to the correct target. The results showed that the speed at which the participants correctly oriented their attention to the target of an observed action was influenced by the availability of eye-gaze information, whereas manipulating action information influenced the accuracy of the predictions. Interestingly, the combined availability of action and gaze cues provided the most reliable prediction cues for both prediction speed and accuracy. These results show that when observing a human agent performing a goal-directed action, participants appeared to prioritize attention to the agent's eyes, and that information from both the agent's eye-gaze and action cues were important for predicting the target to which the agent intended to act toward (the goal of the observed action).



These findings are consistent with other studies showing a tendency for participants to attend to agent's eye-gaze and looking direction (Senju and Hasegawa, 2005; Conty et al., 2006), an effect that is perhaps not surprising given the use of eye contact to establish communicative links between individuals (Farroni et al., 2002). In the case of action observation, the observer could glean information about the agent's intentions through the establishment of joint attention (Driver et al., 1999). When performing actions, the agent usually attends to the object that they intend to act toward (Johansson et al., 2001; Land and Hayhoe, 2001) and therefore, this information appears to constitute a reliable source of predictive information.

Interestingly, there was proportionally very little attention allocated to the hand region during the observation of video clips during the FULL condition. However, observation of the agent's eye-gaze and action cues during the FULL condition resulted in greater prediction accuracy compared to observation of only the agent's eye-gaze cues during the GAZE condition. This suggests that the observation of hand trajectories must use peripheral vision, perhaps serving to reduce the ambiguity regarding the predicted action goal determined by the agent's eye-gaze cues, and contributing to increased prediction accuracy. These findings support and extend previous findings by Webb et al. (2010) who presented participants with video clips depicting one of two human agents performing reach-to-grasp actions to one of three different targets aligned horizontally. Participants were asked to observe the videos and determine the agent and the to-be-grasped object. Before each video, the identity of the agent or the target was unknown to the participant. The results showed that agent's eye-gaze direction and hand trajectory information were important in guiding the observer's gaze to the correct target. Our data add to these results by evaluating the relative contribution of both action and gaze cues in guiding the observer's gaze behaviors, and further, by showing that direct foveal attention to the observed hand trajectories was not necessary to cause the prediction effects, suggesting that action trajectories must have been attended to with peripheral vision.

The discrepancies between action and gaze information processing during action observation could somehow explain the trade-off between prediction speed and prediction accuracy. On the one hand, when no action information was available, the quick processing of the agent's gaze was made at the expense of accuracy, where perhaps an insufficient amount of information regarding the goal of the action had not yet been gathered. On the contrary, the absence of gaze information forced the observers to gather more information about the intended goal of the action from early motor information (hence the hand ROI was fixated for longer in the ACTION condition compared to other conditions). As it is less obvious for the observers to rule out the intended goal from the agent's action, they had to extrapolate the most likely target from the early trajectory of the agent's hand. Two possible targets were available in our design. Therefore, for the observers to make a correct prediction only based on the observation of the agent's hand trajectory, more time was needed to exclude the alternative object.

Flanagan and Johansson (2003) suggested that eye movements during action observation were proactive rather than reactive.

Our results provide further support for this claim, and give insights into how different cues are processed together in order to provide reliable predictive cues about ongoing actions. In everyday life situations, the targets of our actions are not fully predictable to the observer. Our actions can be oriented to a target presented with multiple other objects, differing in sizes, colors, or even locations. During visually guided action execution, the agent must extract the various features from the selected target object and position in order to grasp the object successfully. For example, the agent typically will pre-shape their hand to the size of the intended target object, with a slight over-grasp allowing for an efficient grip placement on the object (Jeannerod et al., 1995; Eastough and Edwards, 2007), and they will monitor the position of other objects in proximity to the target, making critical kinematic modifications to avoid the obstacle positions. This necessity of visual inputs for action control legitimates the anticipative nature of agent's eye movements during action execution (Johansson et al., 2001). Similar mechanisms were hypothesized during action observation conditions, as the human motor system is also involved in processes helping us to perceive the action of others (Wilson and Knoblich, 2005). Bach et al. (2011) provided evidence that the motor representations elicited during action observation were bi-determined. Not only did they match the observed actions, but they also reflected the proprieties of the goal objects that they were directed to. Accordingly, It has already been shown that observers are efficient in predicting future hand-object interactions by relying on hand pre-shaping cues (Ambrosini et al., 2011). Along the same lines, gaze cuing toward an object has been shown to provide a consistent indicator of a future interaction with that object, allowing the observers to predict the short term course of ongoing actions (Pierro et al., 2006). Our data take these findings one-step further by highlighting reliable and successful identification of the intended goal object from the processing of both the agent's action and gaze cues compared to each cue in isolation. Action and gaze cues provided different, but complementary advantages for action prediction, indicating the implementation of different observation strategies depending on the nature of the information available.

The second aim of this study was to better understand the complementary effects that gaze and action cues could have on subsequent action execution (i.e., the action priming effect). By modifying the cues during the observation conditions, we evaluated the effects of spatial and object congruency on subsequent action performance. We found that participants responded with a faster reaction time in actions executed after the observation of the ACTION or FULL video conditions, than after the observation of the GAZE video condition (without action information). This suggests that action information more than gaze information contributed to the action priming effect, suggesting that the action cue had an impact on motor planning processes. An alternative suggestion though could be that the slower reaction time to the gaze cue relative to the other cues might have been a consequence of the lower rate of accurate eye-gaze target prediction during observation. Reaction time was also moderated for spatial congruency, showing that spatial incongruent actions were initiated faster than spatial congruent actions. This effect suggests

that the observation of action was represented in a frame of reference centered on the observed agent; thus, when the participant observed a right hand action to a target object that was on the right of the agent (left of the participant), action was primed when the participant made a right hand action to a target on the right of the participant (the spatially incongruent target). This suggests priming between the observation and execution of action kinematics (the agent's right hand action primed the observer's right hand action; see Hardwick and Edwards, 2011 for similar priming of kinematic trajectories). This same effect was found for the dependent variables of time to peak velocity and time to peak grasp aperture. Further interaction analyses showed that the observation of the FULL condition (with gaze and action cues) compared to GAZE and ACTION conditions (with only one cue) caused participants to have a faster time to peak grasp aperture for the spatial incongruent compared to congruent actions, therefore replicating the main effect.

Counter evidence to the kinematic priming effects discussed above was shown for the interaction analyses of the ACTION condition (where no gaze information was presented). Reaction time and time to peak grasp aperture were earlier for congruent than incongruent target objects, supporting the idea of priming driven by common goals. The combined presentation of action and gaze cues (in the FULL condition) induced a more refined pattern of priming, sensitive to modulations of both object congruency and spatial congruency. The peak grip aperture was earlier for spatial incongruent actions (kinematic congruent) and later for spatially congruent actions (kinematic incongruent), only when the objects were congruent (similarity of goals). It is worth mentioning here that faster time to peak grip aperture is usually linked to a longer deceleration phase, allowing for a better control over the end-phase of the action and to adapt the hand to the state of the target (Jeannerod, 1994). This suggests that in the present findings, information regarding action goals and kinematic trajectories were important for the prime effect to appear and this probably improved grasp performance.

Supporting both the notion of a goal-driven priming and a kinematic priming, these data shed important light on the information extracted and represented during action observation. There is no doubt that the notion of goal is important in executed and observed actions, as shown by the goal-coding preferences of the motor system (Rizzolatti et al., 1996; Umiltà et al., 2001; Fogassi et al., 2005). The extraction of goal information from observed actions has been attributed to mirror neuron system function, where action goals are understood through the observed action resonating with the observer's own motor system. This mechanism has been proposed to allow for the prediction of the action goal based on simulation and perhaps prior experience of action execution (see Rizzolatti and Craighero, 2004; Rizzolatti and Sinigaglia, 2010 for reviews).

In the scientific literature to date, there has been little investigation to understand whether the presence of eye-gaze and action information during action observation differentially moderate the action priming effect. For example, in Edwards et al. (2003), both eye-gaze and action cue information were presented, and either information could have caused the reported action prime

effects. This point is important given the suggestion of Liepelt et al. (2008) that both action kinematics and the representation of goals could contribute to the action priming effects. The fact that the priming was specific to matched action kinematics and matched action goals independently in the ACTION condition here illustrates this bi-determination of motor representations. This might suggest that goal attribution and kinematic priming use independent cue information from the observed action, perhaps implying that they involve two independent cognitive processes that co-occur in parallel. This rationale is consistent with proposals suggesting that observed action representation do not solely rely on kinematic matching, but also require top down goal attribution (Jacob and Jeannerod, 2005; de Lange et al., 2008). In this sense, our results also suggest that the common language between perception and action could vary regarding a degree of abstraction, ranging from a very close representation of the action (kinematics-related) to a more global form of representation (goal-related).

However, as mentioned above, the combined presentation of action and gaze (in the FULL condition) elicited a more fine-grained profile of priming for the time to peak grip aperture. It seems that in this later kinematic component, goal-related priming and kinematic-related priming operated complementarily. In other words, similar goals led to either facilitated action execution if observers' hand trajectories matched that of the agent, or slowed action execution if there was a mismatch between observed kinematics and executed ones (see Ondobaka et al., 2012 for similar findings). These authors stated that when an agent's action intention is relevant for the observer's action execution, the kinematic-related priming is moderated by top-down goal ascription. We suggest that this is due to the perceived intentional value conveyed by an agent reaching for a target while his attention is directed toward the target of his reach. Jellema et al. (2000) described a population of cells in the anterior part of the superior temporal sulcus (aSTS) that responds preferentially to observed reaching action when the agent pays attention to the target of reach, compared to when attention is made elsewhere. According to the authors, eye-gaze in addition to an action would convey useful information to interpret the action as intentional. Under this interpretation, a correspondence between the agent's direction of attention and reaching action would refine the observers' motor representation to match the intention of the agent. Perceived eye-gaze direction could constitute a cue that allows the observers' motor system to distinguish different motor programs aiming for the same goal. This stronger visuo-motor congruency would explain how and why actions with similar goals and different kinematics produced competitive motor responses in the FULL condition.

Conclusion

In this study, we showed that agent's gaze and action differentially, but complementarily contributed to an early representation of the action goal. We suggest that once the goal

representation is understood by the observer's motor system, the diversity of the visual cues available influenced the level of abstraction of the motor representation elicited. We showed that action cues permitted goal-related priming and kinematic-related priming independently, whereas combined gaze and action information triggered a more refined representation illustrating a specific intended action kinematics toward the selected goal. In this case, observers appeared engaged in a communicative link with the agent, maybe through the establishment of joint attention. This permitted for the elicitation of richer motor representations, probably indicating the understanding of the observed motor intention.

References

- Ambrosini, E., Costantini, M., and Sinigaglia, C. (2011). Grasping with the eyes. *J. Neurophysiol.* 106, 1437–1442. doi: 10.1152/jn.00118.2011
- Ambrosini, E., Sinigaglia, C., and Costantini, M. (2012). Tie my hands, tie my eyes. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 263–266. doi: 10.1037/a0026570
- Bach, P., Bayliss, A. P., and Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychon. Bull. Rev.* 18, 171–176. doi: 10.3758/s13423-010-0029-x
- Becchio, C., Bertone, C., and Castiello, U. (2008). How the gaze of others influences object processing. *Trends Cogn. Sci. (Regul. Ed.)* 12, 254–258. doi: 10.1016/j.tics.2008.04.005
- Castiello, U. (2003). Understanding other people's actions: intention and attention. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 416–430. doi: 10.1037/0096-1523.29.2.416
- Conty, L., Tijus, C., Hugueville, L., Coelho, E., and George, N. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: a behavioural study. *Spat. Vis.* 19, 529–545. doi: 10.1163/156856806779194026
- de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Curr. Biol.* 18, 454–457. doi: 10.1016/j.cub.2008.02.057
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., and Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cogn.* 6, 509–540. doi: 10.1080/135062899394920
- Eastough, D., and Edwards, M. G. (2007). Movement kinematics in prehension are affected by grasping objects of different mass. *Exp. Brain Res.* 176, 193–198. doi: 10.1007/s00221-006-0749-3
- Edwards, M. G., Humphreys, G. W., and Castiello, U. (2003). Motor facilitation following action observation: a behavioural study in prehensile action. *Brain Cogn.* 53, 495–502. doi: 10.1016/S0278-2626(03)00210-0
- Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9602–9605. doi: 10.1073/pnas.152159999
- Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771. doi: 10.1038/nature01861
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. doi: 10.1126/science.1106138
- Gesierich, B., Bruzzo, A., Ottoboni, G., and Finos, L. (2008). Human gaze behaviour during action execution and observation. *Acta Psychol.* 128, 324–330. doi: 10.1016/j.actpsy.2008.03.006
- Griffiths, D., and Tipper, S. P. (2009). Priming of reach trajectory when observing actions: hand-centred effects. *Q. J. Exp. Psych.* (2006) 62, 2450–2470. doi: 10.1080/17470210903103059
- Hardwick, R. M., and Edwards, M. G. (2011). Observed reach trajectory influences executed reach kinematics in prehension. *Q. J. Exp. Psychol.* 64, 1082–1093. doi: 10.1080/17470218.2010.538068
- Howell, G., and Lacroix, G. L. (2012). Decomposing interactions using GLM in combination with the COMPARE, LMATRIX and MMATRIX subcommands in SPSS. *Tutor. Quant. Methods Psychol.* 8, 1–22.
- Jacob, P., and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends Cogn. Sci. (Regul. Ed.)* 9, 21–25. doi: 10.1016/j.tics.2004.11.003
- Jeannerod, M. (1994). The timing of natural prehension movements. *J. Mot. Behav.* 16, 235–254. doi: 10.1080/00222895.1984.10735319
- Jeannerod, M., Arbib, M. A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320. doi: 10.1016/0166-2236(95)93921-J
- Jellema, T., Baker, C. I., Wicker, B., and Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain Cogn.* 44, 280–302. doi: 10.1006/brcg.2000.1231
- Johansson, R. S., Westling, G., Bäckström, A., and Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *J. Neurosci.* 21, 6917–6932. doi: 10.1136/jnnp.46.6.533
- Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., and Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nat. Neurosci.* 7, 1299–1301. doi: 10.1038/nn1355
- Land, M. F., and Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 352, 1231–1239. doi: 10.1098/rstb.1997.0105
- Land, M. F., and Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Res.* 41, 3559–3565. doi: 10.1016/S0042-6989(01)00102-X
- Langton, S., Watt, R., and Bruce, I. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends Cogn. Sci. (Regul. Ed.)* 4, 50–59. doi: 10.1016/S1364-6613(99)01436-9
- Liepert, R., Von Cramon, D. Y., and Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *Neuroimage* 43, 784–792. doi: 10.1016/j.neuroimage.2008.08.007
- Ondobaka, S., de Lange, F. P., Newman-Norlund, R. D., Wiemers, M., and Bekkering, H. (2012). Interplay between action and movement intentions during social interaction. *Psychol. Sci.* 23, 30–35. doi: 10.1177/0956797611424163
- Pierro, A. C., Becchio, C., Wall, M. B., Smith, A. T., Turella, L., and Castiello, U. (2006). When gaze turns into grasp. *J. Cogn. Neurosci.* 18, 2130–2137. doi: 10.1162/jocn.2006.18.12.2130
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0
- Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–74. doi: 10.1038/nrn2805
- Rotman, G., Troje, N. F., Johansson, R. S., and Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *J. Neurophysiol.* 96, 1358–1369. doi: 10.1152/jn.00227.2006
- Salama, I. M., Turner, S., and Edwards, M. G. (2011). Automatic priming of grip force following action observation. *Q. J. Exp. Psych.* (2006) 64, 833–838. doi: 10.1080/17470218.2011.572172

Acknowledgments

The research was funded by an FRS- FNRS: FRFC grant (2.4587.12). We thank Pierre Mahau and Dominique Hougardy for technical support.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00569/abstract>

- Sartori, L., Becchio, C., and Castiello, U. (2011). Cues to intention: the role of movement information. *Cognition* 119, 242–252. doi: 10.1016/j.cognition.2011.01.014
- Schneider, W., Eschmann, A., and Zuccolotto, A. (2002). *E-Prime Reference Guide*. Pittsburgh, PA: Psychology Software Tools, Inc.
- Senju, A., and Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cogn.* 12, 127–144. doi: 10.1080/13506280444000157
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing. A neurophysiological study. *Neuron* 31, 155–165. doi: 10.1016/S0896-6273(01)00337-3
- Webb, A., Knott, A., and Macaskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta Psychol.* 133, 51–56. doi: 10.1016/j.actpsy.2009.09.001
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psych. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Letesson, Grade and Edwards. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Individual differences in reading social intentions from motor deviants

Daniel Lewkowicz, Francois Quesque, Yann Coello and Yvonne N. Delevoye-Turrell*

SCALab, UMR CNRS 9193, Department of Psychology, Université de Lille, Villeneuve-d'Ascq, France

OPEN ACCESS

Edited by:

Maurizio Gentilucci,
University of Parma, Italy

Reviewed by:

Francesca Ferri,
University of Ottawa, Canada
Alice C. Roy,
Université de Lyon, France

*Correspondence:

Yvonne N. Delevoye-Turrell,
SCALab, UMR CNRS 9193,
Department of Psychology, Université
de Lille, Rue du Barreau,
Villeneuve-d'Ascq, 59653
Nord, France
yvonne.delevoye@univ-lille3.fr

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 08 April 2015

Accepted: 26 July 2015

Published: 17 August 2015

Citation:

Lewkowicz D, Quesque F, Coello Y
and Delevoye-Turrell YN (2015)
Individual differences in reading social
intentions from motor deviants.
Front. Psychol. 6:1175.
doi: 10.3389/fpsyg.2015.01175

As social animals, it is crucial to understand others' intention. But is it possible to detect social intention in two actions that have the exact same motor goal? In the present study, we presented participants with video clips of an individual reaching for and grasping an object to either use it (personal trial) or to give his partner the opportunity to use it (social trial). In Experiment 1, the ability of naïve participants to classify correctly social trials through simple observation of short video clips was tested. In addition, detection levels were analyzed as a function of individual scores in psychological questionnaires of motor imagery, visual imagery, and social cognition. Results revealed that the between-participant heterogeneity in the ability to distinguish social from personal actions was predicted by the social skill abilities. A second experiment was then conducted to assess what predictive mechanism could contribute to the detection of social intention. Video clips were sliced and normalized to control for either the reaction times (RTs) or/and the movement times (MTs) of the grasping action. Tested in a second group of participants, results showed that the detection of social intention relies on the variation of both RT and MT that are implicitly perceived in the grasping action. The ability to use implicitly these motor deviants for action-outcome understanding would be the key to intuitive social interaction.

Keywords: perception, action, social cognition, intention, observation, kinematics

Introduction

Understanding what a conspecific is doing represents a crucial ability for our everyday social interactions. However, perceiving an action and understanding the reason that drives this behavior may arise from different processes (Spaulding, 2015). As highly social species, it is crucial for us to perceive others' mental states and to predict what they plan to do in order to adapt and coordinate our own behavior to the surrounding context (Hamilton, 2009; Sebanz and Knoblich, 2009). As such, our ability to understand the goal of others' actions relies on a variety of sources (Frith and Frith, 2006). For example, declarative knowledge (Fehr and Fischbacher, 2004) and indirect interaction (Singer et al., 2004) are indices that are used when judging the reason of others' behavior. Contextual cues, such as environmental and physical constraints of an action also help to detect the aim of observed actions (Brass et al., 2007; Stapel et al., 2012). However, experimental evidences now support the hypothesis that humans have the ability to predict the action-outcome goals on the basis of the observation of its early kinematics only (Orliaguet et al., 1996; Knoblich and Flach, 2001; Sebanz and Shiffrar, 2009). Indeed, it has been shown that observers are sensitive to early differences in visual kinematics and can use them to discriminate between movements performed with different object-oriented motor intentions (Méary et al., 2005; Manera et al., 2011; Sartori et al., 2011). However, it is the case that most gestures are socially oriented: I can reach for a cup and place it on a table in order to use it myself, but often I will reach for an object to give it to my partner. The question that

will be considered here is then: Can my partner detect in a predictive manner whether the cup that I am grasping for is for her or not, simply by observing my hand moving?

After considering the literature that discusses how intentions may shape movement kinematics, we will state the differences that are rarely made between motor and social intentions in experimental settings. More specifically, we will reveal the individual differences in the ability to detect social intention when simply observing the motor deviants contained within 3D movement kinematics.

Kinematic studies in humans have shown that different motor intentions can shape the spatio-temporal characteristics of a reach-to-grasp movement depending on the goal of the executed sequence (Marteniuk et al., 1987; Armbrüster and Spijkers, 2006; Ansuini et al., 2008; Naish et al., 2013). For example, people tend to produce slower motor actions when grasping an object with the intention to place it accurately rather than with the intention to throw it (Marteniuk et al., 1987; Louis-Dam et al., 1999). In addition, Jacob and Jeannerod (2005) distinguished two types of intentions. The motor intention refers to the mental state that causes the execution of voluntary action (e.g., to put a glass on a table). However, the same motor intention could involve a conspecific (e.g., put the glass on a table for your child) or not (e.g., put the glass on the table to drink from it). This later level of description is referred to as the social intention that is, the intention to affect a conspecific's behavior. According to these same authors, only the motor intention influences the execution of an action, since the same spatial constraints could serve different social intentions. This is known in the literature as the Dr. Jekyll and Mr. Hyde paradox (Jacob and Jeannerod, 2005). Interestingly, recent studies have shade doubt on these affirmations by showing that specific changes in the kinematics of the arm and hand movements can be revealed when investigating the effects of the social context on the execution of motor sequences (Ferri et al., 2011a; Gianelli et al., 2011; Innocenti et al., 2012; Scorolli et al., 2014). But more specifically, it has been suggested that when endorsing a social intention, humans tend to amplify the spatio-temporal parameters of their movements. When planned with a social intention in mind, a subject's hand tends to move with higher hand paths (Becchio et al., 2008; Quesque et al., 2013; Quesque and Coello, 2014), slower velocities (Becchio et al., 2008; Lewkowicz et al., 2013) and longer movement durations (Ferri et al., 2011b; Quesque et al., 2013; Quesque and Coello, 2014). From these variations in execution, it could then be possible for an observer to distinguish different social goals driving similar motor actions.

In the present contribution, we defined the kinematic *deviances* due to social intentions as the systematic difference between the kinematic features [e.g., movement time (MT), peak velocity, peak height] of two executed movements that have the same motor constraints (e.g., start and stop position, object shape, target shape, object initial, and final position) but executed for different social intents. The use of common kinematic features of movements is an important step for researchers to quantify accurately the deviances due to social intentions (Ansuini et al., 2014). Nonetheless, we underline that our definition of the kinematic *deviance* is not restricted to a specific parameter. Rather, we hypothesize that it is

a mechanism that affects multiple components of the movement and its preparation. Thus, the expression of kinematic *deviance* in terms of specific kinematic features could vary depending of the type of action, the target object position and shape, and various other motor constraints. In other words, when changing the motor constraints of an action, one would also change its social *deviance*. Hence, to characterize the kinematic *deviance* due to social intention one needs to disentangle the multiple kinematic features to determine the potential candidates. By controlling precisely the external constraints of executed movements in real-time (Lewkowicz and Delevoeye-Turrell, 2015), it is possible to verify that the significant deviances of kinematic features are not due to specific motor constraints but rather to internal determinants (see also Ansuini et al., 2015), which would give scientific basis for a better understanding of the Dr. Jekyll and Mr. Hyde paradox (Jacob and Jeannerod, 2005). Whereas it has already been shown that the early deviants of kinematic features could be directly exploited to help detect the underlying intention of an observed action (Sartori et al., 2011; Lewkowicz et al., 2013), it is still unclear whether the sensitivity to kinematic deviances is in relationship with the motor expertise or the social skills of the external observer.

A number of recent studies have shown that motor training directly influences action observation (Hecht et al., 2001; Casile and Giese, 2006). Especially in the case of very skilled observers, for example, in sports (Abernethy and Zawi, 2007; Abernethy et al., 2008; Aglioti et al., 2008), and artistic activity (Calvo-Merino et al., 2005, 2006), experts systematically outmatch novices in recognizing and predicting the outcome of observed action but also in detecting deceptive intentions (Jackson et al., 2006; Cañal-Bruland and Schmidt, 2009; Sebanz and Shiffrar, 2009). These results are in line with the hypothesis that common codes for perception and action (Prinz, 1997; Hommel et al., 2001) can be used to simulate observed actions and thus, gain a better prediction and understanding of motor outcome (Blakemore and Decety, 2001; Jeannerod, 2001; Wolpert et al., 2003; Grush, 2004; Wilson and Knoblich, 2005; Uithol et al., 2011). In addition, within the framework of the mirror neuron system (Cattaneo and Rizzolatti, 2009), it has been claimed that the same mechanisms would be involved during the imagery of a motor act directed to an object and the actual execution of that same motor act (e.g., Jeannerod and Decety, 1995; Ehrsson et al., 2003; Decety and Grèzes, 2006). The ability to detect social deviants should then be correlated to motor expertise and simulation abilities.

The processing of others' movements is also very important for communication and adaptive social behavior. Individuals who exhibit deficits in visual biological motion processing are also compromised on daily-life social perception (see Pavlova, 2012, for a review). When one interacts with another person, it is assumed implicitly that common thoughts are shared. Thus, in social contexts, we unconsciously spend time predicting the behavior of others on the basis of what we would do ourselves in the same situation. One may up to a certain extent try to place our own self within the other person's mind, beliefs and desires. This complex cognitive function is referred to as having a "theory of mind" (Premack and Woodruff, 1978), taking an intentional stance (Dennett, 1987), or mentalizing (Frith, 1989).

Mentalizing has been studied using a wide range of tasks including reading stories (Fletcher et al., 1995; Saxe and Kanwisher, 2003), looking at cartoons (Brunet et al., 2000; Gallagher et al., 2000), and watching simple animations (Castelli et al., 2000). It has recently been proposed that during action observations the neural basis of the “theory of mind” is more recruited when the observer is explicitly asked to interpret the scene in terms of high-level goals than it is when focusing on lower-level intentions (Van Overwalle and Baetens, 2009). In such a case, recognizing social deviants may be associated to the same mechanisms, which participate in the recognition process of body and face for social cognition.

In the current study, our goal was to test whether by maintaining the motor intention identical an observer is still able to dissociate between social and personal intentions in movements performed toward an object. After recording trials of actors performing social and personal reach to grasp actions and verifying that the kinematics were indeed dissociable, we conducted two experiments of action observation in which the participants’ task was to categorize trials as a function of their social scope. In *Experiment 1*, we were specifically interested in the individual differences that may be observed in the ability to read social intentions. In order to gain an insight in the psychological factors that may be involved in the capacity of participants to understand the social action-outcome, we used questionnaires to capture both social cognition and motor imagery abilities. In *Experiment 2*, we probed the *nature* of the kinematic deviances of observed movements, which contributed to the categorizing of social and personal intentions. For this, we used post-recording treatments in order to control precisely for the amount of temporal information available within the movie clips. Through the alterations of specific properties of 3D motion kinematics, we were able to test the effects of a progressive normalization of deviances on the participants’ ability to categorize the action as being personal or social.

Experiment 1: Inter Individual Differences to Recognize Social Patterns

In this first study, we tested whether the ability to recognize social intention through the decoding of social deviants was related to motor imagery and/or social cognition abilities.

Materials and Methods

Participants

Twenty-five healthy young adults (seven males; mean age: 24.7; SD: 3.0) participated in the experiment. All had normal or corrected-to-normal vision and had no prior knowledge of the experimental goals. They gave informed consent before participating in the experimental session that lasted approximately 30 min. The protocol received approval from the ethics committee for Human Sciences of the University of Lille 3.

Apparatus and Stimuli

Stimuli

To create the experimental material, we filmed two naïve adults seated at a table, facing each other, and participating in a short

cooperative game. The game consisted in displacing a little wooden dowel (width 2 cm; height 4 cm) between the thumb and the index finger to different locations. Their sequential actions were time-locked to a series of broadcasted sounds. The first move of the game was always performed by the same member of the dyad (named here, the “actor”) and consisted in displacing the dowel from an initial location to a central target. After this preparatory action, a subsequent main action was to be performed either by the actor (*personal* condition) or by the partner (*social* condition). Two blocks of 15 trials were performed: In one block, the actor performed all the preparatory and the main actions, the partner being just an observer. In the other block, the actor performed the preparatory actions and the main actions were always performed by the partner. Meanwhile, the actor’s movements were recorded using a video camera (Logitech webcam model c270) to record the scene. In addition, four Oqus infrared cameras (Qualisys system) were used to record the upper-body kinematics. Five infrared reflective markers were placed on the index (base and tip), the thumb (tip), the wrist (scaphoid and pisiform) of the actor; one marker was placed at the top of the object. The calibration of the cameras provided the means to reach a standard deviation smaller than 0.2 mm, at a 200 Hz sampling rate.

A particular attention was taken to suppress all contextual information from the video clips (see **Figure 1A**). Only the arm of the actor and the target object were framed within the video clips of the 30 preparatory actions. The video clips that were used as stimuli consisted in a sequential action of two motor elements (1) reach to grasp and (2) move to place. The video clips were cut exactly one frame after the actor finished placing the object. Movies were compressed with FFdshow codec (MJPEG) at 30 frames per second with a screen resolution of 640×480 pixels. 3D kinematics were analyzed with RTMocap toolbox (Lewkowicz and Delevoye-Turrell, 2015). Positional data points were filtered using a dual fourth-order Butterworth low-pass filter ($f_c = 15$ Hz; forward and backward) and tangential 3D instantaneous velocities were calculated. A threshold of $20 \text{ mm} \cdot \text{s}^{-1}$ was used to determine the onset of movement (reaction time, RT). All velocity trajectories were bell shaped and consisted in two “bells,” the first corresponding to the reach to grasp element, the second being the move to place element of the preparatory action. The amplitude of peak velocity of the first element (APV_1) was extracted using the local maxima (first 0-crossing of acceleration). The end of the first element was determined as the time of occurrence of the local minima (second 0-crossing of acceleration) between the first and the second element-peaks (see *vertical arrow* in **Figure 1**). The duration of the first element (MT_1) was calculated as the time interval between the onset and the end of the first element. The amplitude of the peak height of trajectories (APH_1) was defined as the maximum z coordinate of the wrist measured in the grasping element and the lift to place element. APV_2 , MT_2 , and APH_2 are the corresponding kinematic parameters described above but extracted from the second move to place element of the motor sequence. **Table 1** presents the characteristics of the movement parameters that were measured, e.g., RT, MT, peak wrist velocity, and height of hand trajectory. **Figure 2** presents the scatterplot of amplitude of peak velocity against MT in order to confirm none

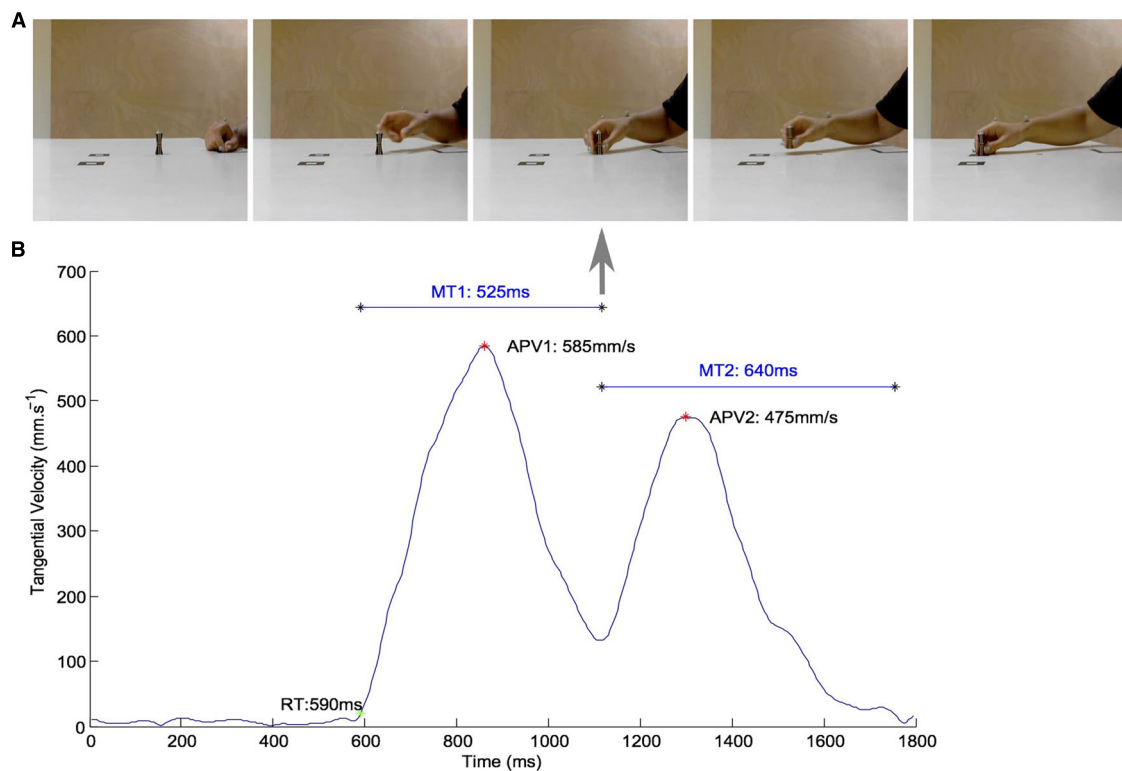


FIGURE 1 | (A) A typical example of the video stimuli that was used both in Experiments 1 and 2 to test the role of motor deviants for the categorization of social and personal object-centered actions. One can note the neutral context that was used with the placement of 3D reflexive markers that provided us the means to verify the kinematic deviants between social and personal movements

during the validation phase of the video database. **(B)** Velocity curves of the corresponding trial illustrating the double bell shaped profiles that are observed in the present reach to grasp task. Reaction times (RT in ms) and movement times of the first element of the sequence (MT of reach in ms) may have been used by the observers to dissociate social from personal actions.

TABLE 1 | Mean kinematic parameters of the preparatory action for both the personal and the social trials.

	RT	APV1	APV2	MT1	MT2	APH1	APH2
	**	*	*	***	*	**	*
Personal intention	616	582	525	440	508	58	63
Social intention	702	547	457	478	545	60	65
Median values	665	572	487	457	533	59	64
Frequency of personal trials > median	4/15	10/15	5/15	3/15	5/15	4/15	5/15
Frequency of social trials > median	11/15	5/15	10/15	12/15	10/15	11/15	10/15

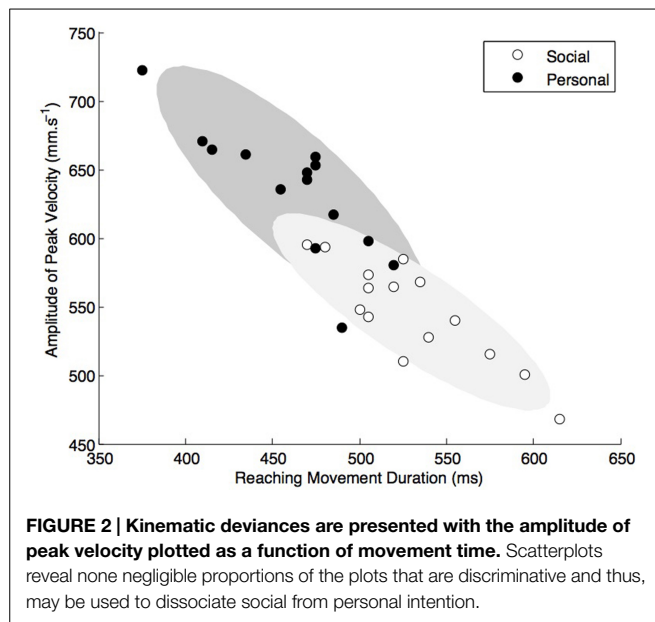
For each parameter, the median values for the totality of the trials are reported and the frequency of trials superior to this value is specified in each condition. RT, reaction time; APV, amplitude of peak velocity; MT, movement time; APH, amplitude of peak hand height, for the first (1) reaching element or the second (2) grasping element. The asterisks revealed the parameters for which significant differences were found between the two distributions in the personal and the social conditions using the median test (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

negligible proportions of the plots that are discriminative between social and personal trials. Using comparison to the median values, pre-analysis confirmed the possibility to dissociate personal from social trials on the basis of RT, MT and height of grasping phase (APH).

Individual evaluations of social and imagery sensitivity

The *Reading the Mind in the Eyes Test*, which will be referred to as the RME-test in the following sections (Baron-Cohen et al., 1997, 2001) was designed to measure each individuals' sensitivity to social cues and in particular the participants' ability

to understand others' complex mental states. This test has shown a high potential to distinguish an individual's tendency to attend to others' intentions in joint cognitive tasks (Ruys and Aarts, 2010). In the RME-test, participants were required to categorize eye-regions of 36 facial expressions by selecting a mental state label that matched the perceived expression, selecting one out of the four terms proposed. In the present experiment, participants completed a French version of this test (Prevost et al., 2014) and were encouraged to select the appropriate term as fast as possible. Overall, the more people attend to the intentions of others, the higher are their scores on the RME-test. We also administered



a French version (Loison et al., 2013) of the Movement Imagery Questionnaire—Revised Second version (MIQ-RS, Gregg et al., 2010) of the Movement Imagery Questionnaire—Revised (MIQ-R, Hall and Martin, 1997). This questionnaire is a reliable measure of motor imagery that distinguishes kinesthetic motor imagery from visual motor imagery. Participants were required to perform and imagine daily life actions that were similar in the two subscales, involving both upper and lower limbs.

Procedure

Participants were seated at a table in a silent experimental box, facing the experimenter. They took part in a short cooperative game to get familiarized with the paradigm. These pre-test trials consisted in similar manipulative movements than that performed by the actor in the stimuli video. Participants performed 15 trials for which they were required to pick and place a wooden dowel at the center of the table for their own purpose and 15 trials for which the wooden dowel was picked and placed for the experimenter. After this familiarization phase, participants were instructed to watch and categorize previously recorded videos clips from the same two conditions. Participants had to categorize a total of 30 videos (15 social and 15 personal). The instructions before categorization were given orally as follow (“Is the actor placing the dowel for a personal use?” OR “Is the actor placing the dowel to give it to his partner?”).

The videos stimuli in the categorization task were displayed on a gray background on a laptop computer using the *PsychToolbox* for Matlab (Natick, MA, USA). Before each trial, a white fixation cross appeared on the gray screen during a variable interval of 500–1000 ms. After each video presentation, as soon as the clip ended, a blank screen was shown during which participants were prompt to give their decision. They were instructed to categorize each movie clip as fast and as accurately as possible. The response keys were marked with tape placed directly on the azerty computer keyboard (“a” for social and “p” for personal). The response keys

were counterbalanced across participants. No feedback was given during the experiment. Finally, the participants were required to complete the French version of the RME-test and the MIQ-R. The order of presentation of the two tests was also counterbalanced across participants. After the entire completion of the experiment, participants were asked to comment on the general degree of confidence that they had in their answers in the categorization task. Finally, participants obtained a short debriefing period and were thanked for their participation.

Analysis

Response times were calculated as the time interval between the presentation of the last frame of the video and the participant’s key press. For the analyses of the amount of correct responses, it is to note that in our experiment the error in judging one kind of stimulus (e.g., social) was redundant with the correct judgment of the other kind of stimulus (e.g., personal). Consequently, the results were expressed in total percentage of correct responses (Bond and DePaulo, 2006). Scores for each category were compared to the reference constant, i.e., the random answer value of 0.50, with a single sample *t*-test. To test whether the classifications rates would entail any substantial individual differences in the perception of social intention, we performed correlation analyses. We then checked whether the percentage of correct responses was correlated with the social cognition measure and with the motor and visual imagery measures, separately. Final score in the French version of the *RME-test* was computed on 34 items, excluding the items 13 and 23 from analysis as recommended (Prevost et al., 2014). Concerning the imagery measures, the two scores (kinesthetic; visual) were calculated on a 7 points scale. All analyses were conducted two-tailed and the alpha level of significance was set to 0.05.

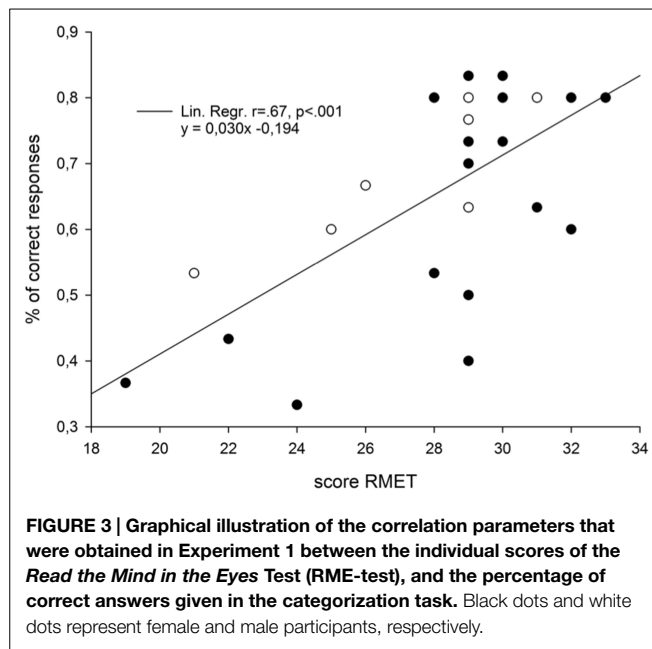
Results

Categorization Performance and Response Time

The results revealed that on average participants were able to categorize the underlying intention above chance level ($M = 65.7\%$, $SD = 15.8$ vs. 50%), $t(24) = 4.980$, $p < 0.001$. There were no significant differences in the percentage of correct categorization for the personal intention ($M = 68\%$, $SD = 19.7$) and the social intention ($M = 63.4\%$, $SD = 19.8$), $t(24) = 0.95$, $p = 0.35$. Moreover, the results revealed no significant effects of the stimulus type on mean response times. Participants categorized the video clips presenting a personal intention as quickly ($M = 600$ ms, $SD = 0.39$) as the video clips presenting a social intention ($M = 570$ ms, $SD = 0.32$), $t(24) = 0.58$, $p = 0.58$.

Correlation With Individual Traits

On average, participants obtained a score of $M = 5.8$, $SD = 1.2$ in visual imagery and $M = 4.8$, $SD = 1.3$ in kinesthetic imagery as assessed by the Movement Imagery Questionnaire. The results revealed an absence of correlation with the percentage of correct categorization for both the visual imagery score ($R = 0.125$, $p = 0.551$) and the kinesthetic imagery score ($R = 0.194$, $p = 0.354$). The results of the *RME-test* revealed a mean score of 28.24 , $SD = 3.5$. Our results showed that the *RME-test* scores were



positively correlated with the percentage of correct categorization ($R = 0.677$, $p < 0.001$), indicating that a higher score in the RME-test is associated to a higher performance in the categorization task (see **Figure 3**). Concerning the degree of relationship between the questionnaires, the RME-test scores were related neither to the kinesthetic imagery scores ($R = 0.006$, $p = 0.975$) nor to the visual imagery scores ($R = 0.278$, $p = 0.178$). Finally, the scores on the two dimensions of the Movement Imagery Questionnaire were not correlated ($R = 0.132$, $p = 0.527$).

Discussion

The aim of this experiment was to test for the individual differences that may be observed in the ability to read social intentions. Firstly, confronted to short video clips of “pick and place” moves, participants were able to categorize the intention (“social” vs. “personal”) of the actor above chance level. Given the effort made to produce stimuli presenting an absence of contextual information, this result confirms the idea that not only motor intention (Méary et al., 2005; Manera et al., 2011; Sartori et al., 2011; Lewkowicz et al., 2013) but also social intention can be inferred from the kinematics of a movement, as suggested by Ansuini et al. (2015). Secondly, it is to note that not all participants were equally talented in performing the task. Particularly, the ability of participants to discriminate between social and personal intentions was highly linked to the scores obtained in the social cognition test but was not related to the scores obtained in the motor imagery questionnaires. Such dissociation corroborates recent findings showing that sensitivity to use subtle cues in biological motion is linked to social but not to motor imagery measures (Miller and Saygin, 2013). More specifically as reported here, the authors showed that form cues correlated more with the social than with the imagery measures suggesting that even if social cognition and motor imagery predict sensitivity to biological motion, these skills tap into different

aspects of perception. In our case, the results comfort the idea that social abilities help detect modulations of trajectories even in very simple and fast motor actions such as a reach to grasp task performed at natural speed.

Experiment 1, gave us the opportunity to assess participants’ ability to perceive social intentions from motor actions. However, it did not give us insights on the actual perceptual cues used by participants to solve the decision task. Consequently, in Experiment 2, we focused on the question of “how” participants could perceive social intentions from motor actions. For this purpose, we used post-recording modifications of videos clips in order to determine which crucial aspects of the kinematic deviants were relevant for participants in making their categorization decision. Finally, during the debriefing sessions of Experiment 1 the vast majority of participants reported that they felt as if they responded randomly in the categorization task, reporting a very low degree of confidence in their responses. However, due to the absence of quantitative measures of the meta-cognitive judgments from the participants, it was not possible to draw straight conclusions. Experiment 2 gave the opportunity to investigate this point more rigorously by obtaining systematic auto-evaluation of metacognitive knowledge through the use of analogical-scales.

Experiment 2: Content Information to Recognize Social Patterns

This study was conducted to assess whether participants could distinguish between social and personal movements even after the specific properties of the 3D motor kinematics were flattened out.

Materials and Methods Participants

Twenty-three healthy young adults (six males; mean age: 25.8; SD: 5.0) participated in the second experiment. All had normal or corrected-to-normal vision and had no prior knowledge of the experimental goals. These participants did not take part in Experiment 1 and gave informed consent before participating in the experimental session that lasted approximately 20 min. All participants completed in a previous session the French version of the RME-test (Prevost et al., 2014) and only those who had a minimal score of 27 (corresponding to the French median score) were selected to take part in the following experiment. The protocol received approval from the ethics committee for Human Sciences of the University of Lille 3.

Apparatus and Stimuli

In this experiment, two-step actions were recorded from a different actor but following the same design as in Experiment 1 in order to generate new stimuli videos. **Table 2** presents the characteristics of actions parameters in the personal and social condition. As expected, significant differences were obtained in the 3D motion kinematics between personal and social trials for many motor parameters and especially those that will be manipulated, i.e., RT and MT of the first element of the motor sequence (MT1).

TABLE 2 | Mean kinematic parameters of the preparatory action for both the personal and the social trials.

	RT	APV1	APV2	MT1	MT2	APH1	APH2
	*		*	**			**
Personal intention	395	590	529	417	501	58	58
Social intention	438	618	487	451	475	63	65
Median values	408	599	509	438	485	61	63
Frequency of personal trials > median	5/15	9/15	11/15	4/15	9/15	6/15	4/15
Frequency of social trials > median	10/15	6/15	5/15	11/15	7/15	9/15	11/15

The asterisks revealed the parameters for which significant differences were found between the two distributions in the personal and the social conditions using the median test ($p < 0.05$; ** $p < 0.01$).

In order to control for the amount of temporal and kinematic information available to participants, we used post-recording modification of the videos. This manipulation led to creation of three types of stimuli. Indeed, depending on the condition, the stimuli that were displayed could be the original video clips ($RT + MT_1$ deviant), video clips normalized according to RTs (MT_1 deviant) or video clips normalized according to the end of the grasping action (*No deviant*).

The modification of each video clip was achieved on-line as follows. First, the mean of the parameters that needed to be homogenized was calculated across all trials (social and personal). Second, the video clips were displayed at an overall refreshment rate so that the display time of this parameter corresponded to the mean pre-determined value. For example, in the MT_1 deviant condition, the parameter that needed to be homogenized was the RT. Thus, using the kinematic data, a deviance ratio was calculated for the section of the video clip corresponding to the overall rate at which the RT section of the video should be presented in order to match the mean pre-determined value. We then interpolated the video frames (30 Hz) with the true refreshment rate of the screen (60 Hz) and replaced each video frame accordingly to the deviance ratio scaled to this final refreshment rate. In other words, the modifications brought to the duration of each video clip was spread out through the successive frames rather than being performed through an abrupt modification a given section of the video (e.g., by removing a frame). This manipulation gave us the opportunity to maintain the majority of the biological content of each movement.

Except for the modifications brought to the videos, the experimental design was identical to the one used in Experiment 1. In addition, analogical scales (10-cm long lines coding for “chance level” to the far left and “high confidence” to the far right) were included at the end of each trial in order to gain information about the metacognitive knowledge that participants’ possessed on their self-evaluation performances.

Procedure

Participants were seated at a table in a silent experimental box and had to perform the categorization task with the same instructions as in Experiment 1. They categorized the three sets of videos in three distinct sessions that were completed in a random order (counter-balanced across participants). After each session, they were asked to auto-evaluate the trust

they had in their present classification rate on analogical scales.

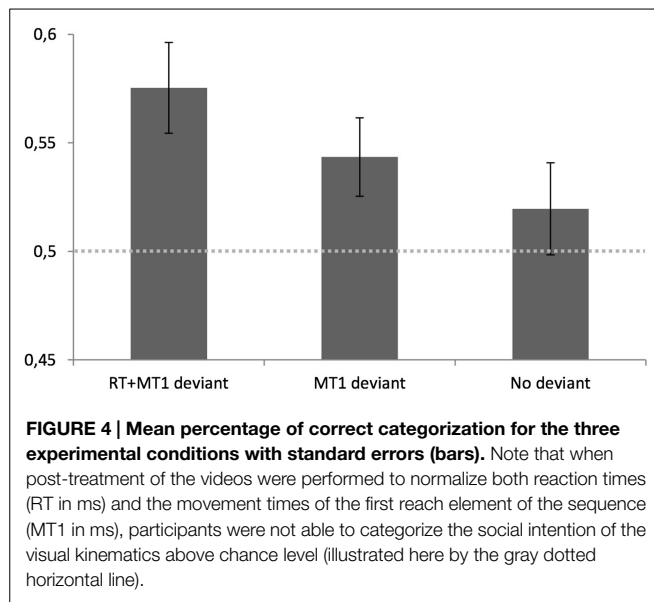
Analysis

Mean percentages of correct responses, mean response times and mean self-evaluation scores were calculated for each condition and submitted to a repeated-measure ANOVA with condition ($RT + MT_1$ deviant, MT_1 deviant, *No deviant*) as within factors. The *post hoc* Bonferroni test was used when needed. We also conducted sub-analyses for the percentages of correct responses: scores for each category were compared to the reference constant, i.e., the random answer value of 0.50, using a single sample *t*-test. All analyses were conducted two-tailed and the alpha level of significance was set to 0.05.

Results

A repeated measures ANOVA revealed an effect of video type [$F(1,22) = 3.02$, $p = 0.05$] on the percentage of correct categorization. *Post hoc* contrast analysis revealed a significant higher rate of correct judgments in the natural condition ($M = 57.5\%$, $SD = 10$) compared to the $RT + MT_1$ deviant condition ($M = 51.9\%$, $SD = 10$; $t = 2.32$, $p < 0.05$). Furthermore, the performances in the MT_1 deviant condition were located in the middle range ($M = 54.3\%$, $SD = 08$) not differing statistically from the two other conditions ($t = -0.22$, $p = 0.83$), suggesting a progressive decrease across the three experimental conditions. Two-sided *t*-tests comparing performances against chance level (50%) in the categorization task revealed that participants were significantly above chance in two of the three conditions (see **Figure 4**). More specifically, participants were able to categorize the underlying intention above chance level when videos were presented in the $RT + MT_1$ deviant condition [$t(22) = 3.6$, $p < 0.01$] and in the MT_1 deviant condition [$t(22) = 2.4$, $p < 0.05$]. However, they were not able to respond above chance level when videos were presented in the *No deviant* condition [$t(22) = 0.9$, $p = 0.37$].

Concerning response times, we found no significant effects of video type [$F(1,22) = 2.19$, $p = 0.15$]. Furthermore, the participants’ responses on the analogical scales used to evaluate metacognitive knowledge about performance self-assessment did not differ between conditions [$F(2,44) = 0.02$, $p = 0.98$]. With an overall mean of 68%, these observations indicate that participants found the task feasible but did not explicitly judge that a certain type of video was harder to categorize than another.



Discussion

The driving question in the second study was to replicate those findings presented in Experiment 1 and assess to what extent kinematic deviants may be used to discriminate social intention in actions that have an exact same motor goal. As in study 1, participants were thus presented with short video clips and were asked to categorize the social intention of the actor. However, these video clips contained different amounts of informative deviants as the videos could be totally informative (original videos as in Experiment 1), partially informative (videos were normalized to RTs) or none informative (videos were normalized to the end of the grasping action). Using video clips of a different naïve actor, we replicated here the results reported in Experiment 1: individuals are able to distinguish between social intention and personal intention through the simple observation of motor kinematics. The fact that the overall categorization performance in the second study was lower than that seen in the first study could be due to the present of fewer kinematic deviances in the stimuli material. It is the case that when comparing trials in the social and the personal conditions, the kinematic analyses revealed more differences in Experiment 1 than in Experiment 2. It is true that in daily social interactions, the actions of certain individuals are easier “to read” than others. This situation—that we all have experienced, is reflected here by the fact that the actor who participated in Experiment 2 had kinematic variances that were less marked than the one participating in Experiment 1. Thus, our findings suggest that the kinematic signature of social intention is difficult to detect within a unique individual. Nevertheless, even if the amount of kinematic information was less present in Experiment 2, we were still able to cancel out the participants’ ability to read social intention through the modification of the kinematic features. Hence, social intention—even if weak, is contained within the kinematic variances of body movement.

The second important result that confirmed our initial hypothesis of the importance of motor deviants for intention reading was that the percentage of correct identification was

proportional to the amount of deviants contained within motor kinematics. The original clips were better categorized than those stimuli that were partially normalized, suggesting that the categorical decisions were based on a spatio-temporal integration of that information contained within the actor’s movements. By asking participants to use analogical scales to self-evaluate performance levels, we furthermore showed that performance levels are not dependent on an explicit conscious decision process. Indeed, even if the percentage of correction identification was significantly affected by the deterioration of the video content, the participants’ metacognitive judgment was not. Participants did not explicitly detect differences in the informative values of the video clips and furthermore, did not judge their performance in the categorization task as being better or worse as a function of the informative content of the videos. Overall, these findings reveal the implicit nature of motor deviants to facilitate social interaction and confirm previous results found in the social literature suggesting that contextual information modulates social behaviors outside of awareness (Knoblich and Sebanz, 2008).

General Discussion

Previous behavioral studies have revealed that the context in which object-oriented actions take place and their relevance for human interactions can affect the way very simple actions are executed (Ferri et al., 2011a; Gianelli et al., 2011; Innocenti et al., 2012; Scorolli et al., 2014). In the present contribution, we were interested in assessing the effects of social context on the temporal and the spatial parameters of hand trajectory in the basic action of reaching for and grasping an object, either to move it for self directed purposes (personal intention) or for the use of the object by a partner (social intention). Our question was the following: Could a naïve observer of the scene detect that the object was going to be reached with a social intention? What in the behavioral dynamics could be used as social cues? This experimental situation is very similar to that observed in daily experiences for which many of our interactions with conspecifics are not conveyed through language. For instance, it has been shown that both structural and dynamic information of body movement through space and time are taken into account for the recognition of point light-display of moving humans (Troje et al., 2005), or for the recognition of another’s emotions when the facial expression is not visible (Atkinson et al., 2004; Meeren et al., 2005). Likewise, in the present contribution, we showed that it is possible for a naïve observer to understand social intention of individuals performing an object-oriented motor action.

Movies were taken from a situation in which a participant picked up and placed an object knowing in advance whether herself or a partner will perform the next action in the sequence. With this method, we created stimuli in which kinematic variants (RT, MT and trajectory height) were the only factor conveying social meaning. Even though the kinematic *variations* due to social intention were small (a few millimeters within a few tens of milliseconds), motor deviants were present in our trajectories in a very repetitive and distinctive way (see **Figure 2**) confirming other

experimental results reported in social oriented tasks (Becchio et al., 2008; Quesque et al., 2013). Here, we confirm in two different sets of actors that human observers are able to exploit these very small kinematic deviances to discriminate the social intention above chance level.

In Experiment 1, we focused on the personal determinants, which could explain inter individual differences in the ability to read the social intention of an action. We thus hypothesized that intention reading would be associated to an individual's competence to either infer complex mental states to others or to use motor imagery to predict motor outcome from movement kinematics. We only found a positive correlation with the social skill as it was previously reported with biological motion processing (Miller and Saygin, 2013). The existence of a close relation between social abilities and the perception of social intention is not surprising as such. Whereas healthy adults are able to perceive intentions (Runeson and Frykholm, 1983; Blakemore and Decety, 2001) and emotions from point-light displays (Dittrich et al., 1996; Pollick et al., 2001; Atkinson et al., 2004; Grezes et al., 2007), this ability seems to be clearly impaired in patients showing deficits in social interactions such as in autism (Blake et al., 2003; Freitag et al., 2008; Parron et al., 2008; Cook et al., 2009; Centelles et al., 2012) and schizophrenia (Kim et al., 2005, 2011). The question that remains is then why does the correct discrimination of social intention not correlate with the motor imagery ability of the observer? We found that increased ability in motor imagery does not in itself help participants to understand correctly the social intention of the movement. One possible interpretation is that the motor imagery questionnaire probes more heavily the explicit processing of motor activity (e.g., goals, conscious monitoring) rather than the implicit sensitivity to subtle kinematic variations.

In Experiment 2, we focused on the hypothesis according to which observers may be able to read the social intention through the exploitation of the kinematic *deviances* between two movements executed with the same motor intention but different social intention. With post-recording treatments, we impoverished the temporal aspects of visual kinematics contained within the video clips to cancel out the ability to read social intention, confirming the central role of these temporal deviants in predicting social outcome. It is now generally accepted that when we execute a movement, we predict the sensory consequences of that movement through generative or forward models (Wolpert et al., 1995, 2003; Wolpert and Miall, 1996). These predictions can then be used to refine motor control problems induced by delayed feedback and sensory noise, but can also play a role to determine the most likely outcome of an observed action (Kilner et al., 2007). It has recently been suggested that a similar system can be used to understand others mental states (Oztop et al., 2005) and more specifically intentions (Ansuini et al., 2015). The results presented here confirm this hypothesis by showing that without temporal deviants, individuals lose the ability to categorize social outcome. These findings indicate that predictive timing may also be the key to the ability of decoding social intention through the observation of motor kinematics. Interestingly, break points were also relevant: RT normalization (in MT₁ deviant condition) was here shown

to also decrease categorization accuracy. This is congruent with previous studies that have shown that individuals are able to infer the subjective confidence of another person simply through the observation of RTs (Patel et al., 2012). Hence, those cognitive states that are based on predictive temporal properties may be correlated to social skills. Future studies need now to generalize these ideas and confirm that social reading is dependent on the accumulation of prediction errors, i.e., not only on *the when* but also on the *how long* of an on-flowing action sequence. Here we suggest that this would be done through the multi-integration of temporal deviants within a bilateral interaction of top-down and bottom up processes (see also Hillebrandt et al., 2014, for a neuro-anatomical account of this perspective).

It is the case that studies have reported gender effects related to social reading (Alaerts et al., 2011; Sokolov et al., 2011). Our results could suffer from the fact that a greater number of female individuals participated in the study. However, the gender main effect was none significant with the male participants performing at similar levels than the female participants both in the RMET and in the categorization task (see **Figure 3**). Furthermore, the tendency for woman to do better than men in the RMET was significantly true in the first version of the test (Baron-Cohen et al., 1997) but this was only marginally the case in the second version of the test (Baron-Cohen et al., 2001), which is the one we used. Finally, recent studies assessing the gender question have shown that men even sometimes do better than woman, e.g., in tasks using point-light displays to recognize human locomotion (Krüger et al., 2013). Hence, our results indicate that individual characteristics are more valuable to predict within gender abilities than the general gender property itself. They are novel and confirm the usefulness of RMET for predicting individual performances in (1) the recognition of body language (Alaerts et al., 2011; Miller and Saygin, 2013) and (2) the ability to detect other's intention through body movements (Ruys and Aarts, 2010), whether that person be a man or a woman. A second point to note is the importance in future studies to assess whether the results presented here can be generalized to more ecological tasks. Indeed, the method presented here using video clips could be further applied to create experimental situations at second-person perspective including, for instance, two participants performing a reach to grasp task in a real interactive situation (see illustrated examples on line through reference keys given in Lewkowicz and Delevoye-Turrell, 2015). Furthermore, demonstrating that similar patterns of results are obtained when not only two but multiple intentional possibilities are presented would provide more ecological validity for the social abilities reported in the present study (see Obhi, 2012).

In conclusion, the present study reveals that the ability to implicitly use motor deviants from observed object-directed actions represents the crucial factor for detecting social intention. Furthermore, this ability seems to depend on individual social cognition skills. Implicit judgments are often considered as intuitive. As such, intuition has been defined in the field of human robotics as our ability for direct knowledge, for immediate insight without explicit reasoning. Intuitive judgments are more or less accessible to individuals depending

on a number of factors (e.g., physical salience, emotional and motivational states, Kahneman, 2003). In the present study, we suggest that an important aspect of intuitive interaction is the power to detect the contained information within the temporal aspects of body movements to prime the social expectancy of an observer.

References

- Abernethy, B., and Zawi, K. (2007). Pick-up of essential kinematics underpins expert perception of movement patterns. *J. Motor Behav.* 39, 353–367.
- Abernethy, B., Zawi, K., and Jackson, R. C. (2008). Expertise and attunement to kinematic constraints. *Perception* 37, 931–948. doi: 10.1068/p5340
- Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., and Wenderoth, N. (2011). Action and emotion recognition from point light displays: an investigation of gender differences. *PLoS ONE* 6:e20989. doi: 10.1371/journal.pone.0020989
- Aglioti, S. M., Cesari, P., Romani, M., and Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nat. Neurosci.* 11, 1109–1116. doi: 10.1038/nn.2182
- Ansuini, C., Cavallo, A., Bertone, C., and Becchio, C. (2014). The visible face of intention: why kinematics matters. *Front. Psychol.* 5:815. doi: 10.3389/fpsyg.2014.00815
- Ansuini, C., Cavallo, A., Bertone, C., and Becchio, C. (2015). Intentions in the brain: the unveiling of Mister Hyde. *Neuroscientist* 21, 126–135. doi: 10.1177/1073858414533827
- Ansuini, C., Giosa, L., Turella, L., Altoè, G., and Castiello, U. (2008). An object for an action, the same object for other actions: effects on hand shaping. *Exp. Brain Res.* 185, 111–119. doi: 10.1007/s00221-007-1136-4
- Armbrüster, C., and Spijkers, W. (2006). Movement planning in prehension: do intended actions influence the initial reach and grasp movement? *Motor Control* 10, 311–329.
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., and Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 717–746. doi: 10.1068/p5096
- Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008). The case of Dr. Jekyll and Mr. Hyde: a kinematic study on social intention. *Conscious. Cogn.* 17, 557–564. doi: 10.1016/j.concog.2007.03.003
- Blake, R., Turner, L. M., Smoski, M. J., Pozdol, S. L., and Stone, W. L. (2003). Visual recognition of biological motion is impaired in children with autism. *Psychol. Sci.* 14, 151–157. doi: 10.1111/1467-9280.01434
- Blakemore, S. J., and Decety, J. (2001). From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* 2, 561–567. doi: 10.1038/35086023
- Bond, C. F., and DePaulo, B. M. (2006). Accuracy of deception judgments. *Pers. Soc. Psychol. Rev.* 10, 214–234. doi: 10.1207/s15327957pspr1003_2
- Brass, M., Schmitt, R. M., Spengler, S., and Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Curr. Biol.* 17, 2117–2121. doi: 10.1016/j.cub.2007.11.057
- Brunet, E., Sarfati, Y., Hardy-Bayle, M. C., and Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage* 11, 157–166. doi: 10.1006/nimg.1999.0525
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007
- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr. Biol.* 16, 1905–1910. doi: 10.1016/j.cub.2006.07.065
- Cañal-Bruland, R., and Schmidt, M. (2009). Response bias in judging deceptive movements. *Acta Psychol.* 130, 235–240. doi: 10.1016/j.actpsy.2008.12.009
- Casile, A., and Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Curr. Biol.* 16, 69–74. doi: 10.1016/j.cub.2005.10.071
- Castelli, F., Happe, F., Frith, U., and Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* 12, 314–325. doi: 10.1006/nimg.2000.0612
- Cattaneo, L., and Rizzolatti, G. (2009). The mirror neuron system. *Arch. Neurol.* 66, 557–560. doi: 10.1001/archneurol.2009.41
- Centelles, L., Assaiante, C., Etchegoyhen, K., Bouvard, M., and Schmitz, C. (2012). From action to interaction: exploring the contribution of body motion cues to social understanding in typical development and in autism spectrum disorders. *J. Autism Dev. Disord.* 43, 1140–1150. doi: 10.1007/s10803-012-1655-0
- Cook, J., Saygin, A. P., Swain, R., and Blakemore, S. J. (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. *Neuropsychologia* 47, 3275–3278. doi: 10.1016/j.neuropsychologia.2009.07.010
- Decety, J., and Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain Res.* 1079, 4–14. doi: 10.1016/j.brainres.2005.12.115
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge: MIT Press.
- Dittrich, W. H., Troscianko, T., Lea, S. E., and Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception* 25, 727–738. doi: 10.1068/p250727
- Ehrsson, H. H., Geyer, S., and Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part specific motor representations. *J. Neurophysiol.* 90, 3304–3316. doi: 10.1152/jn.01113.2002
- Fehr, E., and Fischbacher, U. (2004). Social norms and human cooperation. *Trends Cogn. Sci.* 8, 185–190. doi: 10.1016/j.tics.2004.02.007
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2011a). Social requests and social affordances: how they affect the kinematics of motor sequences during interactions between conspecifics. *PLoS ONE* 6:e15855. doi: 10.1371/journal.pone.0015855
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2011b). To me or to you? When the self is advantaged. *Exp. Brain Res.* 203, 637–646. doi: 10.1007/s00221-010-2271-x
- Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., et al. (1995). Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 57, 109–128. doi: 10.1016/0010-0277(95)00692-R
- Freitag, C. M., Konrad, C., Häberlen, M., Kleser, C., Von Gontard, A., Reith, W., et al. (2008). Perception of biological motion in autism spectrum disorders. *Neuropsychologia* 46, 1480–1494. doi: 10.1016/j.neuropsychologia.2007.12.025
- Frith, C. D., and Frith, U. (2006). How we predict what other people are going to do. *Brain Res.* 1079, 36–46. doi: 10.1016/j.brainres.2005.12.126
- Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford: Basil Blackwell.
- Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of “theory of mind” in verbal and nonverbal tasks. *Neuropsychologia* 38, 11–21. doi: 10.1016/S0028-3932(99)00053-6
- Gianelli, C., Scorolli, C., and Borghi, A. M. (2011). Acting in perspective: the role of body and language as social tools. *Psychol. Res.* 77, 40–52. doi: 10.1007/s00426-011-0401-0
- Gregg, M., Hall, C., and Butler, A. (2010). The MIQ-RS: a suitable option for examining movement imagery ability. *Evid. Based Complement. Alternat. Med.* 7, 249–257. doi: 10.1093/ecam/nem170
- Grèzes, J., Pichon, S., and De Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage* 35, 959–967. doi: 10.1016/j.neuroimage.2006.11.030
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–396. doi: 10.1017/S0140525X04000093
- Hall, C. R., and Martin, K. A. (1997). Measuring movement imagery abilities: a revision of the movement imagery questionnaire. *J. Ment. Imagery* 21, 143–154.

Acknowledgments

Research was supported by grants from the French Research Agency ANR-2009-CORD-014-INTERACT and ANR-11-EQPX-0023. DL and FQ were financed by the region Nord-Pas-de-Calais (France).

- Hamilton, A. F. de C. (2009). Goals, intentions and mental states: challenges for theories of autism. *J. Child Psychol. Psychiatry* 50, 881–892. doi: 10.1111/j.1469-7610.2009.02098.x
- Hecht, H., Vogt, S., and Prinz, W. (2001). Motor learning enhances perceptual judgment: a case for action-perception transfer. *Psychol. Res.* 65, 3–14. doi: 10.1007/s004260000043
- Hillebrandt, H., Friston, K. J., and Blakemore, S. J. (2014). Effective connectivity during animacy perception—dynamic causal modelling of Human Connectome Project data. *Sci. Rep.* 4, 6240. doi: 10.1038/srep06240
- Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878. doi: 10.1017/S0140525X01000103
- Innocenti, A., De Stefani, E., Bernardi, N. F., Campione, G. C., and Gentilucci, M. (2012). Gaze direction and request gesture in social interactions. *PLoS ONE* 7:e36390. doi: 10.1371/journal.pone.0036390
- Jackson, R. C., Warren, S., and Abernethy, B. (2006). Anticipation skill and susceptibility to deceptive movement. *Acta Psychol.* 123, 355–371. doi: 10.1016/j.actpsy.2006.02.002
- Jacob, P., and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends Cogn. Sci.* 9, 21–25. doi: 10.1016/j.tics.2004.11.003
- Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, 103–109. doi: 10.1006/nimg.2001.0832
- Jeannerod, M., and Decety, J. (1995). Mental motor imagery: a window into the representational stages of action. *Curr. Opin. Neurobiol.* 5, 727–732. doi: 10.1016/0959-4388(95)80099-9
- Kahneman, D. (2003). A Perspective on Judgment and Choice. *Am. Psychol.* 58, 697–720. doi: 10.1037/0003-066X.58.9.697
- Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cogn. Process.* 8, 159–166. doi: 10.1007/s10339-007-0170-2
- Kim, J., Doop, M. L., Blake, R., and Park, S. (2005). Impaired visual recognition of biological motion in schizophrenia. *Schizophr. Res.* 77, 299–307. doi: 10.1016/j.schres.2005.04.006
- Kim, J., Park, S., and Blake, R. (2011). Perception of biological motion in schizophrenia and healthy individuals: a behavioral and fMRI study. *PLoS ONE* 6:e19971. doi: 10.1371/journal.pone.0019971
- Knoblich, G., and Flach, R. (2001). Predicting the effects of actions: interactions of perception and action. *Psychol. Sci.* 12, 467–472. doi: 10.1111/1467-9280.00387
- Knoblich, G., and Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philos. Trans. R. Soc. Lond.* 363, 2021–2031. doi: 10.1098/rstb.2008.0006
- Krüger, S., Sokolov, A. N., Enck, P., Krägeloh-Mann, I., and Pavlova, M. A. (2013). Emotion through locomotion: gender impact. *PLoS ONE* 8:e81716. doi: 10.1371/journal.pone.0081716
- Lewkowicz, D., and Delevoeye-Turrell, Y. (2015). Real-Time Motion Capture Toolbox (RTMocap): an open-source code for recording 3-D motion kinematics to study action-effect anticipations during motor and social interactions. *Behav. Res. Methods* doi: 10.3758/s13428-015-0580-5 [Epub ahead of print].
- Lewkowicz, D., Delevoeye-Turrell, Y., Bailly, D., Andry, P., and Gaussier, P. (2013). Reading motor intention through mental imagery. *Adapt. Behav.* 21, 315–327. doi: 10.1177/1059712313501347
- Loison, B., Moussaddaq, A. S., Cormier, J., Richard, I., Ferrapie, A. L., Ramond, A., et al. (2013). Translation and validation of the French Movement Imagery Questionnaire—Revised Second version (MIQ-RS). *Ann. Phys. Rehabil. Med.* 56, 157–173. doi: 10.1016/j.rehab.2013.01.001
- Louis-Dam, A., Orliaguet, J.-P., and Coello, Y. (1999). “Perceptual anticipation in grasping movement: when does it become possible?” in *Studies in Perception and Action*, eds M. G. Grealy and J. A. Thomson (London: Lawrence Erlbaum Associates), 34–39.
- Manera, V., Becchio, C., Cavallo, A., Sartori, L., and Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Exp. Brain Res.* 211, 547–556. doi: 10.1007/s00221-011-2649-4
- Marteniuk, R. G., MacKenzie, C. L., Jeannerod, M., Athenes, S., and Dugas, C. (1987). Constraints on human arm movement trajectories. *Can. J. Psychol.* 41, 365–378. doi: 10.1037/h0084157
- Méry, D., Chary, C., Palluel, R., and Orliaguet, J.-P. (2005). Visual perception of pointing and writing movements. *Perception* 34, 1061–1067. doi: 10.1068/p3388
- Meeren, H. K. M., van Heijnsbergen, C., and de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16518–16523. doi: 10.1073/pnas.0507650102
- Miller, L. E., and Saygin, A. P. (2013). Individual differences in the perception of biological motion: links to social cognition and motor imagery. *Cognition* 128, 140–148. doi: 10.1016/j.cognition.2013.03.013
- Naish, K. R., Reader, A. T., Houston-Price, C., Bremner, A. J., and Holmes, N. P. (2013). To eat or not to eat? Kinematics and muscle activity of reach-to-grasp movements are influenced by the action goal, but observers do not detect these differences. *Exp. Brain Res.* 225, 261–275. doi: 10.1007/s00221-012-3367-2
- Obhi, S. S. (2012). The amazing capacity to read intentions from movement kinematics. *Front. Hum. Neurosci.* 6:117. doi: 10.3389/fnhum.2012.00162
- Orliaguet, J. P., Viallon, S., Coello, Y., and Kandel, S. (1996). Perceptual anticipation in sequential grasping movements. *Int. J. Psychol.* 31, 419.
- Oztop, E., Wolpert, D., and Kawato, M. (2005). Mental state inference using visual control parameters. *Cogn. Brain Res.* 22, 129–151. doi: 10.1016/j.cogbrainres.2004.08.004
- Parron, C., Da Fonseca, D., Santos, A., Moore, D., Monfardini, E., and Deruelle, C. (2008). Recognition of biological motion in children with autistic spectrum disorders. *Autism* 12, 261–274. doi: 10.1177/1362361307089520
- Patel, D., Fleming, S. M., and Kilner, J. M. (2012). Inferring subjective states through the observation of actions. *Proc. R. Soc. Lond. B. Biol. Sci.* 279, 4853–4860. doi: 10.1098/rspb.2012.1847
- Pavlova, M. A. (2012). Biological motion processing as a hallmark of social cognition. *Cereb. Cortex* 22, 981–995. doi: 10.1093/cercor/bhr156
- Pollick, F. E., Paterson, H. M., Bruderlin, A., and Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition* 82, B51–B61. doi: 10.1016/S0010-0277(01)00147-0
- Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a “theory of mind”? *Behav. Brain Sci.* 4, 515–526. doi: 10.1017/S0140525X00076512
- Prevost, M., Carrier, M. E., Chowne, G., Zekowicz, P., Joseph, L., and Gold, I. (2014). The Reading the Mind in the Eyes test: validation of a French version and exploration of cultural variations in a multi-ethnic city. *Cogn. Neuropsychiatry* 19, 189–204. doi: 10.1080/13546805.2013.823859
- Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154. doi: 10.1080/713752551
- Quesque, F., and Coello, Y. (2014). For your eyes only: effect of confederate’s eye level on reach-to-grasp action. *Front. Psychol.* 5:1407. doi: 10.3389/fpsyg.2014.01407
- Quesque, F., Lewkowicz, D., Delevoeye-Turrell, Y. N., and Coello, Y. (2013). Effects of social intention on movement kinematics in cooperative actions. *Front. Neurobot.* 7:14. doi: 10.3389/fnbot.2013.00014
- Runeson, S., and Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. *J. Exp. Psychol. Gen.* 112, 585. doi: 10.1037/0096-3445.112.4.585
- Ruys, K. I., and Aarts, H. (2010). When competition merges people’s behavior: interdependency activates shared action representations. *J. Exp. Soc. Psychol.* 46, 1130–1133. doi: 10.1016/j.jesp.2010.05.016
- Sartori, L., Becchio, C., and Castiello, U. (2011). Cues to intention: the role of movement information. *Cognition* 119, 242–252. doi: 10.1016/j.cognition.2011.01.014
- Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind.” *Neuroimage* 19, 1835–1842. doi: 10.1016/S1053-8119(03)00230-1
- Scorilli, C., Miatton, M., Wheaton, L. A., and Borghi, A. M. (2014). I give you a cup, I get a cup: a kinematic study on social intention. *Neuropsychologia* 57, 196–204. doi: 10.1016/j.neuropsychologia.2014.03.006
- Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Top. Cogn. Sci.* 1, 353–367. doi: 10.1111/j.1756-8765.2009.01024.x
- Sebanz, N., and Shiffrar, M. (2009). Detecting deception in a bluffing body: the role of expertise. *Psychon. Bull. Rev.* 16, 170–175. doi: 10.3758/PBR.16.1.170
- Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662. doi: 10.1016/S0896-6273(04)00014-5
- Sokolov, A. A., Krüger, S., Enck, P., Krägeloh-Mann, I., and Pavlova, M. A. (2011). Gender affects body language reading. *Front. Psychol.* 2:16. doi: 10.3389/fpsyg.2011.00016

- Spaulding, S. (2015). On direct social perception. *Conscious. Cogn.* doi: 10.1016/j.concog.2015.01.003 [Epub ahead of print]
- Stapel, J. C., Hunnius, S., and Bekkering, H. (2012). Online prediction of others' actions: the contribution of the target object, action context and movement kinematics. *Psychol. Res.* 76, 434–445. doi: 10.1007/s00426-012-0423-2
- Troje, N. F., Westhoff, C., and Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. *Percept. Psychophys.* 67, 667–675. doi: 10.3758/BF03193523
- Uithol, S., van Rooij, I., Bekkering, H., and Haselager, P. (2011). Understanding motor resonance. *Soc. Neurosci.* 6, 388–397. doi: 10.1080/17470919.2011.559129
- Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584. doi: 10.1016/j.neuroimage.2009.06.009
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460
- Wolpert, D. M., and Miall, R. C. (1996). Forward models for physiological motor control. *Neural Netw.* 9, 1265–1279. doi: 10.1016/S0893-6080(96)00035-4
- Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 593–602. doi: 10.1098/rstb.2002.1238
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882. doi: 10.1126/science.7569931

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lewkowicz, Quesque, Coello and Delevoeye-Turrell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Factors affecting athletes' motor behavior after the observation of scenes of cooperation and competition in competitive sport: the effect of sport attitude

Elisa De Stefani*, Doriana De Marco and Maurizio Gentilucci

Department of Neuroscience, University of Parma, Parma, Italy

OPEN ACCESS

Edited by:

Andreas B. Eder,
University of Wuerzburg, Germany

Reviewed by:

Yves Paulignan,
Centre National de la Recherche
Scientifique, France
Thorsten Michael Erle,
University of Würzburg, Germany

*Correspondence:

Elisa De Stefani
elidestefani@gmail.com

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 10 March 2015

Accepted: 13 October 2015

Published: 28 October 2015

Citation:

De Stefani E, De Marco D
and Gentilucci M (2015) Factors
affecting athletes' motor behavior
after the observation of scenes
of cooperation and competition
in competitive sport: the effect
of sport attitude.
Front. Psychol. 6:1648.
doi: 10.3389/fpsyg.2015.01648

Aim: This study delineated how observing sports scenes of cooperation or competition modulated an action of interaction, in expert athletes, depending on their specific sport attitude.

Method: In a kinematic study, athletes were divided into two groups depending on their attitude toward teammates (cooperative or competitive). Participants observed sport scenes of cooperation and competition (basketball, soccer, water polo, volleyball, and rugby) and then they reached for, picked up, and placed an object on the hand of a conspecific (giving action). Mixed-design ANOVAs were carried out on the mean values of grasping-reaching parameters.

Results: Data showed that the type of scene observed as well as the athletes' attitude affected reach-to-grasp actions to give. In particular, the cooperative athletes were speeded when they observed scenes of cooperation compared to when they observed scenes of competition.

Discussion: Participants were speeded when executing a giving action after observing actions of cooperation. This occurred only when they had a cooperative attitude. A match between attitude and intended action seems to be a necessary prerequisite for observing an effect of the observed type of scene on the performed action. It is possible that the observation of scenes of competition activated motor strategies which interfered with the strategies adopted by the cooperative participants to execute a cooperative (giving) sequence.

Keywords: scenes of cooperation and competition, expert athletes, cooperative/competitive attitude, kinematics, social interaction

INTRODUCTION

A growing number of behavioral and neurophysiological studies have demonstrated that perception and action have a common coding (Rizzolatti and Craighero, 2004; Rizzolatti et al., 2014). The concept of affordances, as originally postulated by Gibson (1978), refers to the possibilities for action that emerge from the interactions of an organism with its environment.

Further evidence has demonstrated that activation of affordances is modulated not just by the physical properties of objects, but also by the social context in which an action is performed (Mason and Mackenzie, 2005; Georgiou et al., 2007; Meulenbroek et al., 2007; Becchio et al., 2008; Sartori et al., 2009; Ferri et al., 2010, 2011, 2014; Innocenti et al., 2012). Indeed, social behavior during interaction with conspecifics (i.e., different intentions of the agent or the observer) can interact with affordance instantiation and modify the kinematics of the actions. The ability to read others' intentions plays an important role in sports, as athletes need to perceive the action capabilities of their opponents and their teammates in order to be aware of ever-changing opportunities for action afforded by a sport situation (Passos et al., 2009; Correia et al., 2012; Vilar et al., 2012).

Throughout the course of a game, players can implement both defensive and offensive behaviors. It is possible that these behaviors lead to the development of certain skills to either cooperate or compete with teammates. Moreover, with the development of expertise in a sport, athletes improve specific patterns of interaction, that is, a personal predisposition to be more cooperative or competitive toward their teammates. We refer to these specific strategies using the term "attitude": a predisposition toward a specific motor behavior in response to an actual sport setting.

It is well known that the observation of an action activates a process of simulation (Buccino et al., 2004b). For transitive actions (directed upon an object), the same act done by another agent corresponds to the activation of an internal motor representation of that act. This simulation is used to understand the goal of the movement (Buccino et al., 2001, 2004a; Iacoboni et al., 2005). In the case of intransitive actions, the simulation is mainly used to understand the intention of the agent (Fadiga et al., 1995; Buccino et al., 2001; Rizzolatti and Craighero, 2004). In summary, the simulation of an observed action allows one to recognize the goal of the observed movement, to infer others' intentions, and to predict the agent's next act. Moreover, this mechanism of intention understanding can modulate a further self-generated action. In other words, the observation of an action can influence the motor response of a subsequent action. This happens often in a sport context: actions are frequently executed in the presence of another acting individual whose intentions can be cooperative or competitive. Consequently, the observation of sport scenes of cooperation and competition can differently affect the subsequent action of the observer. We hypothesized that this effect would enhance the cooperative and competitive attitude of an athlete. Athletes that are attuned to simulating sportive actions can be greatly affected, compared to non-athletes, in the execution of a subsequent action after observing sportive scenes of cooperation and competition.

We extended our research to sport expertise by considering athletes' attitudes (cooperative versus competitive). Two main issues were examined in this study: firstly, we were interested in ascertaining whether the sole observation of well-known sport actions in a context of cooperation or competition could influence the kinematics of a cooperative social interaction with a conspecific (giving action). Specifically, we expected that the observation of an action of cooperation could

facilitate a successive executed action of cooperation, making the participant's movement faster. On the other hand, the observation of an action of competition could interfere with the participant's action of cooperation, probably slowing down the movement. Secondly, we were interested in investigating how the kinematics of athletes' actions can be modulated not only by the observation of a specific cooperative/competitive sport action, but also by the attitude of the participants. We hypothesized that the interaction between the participant's attitude (cooperative or competitive) and the type of sport actions observed (an action of cooperation or an action of competition) could modulate a successive motor response, affecting the kinematics of reach-grasp movements performed by participants. Specifically, we expected that the congruence between the participant's attitude (e.g., cooperative attitude) and an observed action (e.g., action of cooperation) could facilitate the execution of a successive movement toward a conspecific, making the participant's action faster. On the other hand, we expected that the incongruence matching (e.g., cooperative attitude versus the observation of an action of competition) could interfere with a successive interaction with a conspecific, presumably slowing down the movement. In other words, we expected facilitation only when the attitude of the participant was congruent with the type of observed action.

MATERIALS AND METHODS

Participants

Twenty right-handed undergraduate students (9 male, 11 female) between the ages of 20 and 28 years (mean = 21.6, $SD = 2.5$) took part in the present experiment. They all practiced a sport more than three times per week ($SD = 1.7$) and they all had experience in one or more of the team sports selected in this study (Table 1). Handedness was assessed through the Edinburgh Inventory (Oldfield, 1971). The participants were students of the degree course of Motor Sciences, Sport and Health (University of Parma) and practiced team sports at the competitive level. Before being included in our study, the participants completed a questionnaire to collect information about what sport they practiced; which position they played; and whether they felt more cooperative with their peers than competitive toward their opponents during a game (see Data Sheet 1). The participants were divided into two groups (cooperative and competitive group) according to their answers. In the competitive group, we included only participants that had clearly exhibited competitive behavior during matches (13 competitive athletes). We used the same criteria for athletes included in the cooperative group (seven cooperative athletes). We excluded the uncertain participants. All participants provided a written informed consent to participate in the study, which has been approved by the local ethical committee (Comitato Etico per Parma) and has been conducted according to the principles expressed in the Declaration of Helsinki.

Apparatus, Stimuli, and Procedure

The participants sat comfortably in front of a table on which they placed their right hand with the thumb and index finger

TABLE 1 | Participants' characteristics.

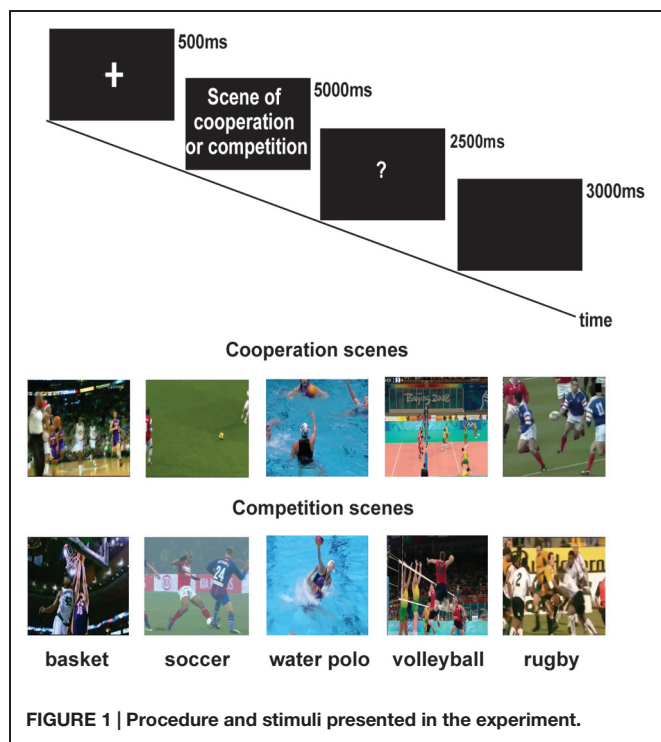
Participants	Age	Attitude	Sport	Frequency	Sex	Expertise
1	20	cooperative	basket	>3 days, a week	M	more than 1 years
2	20	competitive	volleyball	>4 days a week	F	more than 1 years
3	20	competitive	volleyball	>4 days, a week	F	more than 1 years
4	28	competitive	water polo	>4 days a week	F	more than 1 years
5	20	competitive	water polo	>4 days a week	F	more than 1 years
6	20	cooperative	soccer	>3 days a week	M	more than 1 years
7	21	competitive	volleyball	>3 days a week	F	more than 1 years
8	20	competitive	volleyball	>4 days a week	F	more than 1 years
9	20	cooperative	volleyball	>4 days, a week	F	more than 1 years
10	21	competitive	soccer	>4 days a week	M	more than 1 years
11	20	competitive	soccer	>4 days, a week	M	more than 1 years
12	21	cooperative	rugby	>4 days a week	M	more than 1 years
13	21	cooperative	volleyball	>3 days a week	F	more than 1 years
14	20	competitive	volleyball	>4 days a week	F	more than 1 years
15	20	competitive	basket	>3 days a week	F	more than 1 years
16	21	cooperative	soccer	>4 days a week	M	more than 1 years
17	21	competitive	soccer	>4 days, a week	M	more than 1 years
18	26	cooperative	soccer	>4 days a week	M	more than 1 years
19	26	competitive	basket	>4 days, a week	F	more than 1 years
20	25	competitive	soccer	>4 days a week	M	more than 1 years

in pinch position starting position (SP). SP was located along the participants' mid-sagittal plane and was 27 cm away from their chest. An experimenter was seated next to the participant, and she held the palm of her right hand in the supine position (request position). A computer display was placed on a table plane at a distance of 60 cm from the body of the participant sitting in front of it. A wooden cube ($\sim 2 \text{ cm} \times 2 \text{ cm} \times 2 \text{ cm}$) was placed at the center of the table 20 cm in front of participant's SP. Stimuli were presented on the computer display using software developed via MATLAB version 7.7 (R2008b). The stimuli were short videos downloaded from the Internet replicating real matches. Each video lasted five seconds. We selected videos based on the following criterion: (a) the action would involve coordinated sports action among athletes of the same team, or (b) two or more athletes from two different teams would come into contact with each other. Consequently, the actions defined "actions of cooperation"-reproduced situations in which athletes of the same team cooperated in an action of the game (e.g., in volleyball, a pass ball between setter and hitter, see **Figure 1**). In the "actions of competition"-reproduced situations, two athletes from two different teams were opposed (e.g., in a soccer match, the attacker tries to score a goal and the defender marks him). Selected scenes reproduced sports actions in which the participants were experts—that is, five cooperation and five competition scenes from the following sports: basketball, soccer, water polo, volleyball, and rugby (**Figure 1**). In total, 50 scenes were presented. After the presentation of a fixation cross (500 ms), participants viewed one of the 10 videos that lasted 5,000 ms. As soon as they understood whether the action was one of cooperation or competition, they were required to reach for, pick up, and place the wooden cube on the experimenter's hand (giving action). The participants grasped the cube with

their fingers (right hand, precision grip). When a question mark (2,500 ms) appeared on the computer display, the participants were instructed to state out loud whether the just seen action was an action of cooperation or competition (10% catch trials). Subsequently, a black screen was presented (3,000 ms). The participants had to place their hands in SP and then wait for the next trial. In total, the participants responded correctly to the cooperation condition in 99% of the cases and in the competition condition in 99.7% of the cases.

Data Recording

The movements of the participants' right arms were recorded using the 3D-optoelectronic SMART system (BTS Bioengineering, Milan, Italy). This system consists of six video cameras that detect infrared reflecting markers (spheres that are 5 mm in diameter) at a sampling rate of 120 Hz. The spatial resolution of the system is 0.3 mm. The infrared reflective markers were attached to the nail of the participants' right thumbs and index fingers, and another marker was attached to the participants' right wrists. The markers attached to the thumb and index finger were used to analyze the grasp kinematics, whereas the marker attached to the wrist was used to analyze the kinematics of reaching and lifting. Manual prehension consists of two components: the proximal component (also known as "the reach"), which is the action of carrying the hand toward an object, and "the grasp" component, during which the fingers are opened and shaped before the contact of the hand with the target (Jeannerod, 1984; Jakobson and Goodale, 1991; Gentilucci et al., 2001). The reach transports the hand toward the object (the reaching action makes the hand move toward an object), and its kinematics depend on the target's extrinsic properties (i.e., location and orientation). The grasp component provides



information on how to open, preshape, and close the hand during the reach in relation to the target's intrinsic properties (i.e., size and shape). The data of the recorded movements was analyzed using software developed via MATLAB version 7.7 (R2008b). Recorded data were filtered using a Gaussian low-pass smoothing filter ($\Sigma = 0.93$). The time course of the reach, grasp, and lift was visually inspected: the beginning of the grasp was considered to be the first frame in which the distance between the two markers placed on the right finger tips increased more than 0.3 mm (spatial resolution of the recording system) with respect to the previous frame. The end of the grasp was the first frame after the beginning of the finger closing, in which the distance between the two right fingers decreased less than 0.3 mm with respect to the previous frame. The beginning of the reach was considered to be the first frame during which the displacement of

the reach marker along any Cartesian body axis increased more than 0.3 mm with respect to the previous frame. To determine the end of the reach, we calculated the first frame following movement onset separately for the X, Y, and Z axes, in which the X, Y, and Z displacements of the reach marker decreased less than 0.3 mm compared to the previous frame. Then, the frame endpoint temporally closer to the grasp end frame was chosen as the end of the reach. The frame immediately succeeding the reach end was considered as the lift beginning, while the lift end corresponded to the frame in which the highest point of the hand trajectory was reached during lifting. The grasp was studied by analyzing the time course of the distance between the index finger and thumb markers. From a pinch position, the grasp component was constituted of an initial phase of finger opening up to a maximum (maximal finger aperture) followed by a phase of finger closing on the object (Jeannerod, 1988).

We measured the following parameters: reach time, time to peak velocity of reach, peak elevation (trajectory maximal height), grasp time, time to maximal finger aperture, peak velocity of finger opening, time to peak velocity of finger opening, and maximal finger aperture.

Data Analysis

Participants were divided into two groups (cooperative attitude versus competitive attitude) according to the questionnaire responses. They resulted in 7 cooperative participants and 13 competitive participants (**Table 1**). Because of the difference in sample size between groups, the homogeneity of variance was primarily verified with Levene's test. Mixed-design ANOVAs were carried out on the mean values of the reaching-grasping parameters (**Table 2**). The within-subject factor was the type of scene (cooperation versus competition) and the between-subject factor was the participants' attitudes (cooperative versus competitive). In all of the analyses, *post hoc* comparisons were performed using the Newman-Keuls procedure. The significance level was fixed at $p = 0.05$. When a factor was significant, we also calculated the effect size (η_p^2). We also carried another mixed-design ANOVA, using gender (male versus female) and type of practiced sport (basketball versus soccer versus water polo versus volleyball versus rugby) as the between-subject factors. All of

TABLE 2 | Mean values and SE of kinematic parameters of reach and grasp action.

Kinematic parameters	Scene of cooperation				Scene of competition			
	Cooperative attitude		Competitive attitude		Cooperative attitude		Competitive attitude	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Reach time (ms)	637	47	513	35	686	50	523	37
Time to peak velocity of reach (ms)	287	21	256	15	313	22	258	16
Peak elevation (mm)	94	5	91	4	96	5	93	4
Grasp time (ms)	607	44	512	32	645	44	504	32
Time to maximal finger aperture (ms)	420	36	316	27	455	37	311	27
Peak velocity of finger opening (mm/s)	233	43	305	31	217	14	309	30
Time to peak velocity of finger opening (ms)	208	28	160	21	238	31	152	23
Maximal finger aperture (mm)	78	3	84	2	77	3	83	2

these final analyses were not significant, and the corresponding p -values are reported as Supplementary Table S1.

RESULTS

Reach

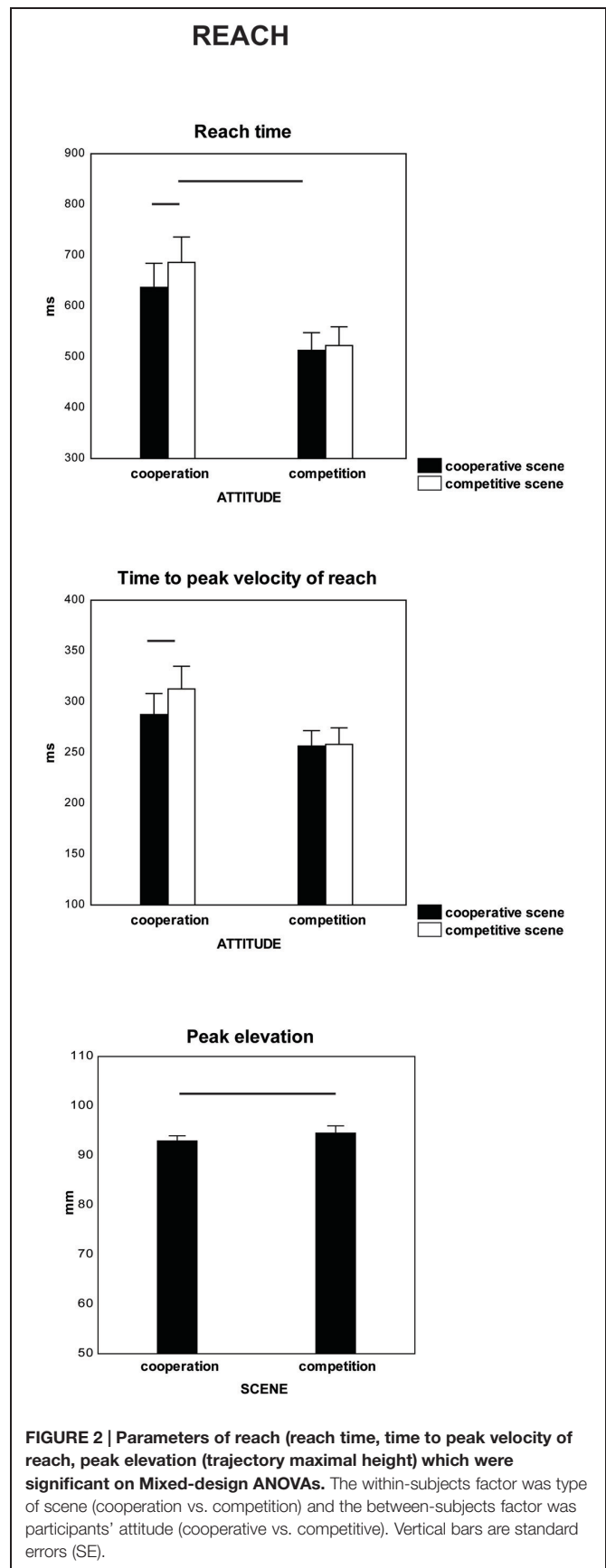
The main factor of the participants' attitudes was significant. There was a significant difference in reach time between cooperative participants and competitive participants [$F(1,18) = 5.74$, $p < 0.028$; cooperative = 662 ms versus competitive = 518 ms].

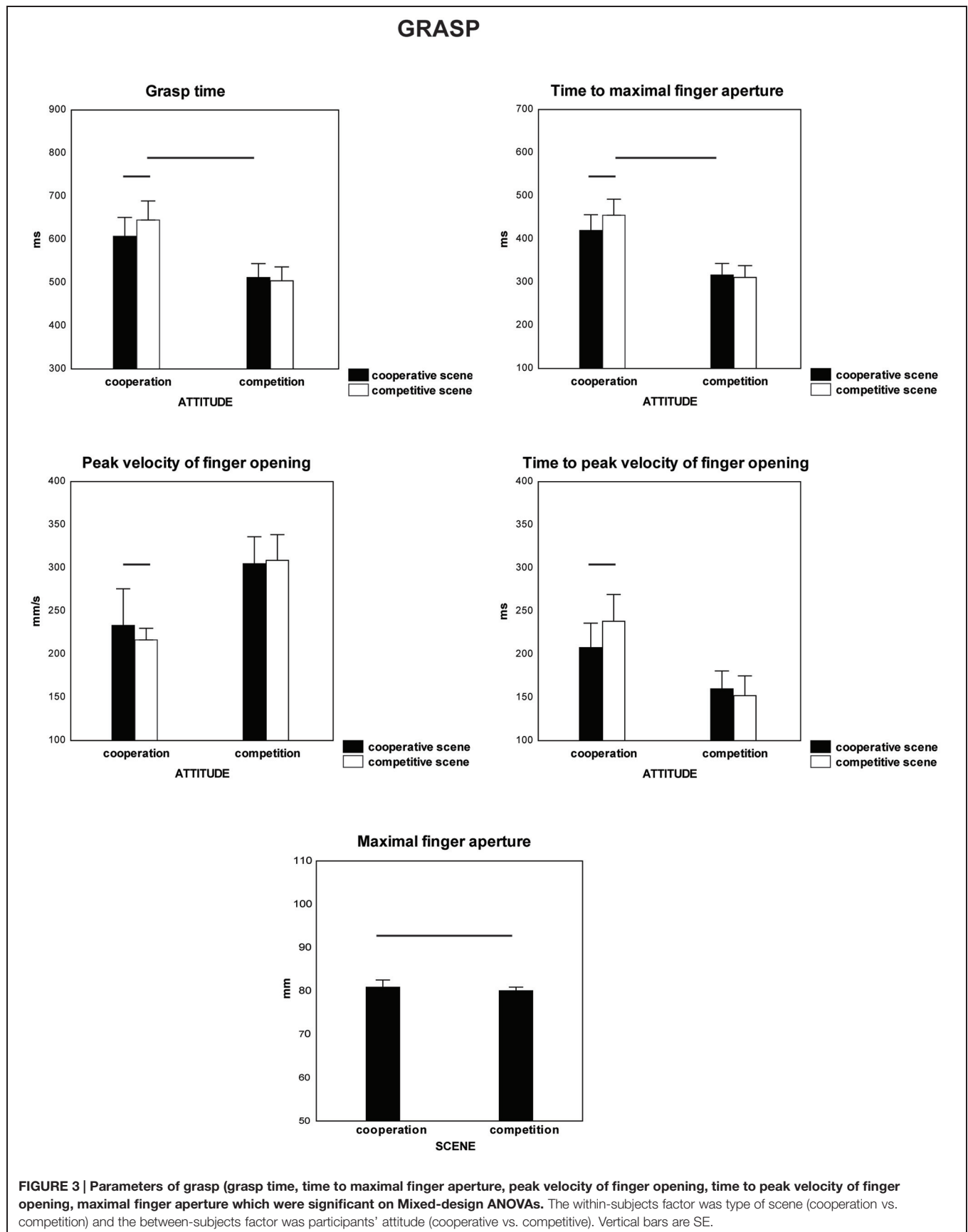
Factor scene affected reach time and time to peak velocity of reach. Scenes of cooperation induced a decrease in both parameters in comparison with scenes of competition [reach time: $F(1,18) = 15$, $\eta_p^2 = 0.45$, $p < 0.00$, 575 ms versus 604 ms; time to peak velocity of reach: $F(1,18) = 6.5$, $\eta_p^2 = 0.27$, $p < 0.02$, 271 ms versus 285 ms]. It is possible that the scenes of cooperation facilitated, and/or the scenes of competition interfered with, the reach (and grasp, see below) because the participants executed a giving (cooperative) action. The interaction between the type of scene and the participants' attitudes also affected reach time [$F(1,18) = 6.8$, $\eta_p^2 = 0.274$, $p < 0.018$] and time to peak velocity of reach [$F(1,18) = 5.01$, $\eta_p^2 = 0.218$, $p < 0.038$, **Figure 2** and **Table 2**]. *Post hoc* comparison showed a significance between types of scene only when the participants were cooperative (reach time: $p = 0.00037$; time to peak velocity of reach: $p = 0.003$). No difference was found between scenes of cooperation and competition when participants were competitive (reach time: $p = 0.384$; time to peak velocity of reach: $p = 0.827$). Finally, scenes of cooperation and competition affected peak elevation differentially [$F(1,18) = 4.7$, $\eta_p^2 = 0.208$, $p < 0.043$, 93 mm versus 95 mm].

Grasp

Competitive participants showed a significant decrease in grasp time and time to maximal finger aperture compared to cooperative participants (grasp time: $F(1,18) = 4.8$, $p < 0.042$, 508 ms versus 626 ms; time to maximal finger aperture: $F(1,18) = 7.5$, $p < 0.013$, 314 ms versus 437 ms).

A significant interaction between the factor type of the scene and the participants' attitudes was found for grasp time [$F(1,18) = 7.24$, $\eta_p^2 = 0.287$, $p < 0.015$] and time to maximal finger aperture [$F(1,18) = 6.35$, $\eta_p^2 = 0.261$, $p < 0.021$, **Table 2** and **Figure 3**]. *Post hoc* comparison showed a significant decrease in the parameters for scenes of cooperation only when the participants were cooperative (grasp time: $p = 0.005$; time to maximal finger aperture: $p = 0.006$). No difference was found between the scenes of cooperation and competition presented to competitive participants (grasp time: $p = 0.533$; time to maximal finger aperture: $p = 0.639$). The interaction between the type of scene and the participants' attitudes showed a trend toward significance for peak velocity of finger opening [$F(1,18) = 3.88$, $\eta_p^2 = 0.177$, $p < 0.064$] and significance for time to peak velocity of finger opening [$F(1,18) = 8.69$, $\eta_p^2 = 0.325$, $p < 0.009$]. *Post hoc* comparisons showed a significant decrease in the two parameters





in the presence of scenes of cooperation only when they were presented to cooperative participants (peak velocity of finger opening: $p = 0.037$; time to peak velocity of finger opening: $p = 0.0039$). Scenes of cooperation and competition differentially affected maximal finger aperture. Participants opened their fingers to a larger degree when grasping the target after seeing scenes of cooperation compared to competition [$F(1,18) = 5.2$, $\eta_p^2 = 0.225$, $p < 0.035$; 81 mm versus 80 mm].

In sum, the participants were facilitated (i.e., faster) when executing actions of cooperation after observing actions of cooperation. This occurred only when they had cooperative attitudes. In general, the competitive participants were faster than the cooperative ones.

DISCUSSION

The aim of the present study was to determine whether and how the matching between the athletes' attitudes (cooperative and competitive attitude) and the observation of sport scenes (actions of cooperation and competition) could influence the kinematics of a successive social interaction. The participants were all expert athletes in at least one of the team sports selected for this study (basketball, soccer, water polo, volleyball, and rugby; **Figure 1**). Before starting the experiment, the athletes were divided into two groups according to their attitude during a game (cooperative versus competitive attitude; see Materials and Methods). The participants had to observe a sport scene of cooperation or competition before performing a motor sequence. They executed a reach-grasp of an object and placed it in the hand of an experimenter who was sitting close to them (a cooperative giving action). Our expectation was that both the participants' attitudes and the type of scene would influence the sequence kinematics.

Firstly, we observed an effect of attitude. The competitive participants were faster than the cooperative ones during the action execution regardless of the observed scene. A possible explanation for this finding is that competitive athletes are generally faster in performing an action than cooperative athletes are. Alternatively, the cooperative athletes could be less competitive, and for this reason, they are slower in performing an action with respect to competitive athletes. A further possible explanation is that the lack of any effect when the scenes of cooperation and competition were presented to the competitive athletes might depend on the inability of these athletes to adopt strategies that are suitable to successfully execute the giving sequence toward a conspecific.

Secondly, we observed an interaction effect between the athletes' attitudes and the type of scene on the reach-grasp temporal parameters. The cooperative participants were faster in their movement when they observed scenes of cooperation, subsequently executing the giving action. On the contrary, these athletes were slower when they observed scenes of competition.

It is possible that the observed action could have been automatically mapped onto participants' motor system, resulting in a facilitation of functionally similar actions. In other words, the observed scene probably acted as a prime stimulus for the subsequent executed action. This facilitation effect would

have been present when the participants observed a scene of cooperation and then had to perform a cooperative motor sequence toward a conspecific. On the other hand, there would have been an interference effect when the participants observed a scene of competition and had to perform a cooperative motor sequence (Chartrand and Bargh, 1999; Brass et al., 2000, 2001; Flanagan and Johansson, 2003; Kilner et al., 2003; Sebanz et al., 2003, 2006; Newman-Norlund et al., 2007; Liepelt et al., 2008; Bekkering et al., 2009). However, the competitive participants did not show any effect. The fact that only the cooperative participants were affected by the type of scene they observed suggests that the effect was more complex than a simple priming.

Only when there was congruence between the attitude and the observed action was it possible to observe changes in the kinematics of a giving action. Specifically, in the case of congruence (i.e., cooperative attitude and observation of a scene of cooperation), the kinematics of the cooperative participants sped up, whereas in the case of incongruence, they slowed down. On the contrary, the competitive athletes seemed to not be directly affected by the experimental conditions. A possible explanation of this result is that they were already faster and, for this reason, the difference between actions of cooperation and competition did not emerge. What would happen if the competitive athletes had to perform a competitive action (e.g., grasp the target and move it away from the conspecific)? Might we expect that the competitive athletes would be faster if they have just observed a scene of competition and slowed down in the case of cooperation? We cannot exclude this possibility. However, we suppose that an action of competition would be performed quickly in order to take away the object as quickly as possible (Georgiou et al., 2007). Consequently, it is possible that the speed of this action may prevent us from observing any effect. However, we believe that deepening these aspects could have interesting implications. For this reason, in future experiments, it would be useful to include a control action, for example, asking the participant to move an object away from the conspecific in order to measure how observing scenes of cooperation and competition affects a competitive action.

Deepening and extending the present results with future studies could have interesting implications for training athletes through the observation of specific sport scenes. For an example, it is possible to speculate that competitive athletes, who were found to be faster in their responses, could be trained to be even faster in their movements through the vision of competitive sport actions.

Finally, we are aware of some limitations in this study. First, we chose to measure the participants' attitudes using a dichotomous item instead of a continuous variable. The reason for our choice was that we wanted to compare the effects of the cooperative and competitive attitude to the videos that were dichotomous (scenes of cooperation and competition). To solve this problem, we included only the athletes who clearly expressed a well-defined position with respect to their attitude, excluding those who were uncertain. Future studies might include sport scenes classified with various degrees of cooperativeness and competitiveness. In this way, it would be possible to compare the participants' attitudes to the

observed scenes in a continuous dimension. Another severe limitation in this study is the very small sample used and the different numbers of males and females and of cooperative and competitive participants (see **Table 1**). For this reason, these findings cannot be generalized to the broader community based on this study alone. In future studies, a larger sample should be used to successfully replicate the present results.

Another important limitation of this study is that we did not use a control group. Future studies might include, for example, a non-athlete group. However, athletes have become more attuned to cooperative and competitive sport situations than non-athletes throughout the course of their sports training. A non-athlete participant group does not have this expertise, so it could be difficult to control the reason why they defined themselves as cooperative or competitive. Another possibility could be to use athletes that play an individual sport, such as dancing or skiing, as a control group. Nevertheless, attention should be paid to their inclusion in the group of cooperative or competitive participants. Finally, another limitation of this study is the lack of a baseline condition against which we could have compared the participants' kinematics after watching the cooperative and

competitive scenes. This aspect is very important, as by including a baseline condition, we could have verified whether watching the different scenes facilitated or interfered with the cooperative participants. Future studies should include a neutral observed scene, for example, a sportive action with just one athlete (e.g., just one soccer player dribbling the ball) as a baseline.

ACKNOWLEDGMENTS

We thank all students of the degree course of Motor Sciences, Sport and Health (University of Parma) who participated in our study. We thank Prof. Francesca Rodighiero for helpful comments on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01648>

REFERENCES

- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition* 106, 894–912. doi: 10.1016/j.cognition.2007.05.004
- Bekkering, H., de Bruijn, E. R. A., Cuijpers, R. H., Newman-Norlund, R., Van Schie, H. T., and Meulenbroek, R. (2009). Joint action: neurocognitive mechanisms supporting human interaction. *Top. Cogn. Sci.* 1, 340–352. doi: 10.1111/j.1756-8765.2009.01023.x
- Brass, M., Bekkering, H., and Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychol. (Amst.)* 106, 3–22. doi: 10.1016/S0001-6918(00)00024-X
- Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn.* 44, 124–143. doi: 10.1006/brcg.2000.1225
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404. doi: 10.1046/j.1460-9568.2001.01385.x
- Buccino, G., Binkofski, F., and Riggio, L. (2004a). The mirror neuron system and action recognition. *Brain Lang.* 89, 370–376. doi: 10.1016/S0093-934X(03)00356-0
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004b). Neural circuits involved in the recognition of actions performed by nonconspecifics: an fMRI study. *J. Cogn. Neurosci.* 16, 114–126. doi: 10.1162/089892904322755601
- Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: the perception-behavior link and social interaction. *J. Pers. Soc. Psychol.* 76, 893–910. doi: 10.1037/0022-3514.76.6.893
- Correia, V., Araújo, D., Duarte, R., Travassos, B., Passos, P., and Davids, K. (2012). Changes in practice task constraints shape decision-making behaviours of team games players. *J. Sci. Med. Sport* 15, 244–249. doi: 10.1016/j.jsams.2011.10.004
- Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* 73, 2608–2611.
- Ferri, F., Busiello, M., Campione, G. C., De Stefani, E., Innocenti, A., Romani, G. L., et al. (2014). The eye contact effect in request and emblematic hand gestures. *Eur. J. Neurosci.* 39, 841–851. doi: 10.1111/ejn.12428
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2010). To me or to you? When the self is advantaged. *Exp. Brain Res.* 203, 637–646. doi: 10.1007/s00221-010-2271-x
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2011). Social requests and social affordances: how they affect the kinematics of motor sequences during interactions between conspecifics. *PLoS ONE* 6:e15855. doi: 10.1371/journal.pone.0015855
- Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771. doi: 10.1038/nature01861
- Gentilucci, M., Benuzzi, F., Gangitano, M., and Grimaldi, S. (2001). Grasp with hand and mouth: a kinematic study on healthy subjects. *J. Neurophysiol.* 86, 1685–1699.
- Georgiou, I., Becchio, C., Glover, S., and Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition*. 102, 415–433. doi: 10.1016/j.cognition.2006.01.008
- Gibson, J. J. (1978). The ecological approach to the visual perception of pictures. *Leonardo* 11, 227–235.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3:e79. doi: 10.1371/journal.pbio.0030079
- Innocenti, A., De Stefani, E., Bernardi, N. F., Campione, G. C., and Gentilucci, M. (2012). Gaze Direction and Request Gesture in Social Interactions. *PLoS ONE* 7:e36390. doi: 10.1371/journal.pone.0036390
- Jakobson, J. S., and Goodale, M. A. (1991). Factors affecting higher-order movement planning: a kinematic analysis of human prehension. *Exp. Brain Res.* 86, 199–208. doi: 10.1007/BF00231054
- Jeannerod, M. (1984). The timing of natural prehension movements. *J. Mot. Behav.* 16, 235–254. doi: 10.1080/00222895.1984.10735333
- Jeannerod, M. (1988). *The Neural and Behavioural Organization of Goal-Directed Movements*. New York, NY, US: Clarendon Press.
- Kilner, J. M., Paulignan, Y., and Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Curr. Biol.* 13, 522–525. doi: 10.1016/S0960-9822(03)00165-9
- Liepelt, R., Von Cramon, D. Y., and Brass, M. (2008). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *Neuroimage*. 43, 784–792. doi: 10.1016/j.neuroimage.2008.08.007
- Mason, A. H., and Mackenzie, C. L. (2005). Grip forces when passing an object to a partner. *Exp. Brain Res.* 163, 173–187. doi: 10.1007/s00221-004-2157-x

- Meulenbroek, R. G. J., Bosga, J., Hulstijn, M., and Miedl, S. (2007). Joint-action coordination in transferring objects. *Exp. Brain Res.* 180, 333–343. doi: 10.1007/s00221-007-0861-z
- Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. G. J., and Bekkering, H. (2007). Exploring the brain basis of joint action: co-ordination of actions, goals and intentions. *Soc. Neurosci.* 2, 48–65. doi: 10.1080/17470910701224623
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Passos, P., Araújo, D., Davids, K., Gouveia, L., Serpa, S., Milho, J., et al. (2009). Interpersonal pattern dynamics and adaptive behavior in multiagent neurobiological systems: conceptual model and data. *J. Mot. Behav.* 41, 445–459. doi: 10.3200/35-08-061
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., and Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706. doi: 10.1152/physrev.00009.2013
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Sartori, L., Becchio, C., Bulgheroni, M., and Castiello, U. (2009). Modulation of the action control system by social intention: unexpected social requests override preplanned action. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1490–1500. doi: 10.1037/a0015777
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76. doi: 10.1016/j.tics.2005.12.009
- Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition* 88, B11–B21. doi: 10.1016/S0010-0277(03)00043-X
- Vilar, L., Araújo, D., Davids, K., and Travassos, B. J. (2012). Constraints on competitive performance of attacker-defender dyads in team sports. *Sports Sci.* 30, 459–469. doi: 10.1080/02640414.2011.627942

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 De Stefani, De Marco and Gentilucci. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Grasping the Agent's Perspective: A Kinematics Investigation of Linguistic Perspective in Italian and German

Claudia Gianelli^{1*}, Michele Marzocchi¹ and Anna M. Borghi^{2,3†}

¹ Division of Cognitive Sciences, University of Potsdam, Potsdam, Germany, ² Department of Psychology, University of Bologna, Bologna, Italy, ³ Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy

OPEN ACCESS

Edited by:

Andriy Myachykov,
Northumbria University, UK

Reviewed by:

Zdenko Kohut,
University of York, UK
Jennifer M. Roche,
Kent State University, USA

*Correspondence:

Claudia Gianelli
claudia.gianelli@uni-potsdam.de

† Present address:

Anna M. Borghi,
Department of Dynamic and Clinical
Psychology, Sapienza University
of Rome, Via degli Apuli 1, Roma,
Italy

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 22 May 2015

Accepted: 09 January 2017

Published: 07 February 2017

Citation:

Gianelli C, Marzocchi M and
Borghi AM (2017) Grasping
the Agent's Perspective:
A Kinematics Investigation
of Linguistic Perspective in Italian
and German. *Front. Psychol.* 8:42.
doi: 10.3389/fpsyg.2017.00042

Every day, we primarily experience actions as agents, by having a concrete perspective on our actions, their means and goals. This peculiar perspective is what allows us to successfully plan and execute our actions in a dense social environment. Nevertheless, in this environment actions are also perceived from an observer's perspective. Adopting such a perspective helps us to understand and respond to other's people actions and their outcomes. Importantly, similar experiences of being agent and observer occur also when actions are not physically acted/perceived but are merely linguistically shared. In this paper we present two exploratory studies, one in Italian and one in German, in which we applied a direct comparison of three singular perspectives in combination with different verb categories. First, second and third person pronouns were combined with action and interaction verbs, i.e., verbs implying an interaction with an object – e.g., grasp – or an interaction with an object and another person – e.g., give. By means of kinematics recording, we analyzed participants' reaching-grasping responses to a mouse while they were presented with the different combinations of linguistic stimuli (pronouns and verb type). Results of Experiment 1 on reaching show that, when they are preceded by YOU, interaction verbs reached the velocity peak earlier than action verbs, since a further motor act will follow. Thus pronouns influence perspective taking and while comprehending language we are sensitive to the motor chain organization of verbs. The absence of the same effects in Experiment 2 is likely due to the fact that, being the pronoun in German mandatory, it is perceived as less salient than in Italian. Overall our result supports the idea that language is grounded in the motor system in a flexible way, and highlights the need for cross-linguistic studies in the field of embodied language processing.

Keywords: perspective taking, action, language comprehension, motor chains, motor system, motor resonance, pronouns, action verbs

INTRODUCTION

Increasing evidence supports the notion that motor processes take place in our brains while we are either observing actions being made by others, or just hearing the verbal description of those actions. In particular, a large amount of data has recently shown that an activation of the motor system (a *motor resonance* process) is present very early during language comprehension,

as revealed by physiological, neuro-imaging, and behavioral studies (for reviews see Fischer and Zwaan, 2008; Toni et al., 2008; Jirak et al., 2010; Glenberg and Gallese, 2012; Meteyard et al., 2012), as well as by computational work (for a review, see Borghi and Cangelosi, 2014). Different methodologies have contributed to characterize this motor activation, supporting the idea that it is not just a side effect of motor imagery but a constitutive part of language comprehension (for a review of the debate on this issue see Fischer and Zwaan, 2008; Mahon and Caramazza, 2008; Toni et al., 2008; Barsalou, 2016). Within behavioral studies, a special role has been played by studies in which kinematics of response movements were recorded (e.g., Boulenger et al., 2006; Dalla Volta et al., 2009; Scorolli et al., 2009; Borghi et al., 2010; Gianelli et al., 2013). By combining the presentation of simple action-related linguistic stimuli (e.g., single verbs, word pairs, short sentences) with a motor task (e.g., reaching-grasping or lifting objects) these studies showed early effects of language processing on motor planning and execution, implying the activation of sensorimotor representations corresponding to the semantic content of linguistic stimuli. However, and despite increasing evidence, the debate regarding the exact timing and nature of the activation of these language-induced sensorimotor representations is still open (see Papeo and Caramazza, 2014).

One open issue regards whether and how these representations are (1) flexible and (2) detailed in terms of the motor components they activate. The present study aims at providing some exploratory data regarding both flexibility of perspective and level of detail of the action components that different perspectives can generate with the help of the powerful tool of kinematics analysis. To this aim, we created a set of stimuli composed of sentence fragments combining a typical perspective-related device, i.e., pronouns, and two categories of action verbs.

Despite being a crucial process in our social interactions (Gallese, 2006; Keysers and Gazzola, 2006), perspective has so far not been a major target of embodied research in the language domain. In contrast, studies on action observation have underlined the importance of perspective, in particular of first-person perspective (Vogt et al., 2003; Jackson et al., 2006; Schütz-Bosbach et al., 2006; Bruzzo et al., 2008; Gianelli et al., 2008).

As pointed out by a recent review (Beveridge and Pickering, 2013), most embodied language studies seem to implicitly assume that readers and listeners activate a first-person – that is an agent – perspective during language comprehension. Likely, this is due to the fact that evidence on motor activation is mainly collected using isolated verbs, rather than verbs embedded within sentences and discourses. But since this assumption is not explicit, the possibility that other perspectives might be activated has not been thoroughly investigated and existing evidence is unclear. Probably also due to the implicit focus on the agent's perspective, pronouns have not been extensively investigated in the recent studies focusing on the motor grounding of language, with the exception of some linguistic studies (MacWhinney, 2005). Pronouns, however, are important as they have at least a double role: first, they allow us to understand who is performing an action (the agent); second, they give us information regarding

the involvement of someone/something else (e.g., patient object) in the action. To our knowledge only a few studies contrasted the motor effects of first and third person action verbs, but they obtained contrasting results. For example, Tomasino et al. (2007) found no difference between first and third person German action verbs with an fMRI study, while in a more recent TMS study Papeo et al. (2011) found a modulation of motor-evoked potentials during processing of first-person but not of third-person Italian action verbs. Furthermore, these results were difficult to compare due to the different languages (German, Italian), to the different techniques (fMRI, TMS) and to the different task. Other TMS studies, such as the study by Buccino et al. (2005) and a subsequent controlled replication of it (Gianelli and Dalla Volta, 2015) found that stimulation of the hand motor cortex while listening to third-person Italian action verbs do indeed modulate motor-evoked potentials as compared to verbs involving other effectors and abstract verbs. However, the latter studies used complete sentences with third-person pronouns instead of infinitive verbs (or sentence fragments) but without direct perspective manipulation. In addition, they used a passive listening task of sentences with only implicit pronoun and the stimulation of the motor cortex occurred when perspective was not yet fully elicited (e.g., in sentences like “cuciva la gonna/she sewed the skirt”). For this reason, although providing evidence in conflict with Papeo et al. (2011), these studies cannot provide final conclusions on this issue.

A small number of behavioral studies have addressed a similar topic but with tasks that did not directly involve the motor system. Brunyé et al. (2009) used a picture-verification task to investigate the perspective adopted when reading action sentences. They compared perspectives implied by the different pronouns (I, You, He) and showed that participants automatically activate an internal perspective when directly addressed as agents (You), whilst activating an observer perspective in the case of He and I pronouns. Interestingly, the same results were obtained also when the task did not explicitly involve a mental simulation, for instance with a memory task (Ditman et al., 2010). However, in this study the linguistic perspective was directly matched with a visual perspective and the authors did not use a motor task.

Differently from passive TMS and fMRI studies, and from behavioral ones, investigations using an explicit motor task might be more informative, as they clearly pose a strong focus on the agent's perspective by requesting participants to perform simple movements as response. However, evidence under this respect is still very limited. For instance, Gianelli et al. (2011) used a novel version of the Action-sentence Compatibility Effect (ACE, Glenberg and Kaschak, 2002) showing that shifting perspective from first to third person was sufficient to prevent the activation of sensorimotor representations, abolishing the behavioral ACE. Critically, the ACE was restored by adding a virtual “body” that allowed participants to know “where” to put themselves in space when taking the third person perspective, thus demonstrating that motor embodied processes are space-dependent. In addition, Gianelli et al. (2013) recently showed how the social and spatial perspective conveyed by the physical presence of another participant and by linguistic productions,

affect a simple reaching-grasping task. However, the focus of these studies was either on complete sentence comprehension (Gianelli et al., 2011) or employed a complex manipulation of social intentions (Gianelli et al., 2013). Moreover, in both studies the agent's perspective (i.e., the actual motor information) was not manipulated and the interplay between linguistic and motor perspective was only limited.

The present study addressed the role of perspective by using sentence fragments starting with the three singular personal pronouns, You, He/She and I.

The first manipulation we introduced, i.e., the use of three pronouns, allowed us to disentangle two alternative hypotheses. In the first one, language structure would exactly reflect the action structure regardless of linguistic perspective, as assumed by standard embodied cognition theories. If this is the case, then while reading simple pronoun-verb pairs we should automatically activate an agent-independent sensorimotor representation. This would imply that similar motor effects should be detected regardless of the pronouns and thus linguistic perspective. In the other, a more flexible view of embodied cognition would predict the activation of different motor patterns as implied by different linguistic pronouns and hence perspectives. If this is the case, then the pronoun YOU would likely activate the agent's perspective thus modulating motor responses, according to the motor information given by the motion verbs (i.e., action vs. interaction pattern). On the contrary, the pronoun I should be perceived as conveying an observer's perspective, thus activating motor information at a lesser and/or different extent since no contextual information was given. Similarly, the HE/SHE pronoun should activate a completely external perspective, thus producing no modulation of kinematics parameters at all.

The second manipulation we introduced concerns the kinds of action verbs we selected. We addressed the possibility to detect if the agent's perspective is activated, and how detailed it is, by manipulating the motor nature of the action verbs composing our sentence fragments. In particular, we decided to focus on the hypothesis that actions are structured into chains of motor acts, informed by the overall action goal. A variety of results obtained initially with monkeys and then with humans (Fogassi et al., 2005; Iacoboni et al., 2005; Cattaneo et al., 2007; Boria et al., 2009; Fabbri-Destro et al., 2009) show that a mechanism of motor chains constitutes one of the basic structures of the motor system. A chain of motor acts is informed by the final action goal, thus motor acts are organized in the chain so that each of them depends on the successive and all depend on the last. Goals characterize both single motor acts and actions as a whole (for a computational model of chained organization in language, see Chersi et al., 2010). Because of these basic properties, the motor chain structure is an ideal target for disentangling whether and how an agent perspective is activated during linguistic processing of actions.

To this aim, we constructed very simple sentence fragments composed by a pronoun and a motion verb, with verbs being divided into two main categories, that we called *action verbs* and *interaction verbs* (AVs, IVs) (e.g., grasp vs. give). Action and interaction verbs differed according to various dimensions (for

a similar approach, see Kemmerer and Gonzalez-Castillo, 2010). First, the two categories differed for the relations they describe and involve: in one case the direct relation subject-object, in the other case the triadic relation subject-object-other subject. Second, they differed for how these relations imply different goals: AVs are actions which may stand alone and whose final goal might be the sole manipulation of an object, whilst IVs directly imply the interaction with another person. Third, they differed as to the organization in motor chains: AVs and IVs share the motor act of reaching-grasping an object, while they differ for the last act of the sequence, the one determining all the others, which might imply or not the presence of another person. Thus, even if the first part of the motor chain is common, the chain is embedded within two different goals, one of which involves the interaction with another person. Previous kinematics literature has shown higher accuracy with actions guided by a social intention (Becchio et al., 2008; Ferri et al., 2011; Gianelli et al., 2013; Scorolli et al., 2014). However, to our knowledge the "social accuracy" effect has been never investigated distinguishing in the linguistic domain. We predict that IVs lead to higher accuracy compared to AVs in correspondence with the planning of the final motor act, the one that implies an interaction with another person and that qualifies the overall goal of the fragment.

In the experiment participants were required to reach and grasp an object (the mouse) while reading a sentence fragment composed by a pronoun and a verb. The task we chose was designed in order to induce participants to pay attention to both the pronoun and the verb: for this reason, once identified the verb and grasped the mouse, they were required to continue the movement and to click the mouse if the pronoun and the verb matched ("io prendevo," I took) and to refrain from continuing the movement if the pronoun and the verb did not match ("io prendeva," You took, wrong in Italian since the pronoun refers to the first person and the verb to the third one). The task allowed us to investigate the development of the effect of linguistic stimuli on the overt action of reaching and grasping, through the analysis of its fine-grained kinematics aspects. Our general aim was to disentangle the final effects of the two components, pronouns and verbs, and at the same time to understand how their effects are combined producing a modulation of various phases of movement kinematics.

First, we intend to test if the pronouns affect the adopted perspective, influencing the motor response. If the agent's perspective is automatically activated, regardless of linguistic perspective, then motor effects should be present in all conditions and thus not significantly differ among these. If the activation of sensorimotor representations is instead flexible, we should then find effects only with pronouns that activate a first-person, that is the agent's, perspective (i.e., YOU).

Second, we intend to test whether the perspective activated by pronouns is modulated by the motor chains implied by the two kinds of verbs, influencing the motor responses. If the perspective-related sensorimotor activations are general and abstract, then no effect of verb category should be detected. If, on the contrary, the degree of activation is such that the typical motor chain organization is activated, then processing AV or IV verbs should produce detectable motor outputs. In particular, the

different structure of AV and IV should be mapped onto specific parameters of the motor response, i.e., those connected with the velocity peak and its latency since they are typically affected by increased accuracy requirements (i.e., namely with those actions that IVs describe).

EXPERIMENT 1

Methods

Participants

Twelve women, aged 18–28, participated in this study, and were recruited among Communication students at the University of Bologna. All participants were right-handed by self-report, native Italian speakers and reported normal or corrected-to-normal vision. All were naive as to the purpose of the experiment and gave their informed consent to the experiment, which was approved by the local Ethics Committee of the University of Bologna.

Procedure

The experiment took place in a soundproof room. Participants sat in front of a laptop, whose LCD monitor was set on a temporal resolution of 60 Hz. The distance between hand and monitor was of 60 cm. Participants started placing their right hand on the table in a pinch position. The target of the subsequent reaching-grasping movement was a mouse, placed in line with the hand of the participant, at a distance of 33 cm. The final position for the mouse movement was set at 50 cm. The hand movement was performed on the right of the laptop, at a distance of 5 cm. This allowed participants to easily perform the movement and look at the screen.

Stimuli

Stimuli consisted of ten Italian verbs referring to manual actions (see **Table 1**). We selected five proper “action” verbs (AVs), which involved a direct relation subject-object (e.g., *to grasp*) and five “interaction” verbs (IVs), involving at least a relation subject-object-subject (e.g., *to give*). A sample of sixteen students evaluated these verbs on two 7-point scales, one aimed to rate how much the verbs implied a relation subject-object (action scale), the other how much the verbs involved another person (interaction scale). An ANOVA performed on the mean ratings (considering two types of verbs and two type of ratings) showed a significant interaction [$F(1,15) = 15, 2, MSE = 21.39, p = 0.001$] between verb type and rating. As predicted, AVs obtained higher values in the action scale, whilst IVs obtained higher values in the interaction scale. Two additional independent groups of ten students each evaluated the same verbs on two 7-point scales for concreteness and abstractness. An ANOVA performed on the mean ratings (considering two types of verbs and two scales) showed a main effect of scale: in general all verbs were evaluated as more concrete than abstract [$F(1,9) = 22.296, MSE = 14.16, p < 0.002$], an expected result since we focused on choosing verbs with a specific action-relatedness. More interestingly, a significant interaction of verb type and scale was also detected [$F(1,9) = 25.857, MSE = 29.93, p < 0.001$].

While the evaluation of IVs tended to be constant along the two scales, ($M = 4.36$ vs. $M = 3.82$), AVs were evaluated higher in the concreteness scale ($M = 5.8$ vs. $M = 2.88$). However, a Newman-Keuls *post hoc* test revealed that AVs and IVs did not significantly differ in the abstractness scale ($p > 0.05$), but they differed in the concreteness scale ($p < 0.05$). This was expected, since we selected AVs as specifically related to object interaction and manipulation, whereas IVs imply a relation with another subject, which can be considered as less concrete. Furthermore, IVs are often related to abstract sentences or expressions, which could explain a tendency to associate them with more abstract contexts. Nevertheless, both AVs and IVs had low scores in the abstractness scale and did not significantly differ: it seems then unlikely that the observed effects were due to this property and not to the experimental manipulations.

For each verb we identified the isolation point (IP), intended as the minimum part of the verb required to understand it and to differentiate it from similar verbs. In our stimuli the IP corresponded to the verbal stem, as showed in **Table 1**. The final set of stimuli was fully balanced for syllables, length, IP duration, and written lexical frequency (ColFIS, Bertinetto et al., 2005).

Each verb was presented in written form in the three singular persons of the Italian past tense in order to compose sentence fragments. In Italian the pronoun can be omitted, as the verb contains information on the person. However, in our case, using both the pronoun and the past tense, we obtained a double reference to the agent. The final set of stimuli comprised 10 verbs, each presented once in combination with one of the three pronouns (30 critical trials). We inserted also 10 catch-trials, i.e., verbs in the same tense as the others but incorrect for the correspondence verb-subject, e.g., “io portava”: in this case the explicit subject is a first person pronoun while the verb refers to the third person. Catch-trials required participants to refrain from completing the movement and were not analyzed further. The experiment was run in a single block of 40 trials.

Experimental Design

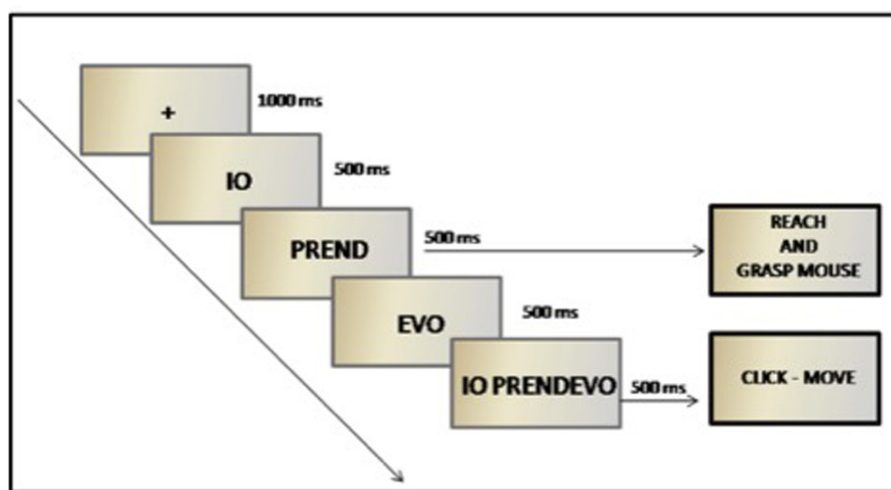
As described in **Figure 1**, each trial started with a fixation cross (1000 ms). Then a pronoun was shown for 500 ms, followed by the first part of the verb (e.g., “prend”) displayed for 500 ms. Subjects were required to pay attention to both the pronoun and the verb, and once they recognized the verb they had to start moving as fast as possible to reach for and grasp the mouse in front of them. During the movement the verb was completed with its suffix (e.g., “evo”) (500 ms). This time was sufficient to accomplish the movement at about the same time in which the complete stimulus “io prendevo” (I took) was presented (time limit of 500 ms). Participants held the hand on the mouse till they decided whether the sentence was correct or not. If correct, they had to click on the left button and then move the mouse to the final position. Otherwise, they had to refrain from moving.

Data Recording and Kinematic Analysis

Movements of the participant’s right hand were recorded using the 3D-optoelectronic SMART system (BTS Bioengineering,

TABLE 1 | Complete list of stimuli in Experiments 1 and 2.

Infinitive	Verb	1st	2nd	3rd	English
Experiment 1-Italian					
afferrare	AV	io afferravo	tu afferravi	egli afferrava	to grasp
alzare	AV	io alzavo	tu alzavi	egli alzava	to raise
portare	AV	io portavo	tu portavi	egli portava	to carry
prendere	AV	io prendevo	tu prendevi	egli prendeva	to take
sol levare	AV	io sollevavo	tu sollevavi	egli sollevava	to lift up
consegnare	IV	io consegnavo	tu consegnavi	egli consegnava	to deliver
dare	IV	io davo	tu davi	egli dava	to give
offrire	IV	io offrivavo	tu offrivavi	egli offriva	to offer
porgere	IV	io porgevo	tu porgevi	egli porgeva	to hand
scambiare	IV	io scambiavo	tu scambiavi	egli scambiava	to exchange
Experiment 2-German					
packen	AV	ich packe	du packst	er packt	to pack
greifen	AV	ich greife	du greifst	er greift	to grasp
heben	AV	ich hebe	du hebst	er hebt	to lift
holen	AV	ich hole	du holst	er holt	to get
ergreifen	AV	ich ergreife	du ergreifst	er ergreift	to seize
schnappen	AV	ich schnappe	du schnappst	er schnappt	to grab
fassen	AV	ich fasse	du fasst	er fasst	to take
bringen	IV	ich bringe	du bringst	er bringt	to bring
reichen	IV	ich reiche	du reichst	er reicht	to hand
liefern	IV	ich liefere	du lieferst	er liefert	to supply
tauschen	IV	ich tausche	du tauschst	er tauscht	to exchange
geben	IV	ich gebe	du gibst	er gibt	to give
bieten	IV	ich biete	du bietest	er bietet	to offer
stiften	IV	ich stifte	du stiftest	er stiftet	to donate

**FIGURE 1 | Experimental procedure.**

Milano, Italy), by means of three infrared cameras at a sampling rate of 60 Hz. Recorded data were filtered using a linear smoothing low pass filter and stored for offline analysis. We used three markers, one applied on the wrist, and the other two on the nail of the index and of the thumb finger respectively.

We considered two components of movement, reaching and grasping, and for each of them we identified different parameters. We avoided considering the act of giving/placing of the mouse due to the high variability of the performed movements.

For the reaching component we analyzed the behavior of the marker placed on the wrist. We considered the reach time, the time to velocity peak (latency), the % of time to velocity peak (normalized with respect to the reach time), and the amplitude of the velocity peak.

To analyze the grasp component we considered the time course of the distance between the two markers posed on the index and on the thumb finger. We analyzed the following parameters: grasp time, maximal finger aperture, time to maximal finger aperture (latency) and percentage of time to maximal finger aperture by means of the software Smart Analyzer and a customized Matlab script. We followed rules and conventions defined by Gianelli et al. (2008) to analyze the different components; in summary: based on the spatial resolution of the system, the reach beginning was defined as the first frame in which the displacement of the wrist marker exceeded 0.3 mm in all Cartesian axes; conversely, to determine the reach end, we first defined the first frame after velocity peak in which the displacement of the reach marker was <0.3 mm along the three axes. The frame (x, y, or z) closer to the grasp end time was selected as reach end. As to the grasp, grasp beginning was defined as the first frame in which the distance between the two markers exceeded 0.3 mm, while grasp end corresponded to the first frame after maximal finger aperture in which the distance between the two markers was less than 0.3 mm. Since reach time and grasp time were defined separately for the two component, normalization with respect to these measures was performed separately for reach and grasp parameters (a similar normalization procedure was applied for instance in Gentilucci, 2002; Gianelli et al., 2008; Ferri et al., 2011).

Data Analysis

Trials with errors (e.g., in the linguistic task, moving when not requested or refraining to do it, anticipated movements, impossibility to correctly separate the reaching of the mouse and the placing during data analysis, etc.) were marked during kinematics analysis and rejected. Participants showing less than 50% of valid trials were excluded from the statistical analysis.

Data analysis was performed only on critical trials (i.e., catch-trial were not analyzed). A repeated measures ANOVA was conducted on the mean values of participants' reaching parameters, considering Verb (AV vs. IV) and Pronoun (I, YOU,

HE) as within-subjects factors. For each significant parameter we report also an estimate of the effect size (η_p^2).

Results

The percentage of errors was negligible (under 1.5%), thus participants correctly understood the word pairs in order to perform the grammatical task and correctly performed the requested motor response. No participant was excluded from data analysis. All results are summarized in **Table 2**.

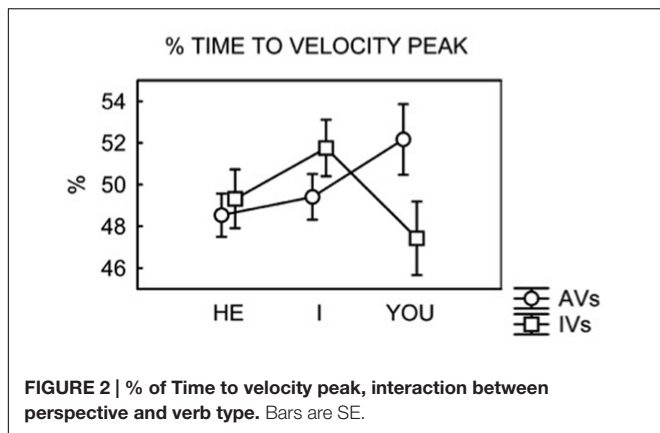
Reaching Component

During the act of reaching we observed no significant main effects of Verb or Pronoun. However, the analysis showed a significant interaction Verb–Pronoun in the normalized % of time to velocity peak, $F(2,22) = 6.48$, $p = 0.006$, $\eta_p^2 = 0.4$. Following our predictions, t -test comparisons were then used to detect the differences between the two kinds of verbs (action vs. interaction) in combination with the three pronouns. A Bonferroni correction for multiple comparisons was applied, with a p -value fixed at 0.01. The result showed that the only significant difference was between AVs and IVs when preceded by the pronoun YOU, $t(11) = 2.81$, $p = 0.008$ (equivalent to a 4.8% difference between conditions). In this sense, IVs showed a shorter time to reaching the velocity peak as compared to AVs. This specific pattern is typically detected at the planning stage when a higher accuracy and the programming of a further motor act are required, as it was the case for IVs and not for AVs. This parameter is thus connected to the activation of the agent's perspective, as activated in a conversational framework by the pronoun YOU. The pronoun I slightly modulated the motor responses but did not reach significance, $t(11) = 1.84$, $p = 0.05$. The same was true for the pronoun HE, as it did not modulate the motor responses at all, $t(11) = 0.56$, $p = 0.3$. No other parameters reached significance¹.

¹Following a reviewers' suggestion, we tested the same parameter with a different normalization procedure (for possible issues connected to the use of normalized parameters, see Whitwell and Goodale, 2013). Namely, instead of normalizing the latency of velocity peak with respect to the reach time (as described in the Section Methods), we used the overall movement time as reference measure. Analyses on this new parameter confirmed the statistical significance of the interaction between Verb and Pronoun [$F(2,22) = 4.087$, $p = 0.031$, $\eta_p^2 = 0.3$]. Three paired sample t -tests (with a corrected p -value of 0.02) confirmed the significant difference between AVs and IVs when preceded by the pronoun YOU [$t(11) = 2.5$, $p = 0.01$, with a 4.3% difference between conditions]. On the contrary, no significant effect was detected for conditions involving the pronouns I and HE ($p = 0.2$ and 0.6, respectively).

TABLE 2 | Summary of results in Experiment 1 and 2, all times are expressed in ms.

	I_Action	You_Action	He_Interaction	I_Interaction	You_Interaction	He_Interaction
Experiment 1 – Italian						
Reach time	742,6	685,6	744,8	705,1	729,8	720,7
Time to velocity peak	367,3	355,6	358,6	360,8	345,6	346,8
% Time to velocity peak	49,4	52,2	48,5	51,8	47,4	49,3
Experiment 2 – German						
Reach time	883,4	869,3	868,5	861,2	879,8	894,5
Time to velocity peak	280,0	277,0	283,0	272,3	276,1	281,0
% Time to velocity peak	31,6	31,7	32,3	31,6	31,0	31,2



The results indicate a specific contribution of the pronoun YOU in activating the agent's perspective and thus modulating one key parameter in the reaching component (the normalized latency of velocity peak), whereas the I perspective did not show significant modulations (Figure 2).

Grasping Component

Repeated measures ANOVAs on grasping parameters showed that no parameter reached significance. In particular, the ANOVA on the time to maximal finger aperture (i.e., time between finger opening and the maximal aperture before grasping the object), did not reach significance, $F(2,22) = 2.08$, $p = 0.148$, $\eta_p^2 = 0.2$. At the qualitative level we can see that the YOU pronoun is the one which mostly modulates the differences between the two verbs, showing overall longer times for the pronoun YOU ($M = 491$ ms) than for the pronoun I ($M = 441$ ms).

Discussion

The results of the study indicate that motor responses are influenced both by the perspective induced by the pronouns and by the different kinds of verbs. We namely found an interaction between the kind of pronoun and the kind of verbs in the analyses on the reaching component. When they were preceded by the YOU pronoun, Interaction verbs reached the velocity peak earlier than Action verbs; this determines a longer deceleration phase. The longer deceleration phase is compatible with the fact that the current motor act is influenced by the next one, i.e., that the action of grasping is influenced by the subsequent action of giving.

The difference we found between Interaction and Action verbs, when preceded by YOU, suggests that when adopting the agent perspective (recruited by the YOU pronoun) we are sensitive to the motor chain structure of verbs. Indeed, while with Action verbs the action terminates once the object is grasped, with Interaction verbs a further motor act follows, since the object has to be given to somebody else.

Furthermore, the finding that the YOU pronoun modulated reaching suggests that, once we read action verbs, we do not automatically assume the agent perspective, but that the adopted perspective is flexible and depends on the presented pronoun.

A research question remains, however, open. While our results demonstrate that in Italian the linguistically presented pronoun influences the motor system, it remains to be determined whether such an influence varies depending on the spoken languages. It is indeed possible that such an influence is present only in languages as Italian where the pronoun assumes salience when mentioned, since it is not mandatory. For this reason in Experiment 2 we adapted the same design and rationale to German stimuli. In the introduction of Experiment 2 we will explain in more detail why we choose to perform a study with a similar inspiration in German.

EXPERIMENT 2

The results of Experiment 1 showed that the combination of pronouns and verbs affects movement execution in a way supporting the hypothesis that an agent's perspective is flexibly activated only under certain conditions and not others (e.g., external perspective). In Experiment 2 we intended to produce a conceptual replication with the same design and rationale but with German stimuli and participants. The reason why we decided to compare Italian and German language is that the role played by the pronouns in the two languages is profoundly different. Italian is a language in which the explicit pronoun can be omitted as the verb already conveys this information (a pro-drop language); however, the relative position of the pronoun and the verb in the sentence is very strict. In German, instead, the use of the pronoun is mandatory, and is very often decisive for revealing the exact subject of a verb. Nevertheless, German speakers are used to a much more flexible sentence construction and word order: previous research, for instance, has shown how this flexibility makes easier for German speakers' to comprehend and produce constructions, such as passive sentences, that result harder to process for other languages' native speakers (see Armon-Lotem et al., 2016).

Both these characteristics can render the pronoun, when mentioned, less salient in German compared to Italian. Consistently with this interpretation, while previous data in Italian (our own, but also Papeo et al., 2011) seem to point to different motor activations according to different perspectives, the only data available in German (Tomasino et al., 2007) suggest that no difference is present between first and third person action verbs. Experiment 2 is therefore aimed at investigating whether the same effects of language-induced perspective and on the motor responses we have found in Experiment 1 are present also in German, a language in which the pronoun is mandatory.

It is worth of notice that, even if we built the experiment in Italian and German starting from the same hypotheses and inspiration, the two experiments are not directly comparable. The choice to use Italian and German had indeed a consequence on the experimental stimuli we selected: in order to be able to correctly identify the verb IP, intended as the minimum part of the verb required to understand it, we had to choose verbs of different tenses in the two languages –past in Italian and present in German (see the method section of Experiments 1 and 2 for further information). Even if the two experiments

are not directly comparable, Experiment 2 can be informative as to how participants process simple sentence fragments and the perspective information they convey, by producing a set of data obtained starting from the same hypotheses and inspirations in a language with different structural characteristics, such as German.

Methods

Participants

In order to estimate the required sample size we used the effect size derived from Experiment 1 in order to establish the aimed sample size for Experiment 2. To this aim we used the software G*Power (Version 3.1.6, University of Duesseldorf) procedure for repeated measures ANOVA and used the effect size estimated derived by the significant Verb*Pronoun interaction in Experiment 1 (setting alpha at 0.05 and the desired power to 0.95). The resulted sample size of sixteen participants was thus used as a stopping rule for data collection in this experiment, with no replacement except in case of technical issues occurred during data recording (e.g., the participant is immediately rejected during the experiment because of the lack of a complete data set). In this case, we tested a total of nineteen participants, of which three did not provide a complete dataset because of technical issues – a sample of sixteen complete data sets thus entered data analysis. All participants were Psychology students at Potsdam University, native German speakers (all women, age 19–35), right-handed by self-report. As confirmed by a standard Edinburgh questionnaire (Oldfield, 1971), they had normal or corrected-to-normal vision and gave their written informed consent as requested by the local Ethics procedures. They took part to the experiment in exchange of course credits.

Procedure

The experiment took place in a soundproof room. Participants sat in front of a PC with the monitor set to a temporal resolution of 60 Hz. The distance between hand and monitor was of 60 cm. Participants started placing their right hand on the table in a pinch position. The target of the subsequent reaching-grasping movement was a mouse, placed in line with the hand of the participant, at a distance of 35 cm. The final position for the mouse movement was set at 50 cm.

Stimuli

Experiment 2 was built on the same principles and categories of experiment one. First, verbs pertaining the two categories were selected resulting in seven AVs and seven IVs (see **Table 1**). In order to select the stimuli, a sample of 34 psychology students recruited at the University of Potsdam filled an online questionnaire in exchange of course credits. All participants were native German speakers and were asked to evaluate each verb (given in the infinitive form) according to the same 7-point scales used for Experiment 1: action and interaction scales, as well as concreteness and abstractness ones.

As in Experiment 1, we first compared the results of the action vs. interaction scales by means of a 2*2 ANOVA with verb type (action, interaction) and rating scale (action, interaction) as factors. The results showed a significant main effect of verb

type [$F(1,33) = 56, 22, MSE = 13.03, p < 0.001$] and an interaction between verb type and rating scale, [$F(1,33) = 47, 52, MSE = 18.46, p < 0.001$]. While both verbs were similarly rated along the action scale, IVs were rated significantly higher in the interaction scales as compared to AVs [paired-sample *t*-test comparison, $t(33) = -588, p < 0.001$].

In a second ANOVA we compared the results of the concreteness vs. the abstractness scales by means of 2*2 ANOVA with verb type (action, interaction) and rating scale (concrete, abstract) as factors. As in the first experiment, a main effect of scale [$F(1,33) = 195.203, MSE = 114.706, p < 0.001$] shows that overall verbs were evaluated higher in the concreteness than in the abstractness scale, as we selected verbs with a specific action-relatedness. This main effect seems to drive the significant interaction we also found between verb type and rating scale [$F(1,33) = 8.901, MSE = 0.972, p = 0.005$]. As in experiment one, AVs and IVs did not differ along the abstractness scale [$t(33) = -1.527, p = 0.136$], while they differed along the concreteness scale with AVs being evaluated slightly higher in the concreteness scale [$M = 4.279$ vs. $4.074, t(33) = 3.956, p < 0.001$]. The same considerations regarding this scale for Experiment 1, hold for these stimuli as well.

No verb was excluded at this stage and all fourteen verbs entered one last linguistic evaluation with the aim of being matched for syllables, length, and written frequency (database: dlexDB). As in the case of experiment one, we selected a tense in which it would be acceptable to split the verb in two between the stem of the verb and the suffix that contains the information relative to tense and subject. To this aim, we selected the present tense of regular German verbs, as it fulfills our requirements, e.g., “Ich greife” vs. “Du greifst” vs. “Er greift” (I grasp, You grasp, He grasps). As a clarification, past tense would not have worked, being respectively “Ich griff,” “Du griffst,” “Er griff,” and for different reasons the same holds for composite forms as the perfect. As we already noticed, compared to Experiment 1, in German the presence of both the personal pronoun and the subject information given by the verb is mandatory (all verbs are listed in **Table 1**).

Each verb was presented in written form in the three singular persons of the German present tense in order to compose sentence fragments. The final set of stimuli comprised 14 verbs, each presented twice in combination with one of the three pronouns (84 critical trials). We inserted also 16 catch-trials, i.e., verbs in the same tense as the others but incorrect for the correspondence verb-subject, e.g., “er greift”: in this case the explicit subject is a third person pronoun while the verb refers to the second person. Catch-trials required participants to refrain from completing the movement and were not analyzed further. The experiment was run in a single block of 100 trials.

Experimental Design

The procedure was the same as described for Experiment 1: each trial started with a fixation cross (1000 ms). Then a pronoun was shown for 500 ms (“Ich”) followed by the first part of the verb (“greif”) displayed for 500 ms. Subjects were required to pay attention to both the pronoun and the verb, and once they recognized the verb they had to start moving as fast as possible

to reach for and grasp the mouse in front of them. During the movement the verb was completed with its suffix (“e”) (500 ms). This time was sufficient to accomplish the movement at about the same time in which the complete stimulus “Ich greife” was presented (time limit of 500 ms). Participants held the hand on the mouse till they decided whether the sentence was correct or not. If correct, they had to click on the left button and then move the mouse to the final position. Otherwise, they had to refrain from moving.

Data Recording and Kinematic Analysis

Movements of the participant’s right hand were recorded by means of a 3D guidance tracking system (Trakstar, Ascension) with a sampling rate of 200 Hz, filtered using a linear smoothing low pass filter and then stored for offline analysis.

The choice of the movement components and movement parameters were guided by the results of the first experiment. As Experiment 1 showed effects pertaining only the reach component of movement, we focused on the analysis of one sensor placed on the participants’ right wrist and analyzed parameters related only to this component. As in the first experiment, we analyzed the reach time, time to velocity peak (latency), % of time to velocity peak (normalized with respect to the reach time), and the amplitude of the velocity peak by means of a customized Matlab script. In this case, reach beginning and end were determined as the first and last frame in which the velocity was > 1 mm/s. Normalization procedures were the same as in Experiment 1.

Data Analysis

Data analysis was the same as in Experiment 1 and it was performed only on the critical trials. Trials with errors (e.g., in the linguistic task, hence moving when not requested or refraining to do it, anticipated movements, impossibility to correctly separate the reaching of the mouse and the placing during data analysis etc.) were marked during kinematics analysis and rejected. Participants showing less than 50% of valid trials were excluded by statistical analysis. Statistical analysis was the same as in Experiment 1.

Results

Three participants were excluded from statistical analysis based on the number of valid trials. The remaining thirteen participants entered the statistical analysis with a total of 92% of analyzed trials equally distributed across all participants and conditions (13 trials on average per condition).

Reaching Component

Statistical analysis (ANOVA) showed no significant main effect or interaction for any of the selected parameters (all $p_s > 0.05$). In particular, the critical parameter of % of velocity peak (significant in Experiment 1) resulted in a $F(2,24) = 0.341$, $p > 0.7$, $\eta_p^2 = 0.028$, with a difference as small as 0.6% between the you-action and you-interaction conditions (all data are summarized in **Table 2**). According to significance testing, then, no effect of perspective was detected in the second experiment, hence

providing no evidence for a similar effect in the Italian and German experiments.

Discussion

Experiment 2 aimed to verify whether the different pronouns and the two different kinds of verbs had an influence on motor response in German, a language chosen because, differently from Italian, pronouns are mandatory while the sentence construction is flexible. The results of Experiment 1 were not replicated. We will discuss the possible reasons of the missing effects in the Section “General Discussion.” To have a better idea of what happened in the two experiments, we analyzed them also using a Bayesian approach.

Exploratory Bayesian Analyses

As already shown, planned analyses for both experiments were based on classical null hypothesis significance testing (NHST) and relative estimation of effect size. Under this respect, Experiment 1 clearly showed a significant modulation of reaching parameters while Experiment 2 showed no significant effects. The significance tests thus leave the contribution of Experiment 2 unclear: how strong is the observed evidence against a modulation of kinematics parameters in German? Nevertheless, is the significant modulation observed in Experiment 1 substantial or just anecdotal?

In order to investigate these issues and complement our results, we performed an additional, exploratory analysis taking a Bayesian approach, with the aim to quantify the observed evidence in terms of odds ratio between the null and the alternative hypothesis. To this aim we report the results of two JZS Bayes factor ANOVA (using JASP, Rouder et al., 2012; Morey and Rouder, 2013; Love et al., 2015) with default prior scales, based on the data on the crucial parameter of % velocity peak for both experiments separately. In addition, and since the % of velocity peak is a normalized measure determined by the latency of velocity peak and the reach time, we tested these two parameters as well, although they did not show any difference in the significance tests.

For Experiment 1, the % of time to velocity peak shows a $BF_{10} = 7.301$ (that is a $BF_{01} = 0.137$) for the model comprising the interaction between the two factors, verb and subject type (as compared to the null) providing substantial evidence in support of the alternative hypothesis. In other terms, the observed data are seven times more likely to occur under H_1 . For Experiment 2, the same parameter produced a $BF_{10} = 0.23$ for the same comparison (that is a $BF_{01} = 4.378$), providing no evidence in favor of the alternative hypothesis, with the observed data being far more likely to occur under H_0 . As to the other parameters, reach time showed comparable BFs in the two experiments ($BF_{10} = 0.735$ vs. 1.083, that is $BF_{01} = 1.36$ vs. 0.92), with the observed data almost equally likely to occur under H_0 or H_1 . On the other side, the latency of velocity peak in Experiment 1 produced a $BF_{10} = 0.420$ ($BF_{01} = 2.381$) and in Experiment 2 $BF_{10} = 0.203$ ($BF_{01} = 4.94$), that is the observed data are more likely to occur two times and almost five times under H_0 than H_1 in both experiments.

GENERAL DISCUSSION

The results we found in the experiment in Italian and in German are quite different. We will first discuss the overall issue of whether language influences the motor system considering the results of the two studies. Then we will discuss more specific issues, i.e., the role of pronouns and verbs in light of the results of the first experiment. Finally we will outline the possible reasons why we found different results in the two languages.

Language and Flexible Involvement of the Motor System

The results of our exploratory kinematics analysis in Experiment 1 showed the presence of distinct motor patterns as influenced both by the perspective elicited by the pronouns and the motor chains elicited by action verbs.

The interaction between verbs and pronouns found in Experiment 1 suggests that the effect of modulation due to language occurs early during the actual movement and is evident in a range of 300–350 ms after stimuli presentation. Interestingly, the parameter in which we found a modulation is connected to the velocity peak. We know well that the velocity peak is the main parameter which is defined in movement planning and it is susceptible to be affected by the various factors (motor factors as in Gentilucci et al., 1997; Dalla Volta et al., 2009 or social factors as in Ferri et al., 2010) under which the movement is executed. Consequently, the effect of our stimuli on this crucial parameter suggests that our stimuli mainly affected the planning stage of action. This early influence of linguistic processing on the motor system suggests that the activation of the motor system is not due to late-occurring imagery processes; it is therefore consistent with the view according to which the activation of the motor and sensorimotor cortices is not just a side effect but effectively contributes to language comprehension.

While the interaction Verb–Pronoun found in Experiment 1 clearly indicate that pronouns and verbs differently influence the motor system, the absence of a perspective-related modulation following German stimuli might point to the activation of agent-independent sensorimotor representations. However, we do not believe that we can draw such a conclusion. Indeed, the absence of a baseline/reference condition (e.g., movement in absence of linguistic stimuli, or following unrelated stimuli) in this design does not allow us to disentangle whether the results in Experiment 2 are the product of an homogenous motor activation for all perspectives or the absence of it. In a more nuanced view, future studies should also clarify the relationship between kinematics and behavioral measures and the magnitude of the same effect at the neurophysiological level. Kinematics results of Experiment 2 might be the product of neural effects similar to Experiment 1 but weaker, which translate into no effect on overt movement execution. Future studies directly comparing the same manipulation with different techniques are highly recommended.

Overall, our results indicate that the influence of language on the motor system is likely not automatic but highly flexible and context dependent. Our results namely showed that the motor system activation was strongly influenced by the used pronoun:

we found evidence for it with the YOU pronoun and not with the third pronoun, and we found that the YOU pronoun differently influenced the motor response depending on the verb with which it was combined. The effect was also modulated by the spoken language, since the interaction Verb–Pronoun was not present in German but only in Italian.

The way the motor planning was affected by language in Experiment 1 was undoubtedly interesting, since both perspectives induced by pronouns and chain organization of verbs seemed to be involved. We will first handle the role of perspective and of action verbs in Experiment 1, and then we will discuss why the same effects were not found in Experiment 2.

Pronouns and Perspective Taking

Results of Experiment 1 clearly reveal that the perspective induced by the pronoun affects the motor system. Specifically, our data show a strong effect of the YOU perspective in modulating both action and interaction verbs, and notably this pattern is present in all our subjects. This complements and extends the results obtained by Brunyé et al. (2009), since we used a motor task and demonstrated that perspective modulates the very first stages of actions planning and subsequently execution. Our preliminary results are also consistent with the previous studies where the strongest compatibility/facilitation effects (Glenberg and Kaschak, 2002; Gianelli et al., 2011) are obtained with sentences using YOU or with the infinitive form of the verbs, where the perspective activated is necessarily the one of the agent.

Our results suggest instead that the perspective elicited while reading the pronoun HE is more abstract and external, so that the motor effects of language processing disappear. This might appear in contrast with the results obtained in a recent behavioral and TMS study by Gianelli and Dalla Volta (2015) who showed a motor facilitation with the use of a passive listening task for third-person Italian sentences. However, in this study the authors used only a passive listening task and implicit agent's attribution (i.e., no pronoun) and stimulated the motor cortex before the agent's information was made explicit. In addition, only third-person sentences were presented, with no perspective manipulation. Further studies are thus needed to investigate under which conditions the third person perspective activates an agent perspective and at which degree. Interestingly, what happens for HE seems to be true – at least partially – for the I perspective as well. The I perspective may involve the subject a bit more than the HE perspective. However, overall our results point to the idea that the role of agent is taken when the YOU pronoun is used. In this condition it appears that the participants are called directly into action and then they re-activate the motor pattern of an action from the point of view of the agent. I and HE constitute external and “observational” perspectives but at different degrees. In an inter-subjective framework, as for example in a conversation, the use of the pronoun I normally refers to the presence of a speaker who is reporting the action from his/her point of view, whereas we (i.e., the readers) are recruited as recipients of his/her speech. In the case of the pronoun HE, a radically external perspective is assumed. Consider for instance a situation in which we and another person

are talking of the actions of a third person: his/her perspective does not involve us directly.

Overall, the results of Experiment 1 suggest that while comprehending language we activate an inter-subjective framework, as the role of the YOU pronoun with interactive verbs indicate. This happens even if we are not directly involved in communication but simply read linguistic stimuli. The adoption of this frame of reference has a very precocious effect as it differently impacts the early stages of movements planning and execution. The activation of a conversational framework has an interesting theoretical implication. Even if our study showed that under certain conditions action organization (e.g., the motor chains) might be reflected in language (e.g., in Italian), language imposes its own constraints on the way actions are conceived, giving relevance to the YOU perspective in taking the agent's perspective. In line with evidence on neural re-use of previously built neural structures (Gallese, 2008; Anderson, 2010, 2014), our work shows that language builds on previously formed structures, such as the action chained organization, but also that it strongly constrains and modifies it (Borghi, 2012, for discussion of this issue), as the importance assumed by the YOU perspective clearly demonstrates.

Verbs and Motor Chains

The motor pattern activated by YOU both with AVs and IVs fits well also with our hypothesis about the organization of actions in motor chains, supporting the notion that an agent's perspective is activated. In fact, IVs result in a shorter time to velocity peak, so that conversely the deceleration phase is longer. This is coherent with evidence on motor planning and control of a sequence of motor acts (Gentilucci et al., 1997): an increasing accuracy in interaction with the object influences arm velocity profiles by decreasing the velocity peak and lengthening the deceleration phase. In this sense the current motor act is influenced by the requests of the successive act. AVs do not imply any particular request of accuracy since there is not a second motor act to plan: namely, the action ends with the grasping of the object. This is not the case for IVs where more accuracy is requested in order to interact with the object: indeed, the object should be grasped and given to somebody else. One could speculate that participants are particularly accurate also due to the fact that IVs do not simply involve a further motor act compared to AVs, but that they also involve a social dimension, guaranteed by the virtual presence of a recipient. However, our data do not allow us to definitively solve this issue since no direct social manipulation was designed.

Cross-Linguistic Differences

Once verified that the Italian pronouns influence perspective taking with action verbs, we performed a conceptual replication of the same study in German (Experiment 2), comparable for task and design – e.g., both studies directly manipulate and compare different perspectives in combination with specific verb categories. The reasons why we were interested in performing the same study in another language, and specifically in German, are many. First, we think it is important to conduct cross-cultural studies. In many cases researchers implicitly assume that the phenomena they find hold across different populations,

while often this is not the case (for a recent review, see Henrich et al., 2010). To make general claims it is therefore important to investigate whether the same phenomenon holds in different populations. Second, we believe it is crucial not only to realize cross-cultural, but also cross-linguistic studies. The last years have been characterized by a resurgence of interest for linguistic relativity, the idea that natural languages shape the way we think and conceptualize the world (Whorf, 1956; Casasanto, 2008; Reines and Prinz, 2009). Once identified a phenomenon – in our case the fact that the perspective induced by pronouns influences the motor system – it is important to verify to what extent such phenomenon is generalizable across different natural languages. Our results suggest that the interaction Verb–Pronoun we found is not generalizable to German, and this has theoretical implications since our results are in line with the idea that the language we speak can differently influence perspective taking. A third specific reason is related to the specific differences of Italian and German in the use of pronouns, which are mandatory in German but not in Italian. As anticipated in the introduction to Experiment 2, we intended to investigate whether the effects found in Italian was replicated in German, a language where pronouns play a different role. Experiment 2 did not yield the same results and instead pointed to the absence of difference between conditions, in particular pertaining the crucial interaction of verb type and pronoun. We will now discuss the possible reasons underlying such a discrepancy.

The first and more crucial difference between Italian and German and the reason why we performed the second experiment in German pertains the role of the pronouns. While the use of pronouns is mandatory in German, it is not in Italian. Our results showed that the difference in processing action and interaction verbs with the YOU pronoun was present only in Italian. We interpret this difference as due to the fact that, since the use of pronouns in Italian is not necessary, their presence might be perceived as more salient. This suggests that not language *per se*, but different natural languages have a different impact on perspective taking.

One further possible explanation of the difference we found between the two experiments concerns the tense of verbs: the two experiments do not fully overlap, since we used past tense in Italian, and present tense in German. Although the choice of the two tenses was due to pure methodological reasons (e.g., in keeping with the methodology used in Experiment 1 and the relative kinematics analysis) and this factor was not manipulated, literature suggests that different tenses might indeed lead to different motor activations, supporting a flexible view of embodied language processing (e.g., Bergen and Wheeler, 2010; Candidi et al., 2010). From this point of view, it is possible that different verb tenses activate motor resonance at a different degree, making a stronger motor resonance more capable to affect motor behaviors than weaker ones, especially when combined with certain perspectives (e.g., more internal ones). However, we tend to exclude that the effect is due to the different tenses used since we found a stronger modulation of the motor system in Italian, i.e., when we used the past tense, than in German, when we used the present tense.

We tend rather to believe that the most plausible explanation of the differences in results is due to the structural differences between Italian and German language and in particular to how pronouns differently influence perspective taking.

On the other hand, it is worth considering for future research that linguistic differences between the two experiments are not limited to the differences in linguistic stimuli *per se*. Indeed, we did compare not only two sets of stimuli but also two groups of native speakers whose linguistic habits are very different. The degree to which these linguistic habits could have affected their motor behavior and the way they handled the linguistic task, cannot be solved but only pointed out by the exploratory data we made available. The study of embodied language processing so far has focused on few languages (with a predominance of English, Italian, French, Dutch and more limitedly German) and the direct comparison of different languages in the same study is in most cases absent. This seems indeed surprising as one would clearly expect that different linguistic and motor experiences would affect the encoding of the corresponding linguistic labels, and hence the re-activation of these experiences in terms of motor resonance. Similarly, if while comprehending language we activate an inter-subjective framework (as suggested by Experiment 1), this might occur differently in two languages, being more or less flexible in different groups of native speakers. Our exploratory study points out the need for future studies performing direct cross-linguistic comparisons, and when possible comparing different groups of speakers (e.g., native vs. not native). At the same time, we believe that the use of kinematics and hence of motion analysis, could constitute a powerful tool for such comparisons, making it possible to use the same motor tasks regardless of the tested language.

REFERENCES

- Anderson, M. L. (2010). Neural re-use as a fundamental organizational principle of the brain. *Behav. Brain Sci.* 33, 245–266. doi: 10.1017/S0140525X10000853
- Anderson, M. L. (2014). *After Phrenology: Neural Reuse and the Interactive Brain*. Cambridge, MA: MIT Press.
- Armon-Lotem, S., Haman, E., de López, K. J., Smoczynska, M., Yatsushiro, K., Szczerbinski, M., et al. (2016). A large-scale cross-linguistic investigation of the acquisition of passive language acquisition of passive. *J. Dev. Linguist.* 23, 27–56.
- Barsalou, L. W. (2016). On staying grounded and avoiding quixotic dead ends. *Psychon. Bull. Rev.* 23, 1122–1142. doi: 10.3758/s13423-016-1028-3
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition* 106, 894–912. doi: 10.1016/j.cognition.2007.05.004
- Bergen, B., and Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain Lang.* 112, 150–158. doi: 10.1016/j.bandl.2009.07.002
- Bertinetto, M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., et al. (2005). *Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS)*. Available at: <http://linguistica.sns.it/CoLFIS/Home.htm>
- Beveridge, M. E., and Pickering, M. J. (2013). Perspective taking in language: integrating the spatial and action domains. *Front. Hum. Neurosci.* 7:577. doi: 10.3389/fnhum.2013.00577
- Borghi, A. M. (2012). “Action language comprehension, affordances and goals,” in *Language and Action in Cognitive Neuroscience. Contemporary Topics in Cognitive Neuroscience Series*, eds Y. Coello and A. Bartolo (London: Psychology Press), 125–143.
- Borghi, A. M., and Cangelosi, A. (2014). Action language integration: from humans to cognitive robots. *Top. Cogn. Sci.* 6, 344–358. doi: 10.1111/tops.12103
- Borghi, A. M., Gianelli, C., and Scorolli, C. (2010). Sentence comprehension: effectors and goals, self and others. An overview of experiments and implications for robotics. *Front. Neurobot.* 4:3. doi: 10.3389/fnbot.2010.00003
- Boria, S., Fabbri-Destro, M., Cattaneo, L., Sparaci, L., Sinigaglia, C., Santelli, E., et al. (2009). Intention understanding in autism. *PLoS ONE* 4:e5596. doi: 10.1371/journal.pone.0005596
- Boulenger, V., Roy, A., Paulignan, Y., Deprez, V., Jeannerod, M., and Nazir, T. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *J. Cogn. Neurosci.* 18, 1607–1615. doi: 10.1162/jocn.2006.18.10.1607
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., and Taylor, H. A. (2009). When you and I share perspectives: pronouns modulate perspective taking during narrative comprehension. *Psychol. Sci.* 20, 27–32. doi: 10.1111/j.1467-9280.2008.02249.x
- Bruzzo, A., Borghi, A. M., and Ghirlanda, S. (2008). Hand-object interaction in perspective. *Neurosci. Lett.* 441, 61–65. doi: 10.1016/j.neulet.2008.06.020
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., and Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Brain Res. Cogn. Brain Res.* 24, 355–363. doi: 10.1016/j.cogbrainres.2005.02.020
- Candidi, M., Fernandez, B. L., Barber, H. A., Carreiras, M., and Aglioti, S. M. (2010). Hand on the future: facilitation of cortico-spinal hand-representation when reading the future tense of hand-related action verbs. *Eur. J. Neurosci.* 32, 677–683. doi: 10.1111/j.1460-9568.2010.07305.x

ETHICS STATEMENT

The study did not involve any risk for participants health or wellbeing. Experiment 1 was performed at the University of Bologna, Department of Psychology and was approved by local ethics committee. Experiment 2 was performed at the University of Potsdam, Division of Cognitive Sciences following the local ethics procedures (written informed consent), but being a behavioral study it was exempt by the requirement of formal approval. All participants were informed as to the procedures involved in the experiment and gave their written informed consent.

AUTHOR CONTRIBUTIONS

CG and AB designed the experiment, CG and MM collected and analyzed the data, CG and AB wrote the paper.

ACKNOWLEDGMENTS

We would like to thank the members of the EMCO group in Bologna, Fabian Chersi and the members of the PECoG group in Potsdam for useful discussions on preliminary versions of this paper. We also would like to thank Zsuzsanna Nemecz and Silvia Mencaraglia for supporting data collection in Potsdam. Experiment 1 was supported by the FP7 project ROSSI, Emergence of communication in RObots through Sensorimotor and Social Interaction, Grant agreement no: 216125.

- Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Lang. Learn.* 58, 63–79. doi: 10.1111/j.1467-9922.2008.00462.x
- Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., et al. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17825–17830. doi: 10.1073/pnas.0706273104
- Chersi, F., Thill, S., Ziemke, T., and Borghi, A. M. (2010). Sentence processing: linking language to motor chains. *Front. Neurobot.* 4:4. doi: 10.3389/fnbot.2010.00004
- Dalla Volta, R., Gianelli, C., Campione, G. C., and Gentilucci, M. (2009). Action word understanding and overt motor behavior. *Exp. Brain Res.* 196, 403–412. doi: 10.1007/s00221-009-1864-8
- Ditman, T., Brunye, T., Mahoney, C., and Taylor, H. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition* 115, 172–178. doi: 10.1016/j.cognition.2009.10.014
- Fabbri-Destro, M., Cattaneo, L., Boria, S., and Rizzolatti, G. (2009). Planning actions in autism. *Exp. Brain Res.* 192, 521–525. doi: 10.1007/s00221-008-1578-3
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2010). To me or to you? When the self is advantaged. *Exp. Brain Res.* 203, 637–646. doi: 10.1007/s00221-010-2271-x
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2011). Social requests and social affordances: how they affect the kinematics of motor sequences during interactions between conspecifics. *PLoS ONE* 6:e15855. doi: 10.1371/journal.pone.0015855
- Fischer, M. H., and Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol. (Colchester)* 61, 825–850. doi: 10.1080/17470210701623605
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. doi: 10.1126/science.1106138
- Gallese, V. (2006). Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Brain Res.* 1079, 15–24. doi: 10.1016/j.brainres.2006.01.054
- Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Soc. Neurosci.* 3, 317–333. doi: 10.1080/17470910701563608
- Gentilucci, M. (2002). Object motor representation and reaching-grasping control. *Neuropsychologia* 40, 1139–1153. doi: 10.1016/S0028-3932(01)00233-0
- Gentilucci, M., Negrotti, A., and Gangitano, M. (1997). Planning an action. *Exp. Brain Res.* 115, 116–128. doi: 10.1007/PL00005671
- Gianelli, C., and Dalla Volta, R. (2015). Does listening to action-related sentences modulate the activity of the motor system? Replication of a combined TMS and behavioral study. *Front. Psychol.* 5:1511. doi: 10.3389/fpsyg.2014.01511
- Gianelli, C., Dalla Volta, R., Barbieri, F., and Gentilucci, M. (2008). Automatic grasp imitation following action observation affects estimation of intrinsic object properties. *Brain Res.* 1218, 166–180. doi: 10.1016/j.brainres.2008.04.046
- Gianelli, C., Farnè, A., Salemme, R., Jeannerod, M., and Roy, A. C. (2011). The agent is right: when motor embodied cognition is space-dependent. *PLoS ONE* 6:e25036. doi: 10.1371/journal.pone.0025036
- Gianelli, C., Scorolli, C., and Borghi, A. M. (2013). Acting in perspective: the role of body and of language as social tools. *Psychol. Res.* 77, 40–52. doi: 10.1007/s00426-011-0401-0
- Glenberg, A. M., and Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex* 48, 905–922. doi: 10.1016/j.cortex.2011.04.010
- Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychon. Bull. Rev.* 9, 558–565. doi: 10.3758/BF03196313
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83. doi: 10.1017/S0140525X0999152X
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3:e79. doi: 10.1371/journal.pbio.0030079
- Jackson, P. L., Meltzoff, A. N., and Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *Neuroimage* 31, 429–439. doi: 10.1016/j.neuroimage.2005.11.026
- Jirak, D., Menz, M., Buccino, G., Borghi, A. M., and Binkofski, F. (2010). Grasping language. A short story on embodiment. *Conscious. Cogn.* 19, 711–720. doi: 10.1016/j.concog.2010.06.020
- Kemmerer, D., and Gonzalez-Castillo, J. (2010). The Two-Level Theory of verb meaning: an approach to integrating the semantics of action with the mirror neuron system. *Brain Lang.* 112, 54–76. doi: 10.1016/j.bandl.2008.09.010
- Keyers, C., and Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Prog. Brain Res.* 156, 379–401. doi: 10.1016/S0079-6123(06)56021-2
- Love, J., Selker, R., Marsman, M., Jamil, T., Verhagen, A. J., Ly, A., et al. (2015). *JASP (Version 0.7.5) [Computer software]*.
- MacWhinney, B. (2005). “The emergence of grammar from perspective taking,” in *The Grounding of Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge, eds D. Pecher and R. Zwaan (Cambridge: Cambridge University Press), 198–223.
- Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004
- Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804. doi: 10.1016/j.cortex.2010.11.002
- Morey, R. D., and Rouder, J. N. (2013). *Bayes Factor: Computation of Bayes Factors for Simple Designs (Version 0.9.4) [Computer Program]*. Available at: <http://CRAN.R-project.org/package=BayesFactor>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Papeo, L., and Caramazza, A. (2014). When “ultraparad” word-related motor activity is not faster than “early.” *Front. Hum. Neurosci.* 8:842. doi: 10.3389/fnhum.2014.00842
- Papeo, L., Corradi-Dell'Acqua, C., and Rumiati, R. I. (2011). “She” is not like “I”: the tie between language and action is in our imagination. *J. Cogn. Neurosci.* 23, 3939–3948. doi: 10.1162/jocn_a_00075
- Reines, M. F., and Prinz, J. (2009). Reviving Whorf: the return of linguistic relativity. *Philos. Compass* 4, 1022–1032. doi: 10.1111/j.1747-9991.2009.00260.x
- Rouder, J. N., Morey, R. D., Speckman, P. L., and Province, J. M. (2012). Default Bayes factors for ANOVA designs. *J. Math. Psychol.* 56, 356–374. doi: 10.1016/j.jmp.2012.08.001
- Schütz-Bosbach, S., Mancini, B., Aglioti, S. M., and Haggard, P. (2006). Self and other in the human motor system. *Curr. Biol.* 16, 1830–1834. doi: 10.1016/j.cub.2006.07.048
- Scorolli, C., Borghi, A. M., and Glenberg, A. M. (2009). Language-induced motor activity in bimanual object lifting. *Exp. Brain Res.* 193, 43–53. doi: 10.1007/s00221-008-1593-4
- Scorolli, C., Miatton, M., Wheaton, L. A., and Borghi, A. M. (2014). I give you a cup, I get a cup: a kinematic study on social intention. *Neuropsychologia* 57, 196–204. doi: 10.1016/j.neuropsychologia.2014.03.006
- Tomasino, B., Werner, C. J., Weiss, P. H., and Fink, G. R. (2007). Stimulus properties matter more than perspective: an fMRI study of mental imagery and silent reading of action phrases. *Neuroimage* 36, T128–T141. doi: 10.1016/j.neuroimage.2007.03.035
- Toni, I., de Lange, F. P., Noordzij, M. L., and Hagoort, P. (2008). Language beyond action. *J. Physiol. Paris* 102, 71–79. doi: 10.1016/j.jphysparis.2008.03.005
- Vogt, S., Taylor, P., and Hopkins, B. (2003). Visuomotor priming by pictures of hands: perspective matters. *Neuropsychologia* 41, 941–951. doi: 10.1016/S0028-3932(02)00319-6
- Whitwell, R. L., and Goodale, M. A. (2013). Grasping without Vision: time normalizing grip aperture profiles yields spurious grip scaling to target size. *Neuropsychologia* 51, 1878–1887. doi: 10.1016/j.neuropsychologia.2013.06.015
- Whorf, B. L. (1956). *Language, Thought, and Reality: Selected Writings*, ed. J. B. Carroll (Cambridge, MA: MIT Press).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Gianelli, Marzocchi and Borghi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Multilingual CID-5: A New Tool to Study the Perception of Communicative Interactions in Different Languages

Valeria Manera¹, Francesco Iani², Jérémy Bourgeois¹, Maciej Haman³, Łukasz P. Okruszek³, Susan M. Rivera⁴, Philippe Robert^{1,5}, Leonhard Schilbach^{6,7}, Emily Sievers⁴, Karl Verfaillie⁸, Kai Vogeley^{6,9}, Tabea von der Lühse¹⁰, Sam Willems⁸ and Cristina Becchio^{2,11*}

¹ CoBTeK Laboratory, Faculty of Medicine, University of Nice Sophia Antipolis, Nice, France, ² Department of Psychology, University of Turin, Turin, Italy, ³ Faculty of Psychology, University of Warsaw, Warsaw, Poland, ⁴ Department of Psychology, Center for Mind and Brain & The MIND Institute, University of California, Davis, CA, USA, ⁵ Centre Mémoire de Ressources et de Recherche, CHU de Nice, Nice, France, ⁶ Department of Psychiatry, University Hospital Cologne, Cologne, Germany, ⁷ Max Planck Institute of Psychiatry, Munich, Germany, ⁸ Laboratory of Experimental Psychology, KU Leuven, Leuven, Belgium, ⁹ Cognitive Neuroscience – Institute of Neuroscience and Medicine (INM3), Research Center Jülich, Jülich, Germany, ¹⁰ Department of Psychiatry and Psychotherapy, Heinrich-Heine-University of Düsseldorf, Rhineland State Clinics Düsseldorf, Düsseldorf, Germany, ¹¹ Department of Robotics, Brain and Cognitive Sciences, Fondazione Istituto Italiano di Tecnologia, Genova, Italy

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Luke Edward Miller,
University of California, San Diego,
USA

Janny Christina Stapel,
Uppsala University, Sweden

*Correspondence:

Cristina Becchio
cristina.becchio@unito.it

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 18 May 2015

Accepted: 26 October 2015

Published: 17 November 2015

Citation:

Manera V, Iani F, Bourgeois J, Haman M, Okruszek ŁP, Rivera SM, Robert P, Schilbach L, Sievers E, Verfaillie K, Vogeley K, von der Lühse T, Willems S and Becchio C (2015) The Multilingual CID-5: A New Tool to Study the Perception of Communicative Interactions in Different Languages. *Front. Psychol.* 6:1724. doi: 10.3389/fpsyg.2015.01724

The investigation of the ability to perceive, recognize, and judge upon social intentions, such as communicative intentions, on the basis of body motion is a growing research area. Cross-cultural differences in ability to perceive and interpret biological motion, however, have been poorly investigated so far. Progress in this domain strongly depends on the availability of suitable stimulus material. In the present method paper, we describe the *multilingual CID-5*, an extension of the CID-5 database, allowing for the investigation of how non-conventional communicative gestures are classified and identified by speakers of different languages. The CID-5 database contains 14 communicative interactions and 7 non-communicative actions performed by couples of agents and presented as point-light displays. For each action, the database provides movie files with the point-light animation, text files with the 3-D spatial coordinates of the point-lights, and five different response alternatives. In the multilingual CID-5 the alternatives were translated into seven languages (Chinese, Dutch, English, French, German, Italian, and Polish). Preliminary data collected to assess the recognizability of the actions in the different languages suggest that, for most of the action stimuli, information presented in point-light displays is sufficient for the distinctive classification of the action as communicative vs. individual, as well as for identification of the specific communicative gesture performed by the actor in all the available languages.

Keywords: point-light display, biological motion, communicative interaction, communicative intention, individual intention, cross-linguistic comparisons, forced choice

INTRODUCTION

Successful gestural communication depends on the recipient understanding and recognizing the intention of the communicative act (Sperber and Wilson, 1986). To do so, the recipient needs to be able to (a) discriminate between communicative gestures and individual actions (intention classification), and (b) identify the specific communicative content conveyed by the gesture

(intention identification). Conventional emblematic communicative gestures, such as the 'okay' sign or the 'thumbs-up' gesture, have a certain format and an explicit meaning established by the conventions of specific communities. It is thus unsurprising that they may have radically different meanings from one society to another, or even within a single communicative tradition. 'The horns,' made by extending the pinkie and index finger while making a fist, for example, is used to ward off the evil eye in traditional Mediterranean cultures. Variants of this gesture were used in Elizabethan England to accuse a man of having an unfaithful wife, in modern England and the US to express a passion for heavy metal music (Casasanto, 2013). Non-conventional gestures, on the contrary, may be more easily understood across cultures. Pointing when giving directions, reaching up to show how tall someone is, gesturing towards an empty seat, are all examples of communicative gestures that can serve as a 'quasi-universal' language (Marsh et al., 2007). Comparison of results obtained in different linguistic contexts and cultures, however, have so far been hindered by lack of evaluation instruments adapted and validated in different languages.

In the present work, we describe the *multilingual CID-5*, a new tool being made available in seven languages for the investigation of how non-conventional communicative gestures are classified and identified in different linguistic and cultural contexts. In the following, we present first a brief background on the point-light technique used to create stimuli included in the multilingual CID-5, and in the original CID-5 database. Next, we provide a detailed description of multilingual CID-5 database, including all the materials available for download. Finally, we present normative data collected to assess the stimulus classification (communicative vs. individual) and intention identification by speakers of seven different languages, namely Chinese, Dutch, English, French, German, Italian, and Polish.

Gestural Communication through Point-light Displays

The point-light technique is a method for representing biological motion through limited visual information (Johansson, 1973). With this method, the movements of a body of a living being are represented by a small number of point lights indicating the major joints of a person performing a given action. Despite the absence of other cues such as contour, color, or texture, observers can quite easily identify what an actor is doing (e.g., Vanrie and Verfaillie, 2004), as well as many features of the actor themselves, including identity (e.g., Loula et al., 2005), gender (e.g., Kozłowski and Cutting, 1977; Pollick et al., 2005; Brooks et al., 2008), age (Montpare and Zebrowitz-McArthur, 1988), emotional state (e.g., Pollick et al., 2001; Atkinson et al., 2004; Clarke et al., 2005), and personality traits (Heberlein et al., 2004).

Given this keen sensitivity to action motion signatures, it is reasonable to expect that people are also able to discern communicative gestures from point-light displays. Along these lines, recent evidence suggests that biological motion information is sufficient for clear classification of a non-conventional action as communicative, as well as for the identification of the

specific communicative intent (Manera et al., 2010, 2011a,b, 2013; Centelles et al., 2013). Furthermore, Manera et al. (2011a, 2013) demonstrated that in the context of a communicative interaction between two point-light agents, observing the communicative gesture of one agent enhances the visual discrimination of a second agent who responds appropriately.

The generalizability of these findings across different cultures and linguistic communities, however, is until now poorly documented. There is evidence that biological motion perception is not necessarily influenced by culture, and that point-light stimuli reproducing simple and putatively universal human actions, such as walking, can be recognized even by indigene populations of Amazonian territories (Pica et al., 2011; see also Barrett et al., 2005), as well as by newborns (Simion et al., 2008). It remains possible, however, that cultural tendencies to display particular non-conventional gestures in certain styles influence intention-from-movement judgments, and that speakers of different languages may classify and describe the same actions differently.

The CID-5 Database

The CID-5 database (Communicative Interaction Database, Five Alternative Forced Choice format, 5AFC) contains 21 full-body point-light stimuli depicting two agents (A and B) engaged either in communicative interactions ($N = 14$) or non-communicative individual actions ($N = 7$) as seen from four different viewpoints. Following Dekeyser et al. (2002), stimuli were constructed by combining motion capture techniques and 3-D animation software to provide precise control over the computer-generated actions and allow the actions of the two agents to be independently manipulated. For each action stimulus, the CID-5 provides (i) coordinate files for each actor; (ii) movie files depicting the action of the two agents as seen from four different perspectives; (iii) five action alternatives describing the action performed by the two agents. The CID-5 database can be freely downloaded from <http://bsb-lab.org/research/>.

Results collected on a sample of 113 Italian speaking participants using these stimuli confirmed that naive observers are able to distinguish communicative and individual gestures, and to identify the correct action description among the five alternatives (Manera et al., 2015). The *multilingual CID-5* extends the CID-5 by providing a translation of the response alternatives into seven different languages: Chinese, Dutch, English, French, German, Italian, and Polish. Furthermore, it provides some normative data to validate the alternative action descriptions in the different languages.

THE MULTILINGUAL CID-5 DATABASE

Building on the CID-5 database, the *multilingual CID-5* database provides a new tool to investigate classification and identification of non-conventional communicative gestures by speakers of different languages. The database is available as Supplementary Material to this article, or from the website of the Biology of Social Behavior Lab, University of Torino (<http://bsb-lab.org/research/>).

Actions

A brief description of each action stimulus is reported in **Table 1**. Stimuli consist of the 21 point-light actions depicting two point-light agents, each consisting of 13 markers indicating head, shoulders, elbows, wrists, hips, knees, and feet. For each stimulus, we report the stimulus classification (communicative vs. individual), a brief description of the actions of agent A and agent B, and the actors' gender.

Stimuli were originally constructed by capturing the movements of four actors, two Italian females and two Dutch males, each wearing 30 reflective spherical markers (Qualisys MacReflex motion capture system; Qualisys; Gothenburg, Sweden, consisting of six 30-Hz position units). For the communicative interactions, the two female and the two male actors worked in pair (a couple of male actors and a couple of female actresses) and were assigned to a 'communicator' and 'responder' role. The communicator (agent A) always initiated the interaction by performing a communicative gesture; the responder (agent B) perceived the communicative gesture and acted in response, based on a predefined interaction plot. To ensure that the responder's action matched the communicator's gesture in all respects (e.g., timing, position, kinematics), interactions were captured in real time, with the actors facing each other, at a distance of approximately 2 m. Individual actions were performed by agent A acting in isolation. Objects (e.g., table, chair, coins, fruits) were present during the production of actions to aid the actors in producing natural movements.

After the capture session, the 2-D data from all the position units were processed offline to calculate the 3-D coordinates of the markers. Missing data points (less than 5%) were filled in manually using customized functions of the Fluey 2 motion

toolkit (MTK, Televirtual). The data from the markers were then imported into Character Studio (Autodesk Inc, 1998). This allowed to animate a biped for each actor, consisting of a transparent skeleton and 13 bright dots attached to the center of the major joints (shoulders, elbows, wrists, hips, knees, and ankles) and the head. To create the actual movie files, the smoothed data were imported into 3-D Studio as moving bright spheres, and all the frames of the action were rendered as avi-files from four different viewpoints. Some manual smoothing was performed to avoid any remaining "jumpy" dot movements. An orthographic projection was used, and there was no occlusion, so no explicit depth cues were available. To create the communicative action stimuli avi-files, data from the two actors of each couple were imported into the same 3D studio environment, making sure that the actors were exactly at the same distance as in the original recording session. To create the individual action stimuli.avi files, the communicator's gesture was substituted with an individual action performed by the same actor, making sure to match stimulus duration. Objects present in the scene during motion capturing were never visible in any of the point-light displays.

Response Alternatives

The 'Response Alternatives' folder contains seven.doc files (Supplementary Data Sheet S1) reporting the list of the five response alternatives for each action stimulus in seven different languages (Chinese, Dutch, English, French, German, Italian, and Polish).

The five alternatives included the correct action description and four incorrect response alternatives. The incorrect response alternatives were generated according to the following criteria.

TABLE 1 | Description of the actions included in the CID-5 database.

Action	Action (sequence) description	Communicative vs. individual	Male/female couple
Choose which one	A asks B to choose between two objects; B takes an object.	Communicative	Male
Come closer	A asks B to come closer; B moves forward.	Communicative	Female
Go out of the way	A asks B to go out of the way; B moves over.	Communicative	Male
Imitate me	A squats down and asks B to imitate him; B squats down.	Communicative	Male
Look at the ceiling	A asks B to look at something behind him on the ceiling; B turns around.	Communicative	Male
Look at the ground	A asks B to look at something on the ground; B squats down.	Communicative	Male
Move this down	A asks B to move something down; B picks something and moves it down.	Communicative	Female
No	A says no; B, who had grasped something, puts that down.	Communicative	Male
Pick this up	A points to B something to pick up; B picks something up.	Communicative	Female
Sit down	A asks B to sit down; B sits down.	Communicative	Male
Squat down	A asks B to squat down; B squats down.	Communicative	Female
Stand up	A asks B to stand up; B, who is sitting, stands up.	Communicative	Female
Stop	A asks B to stop; B, who is walking, stops.	Communicative	Male
Walk away	A asks B to walk away; B takes some steps into the indicated direction.	Communicative	Male
Drink	A drinks; B sits down.	Individual	Male
Jump	A jumps; B picks something up.	Individual	Female
Lateral steps	A makes some lateral steps; B takes something and eats it.	Individual	Male
Look under the foot	A looks under his foot; B moves something.	Individual	Female
Sneeze	A sneezes; B turns around.	Individual	Male
Stretch	A stretches; B moves something.	Individual	Female
Turn over	A turns over; B squats down.	Individual	Female

For each action stimulus (e.g., A asks B to walk away), two incorrect communicative alternatives (e.g., A opens the door for B; A asks B to move something) and two incorrect non-communicative alternatives (A stretches; A draws a line) were generated by modifying the description of the action of agent A. All alternative action descriptions were constructed to be physically compatible with the action performed by agent A. For instance, if agent A performed an arm movement, then reference to arm movement was included in all incorrect response alternatives describing the action stimulus. Finally, to avoid that for communicative stimuli the correct alternative was selected simply based on the congruence between the actions of the two agents (i.e., agent A asks B to perform an action, and agent B responds *accordingly*), for each action stimulus, one of the incorrect communicative alternatives always described a congruent interaction between the two agents (see **Supplementary Table S1**). The description of the action of agent B was the same for all response alternatives.

Translation of the Alternatives

Translations in each language were performed by two independent native speakers. Translators were provided with the English version of the alternatives, and the original point-light movie files. The two translations were then compared, and in case of discrepancies, the translators were asked to decide together which description matched better the English version of the alternative and the corresponding point-light video.

COLLECTION OF PRELIMINARY DATA

Participants

One hundred and forty healthy volunteers (61 male, 79 female; age, $M = 24.9$, $SD = 4.6$, years of education, $M = 15.8$, $SD = 2.2$) took part in this study, 20 for each of the following languages: Chinese, Dutch, English, French, German, Italian, and Polish. Participants were recruited at the University and Polytechnic of Torino, in Italy (Chinese and Italian speakers), at the Katholieke Universiteit Leuven, in Belgium (Dutch speakers), at the University of California at Davis, in the US (English speakers), at the University Hospital Cologne, in Germany (German speakers), at the University of Nice Sophia Antipolis and the Nice University Hospitals, in France (French speakers), and at the University of Warsaw, in Poland (Polish speakers). They received course credits or payment for their participation. Demographic characteristics of the participants of each country are reported in **Table 2** and **Figure 1**. All participants had normal or corrected to normal vision, and were naive as to the purpose of the study. The study was approved by the local ethical committees.

Stimuli and Procedure

Twenty-one point-light actions taken from CID-5 database were employed (Manera et al., 2015), including 14 communicative interactions in which the two agents (A and B) were engaged

TABLE 2 | Participant's demographics.

Native language	Gender <i>N</i> of female (in a sample of 20)	Age, years Mean (SD)	Education (years) Mean (SD)
Chinese	11	24.4 (3.5)	14.5 (1.7)
Dutch	10	25.0 (2.2)	15.6 (1.7)
English	9	21.0 (2.5)	14.9 (1.5)
German	12	26.8 (5.6)	17.3 (1.8)
French	12	29.8 (5.7)	17.6 (2.3)
Italian	10	25.5 (2.6)	16.8 (1.6)
Polish	15	22.1 (2.7)	14.3 (2.1)

in a communicative interaction (e.g., agent A points out at the ceiling, agent B looks at the ceiling) and 7 non-communicative individual actions, in which A and B were acting independently of each other (e.g., A drinks, B sits down). Agent A was positioned on the right side of the screen, and agent B was positioned on the left side of the screen (corresponding to the 125° perspective in the CID-5 database; see the description of the video perspectives reported in Manera et al., 2010) in all the action stimuli. The two agents were displayed simultaneously, with action of agent B (the responder in the communicative stimuli) always following in time the action of agent A (the communicator in the communicative stimuli). Stimuli were presented in a randomized order. Following the procedure used in previous reports (e.g., Manera et al., 2015), each video with the two agents was shown twice consecutively, with the two videos separated by a 500 ms fixation cross. After the second presentation of each video, participants were, firstly, asked to decide whether the two agents were communicating vs. acting independently of each other (intention classification). The question was displayed on

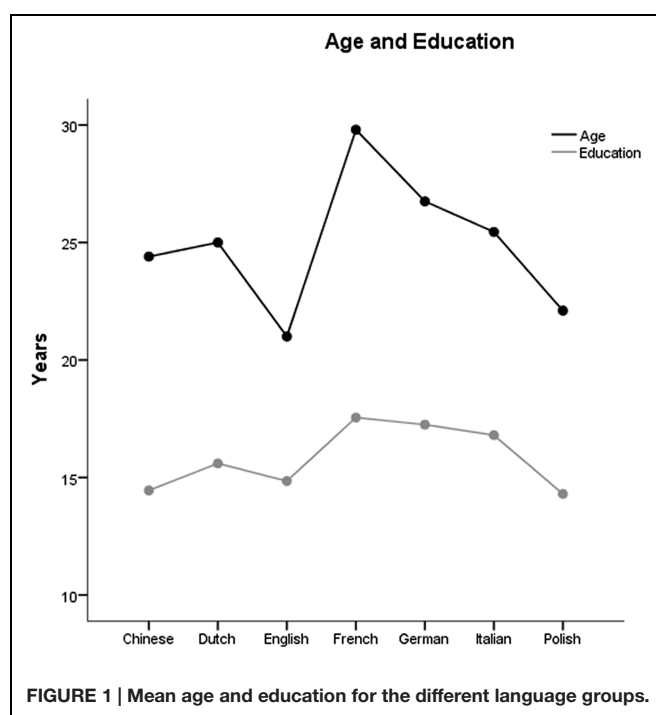


FIGURE 1 | Mean age and education for the different language groups.

the screen until a response was provided. Secondly, participants were asked to select the correct action description among five numbered response alternatives displayed simultaneously (intention identification). The order of the response alternatives was randomized across stimuli. The question was presented on the screen until response, with no time restriction. No feedback concerning response correctness was given to the participants. Depending on the sample, instructions, questions, and response alternatives were presented in Chinese, Dutch, English, French, German, Italian, or Polish.

The Chinese and Polish versions of the procedure were created with E-prime software (Psychology Software Tool, Inc), while the Dutch, English, French, German, and Italian versions of the procedure were created with Presentation software (Neurobehavioral Systems, Inc.) In all language samples, stimuli were displayed on a 14" to 17" LCD screen. The task took approximately 15–20 min to complete.

Results

Demographics

Chi Square analysis revealed no gender differences among the seven samples corresponding to the different languages ($\chi^2 = 4.76$, $p = 0.574$). A between-subject ANOVA with age as dependent variable and language (Chinese, Dutch, English, French, German, Italian, and Polish) as between subject factor revealed a significant difference in age among the different language groups [$F_{(6,133)} = 11.52$, $p < 0.001$]. *Post hoc* comparisons (Bonferroni corrected) revealed that French-speaking participants were significantly older than Chinese-speaking participants ($p < 0.001$), Dutch-speaking participants ($p = 0.003$), English-speaking participants ($p < 0.001$), Italian-speaking participants ($p = 0.010$) and Polish-speaking participants ($p < 0.001$). Dutch-speaking participants were significantly older than English-speaking participants ($p = 0.027$). English-speaking participants were younger than German-speaking participants ($p < 0.001$), and Italian-speaking participants ($p = 0.008$). Finally, German-speaking participants were significantly older than Polish-speaking participants ($p = 0.004$).

A between-subject ANOVA with education as dependent variable and language (Chinese, Dutch, English, French, German, Italian, and Polish) as between subject factor revealed a significant difference in education between the different language groups [$F_{(6,133)} = 11.00$, $p < 0.001$]. *Post hoc* comparisons (Bonferroni corrected) revealed that French-speaking participants had more education years compared to Chinese-speaking participants ($p < 0.001$), Dutch-speaking participants ($p = 0.022$), English-speaking participants ($p < 0.001$), and Polish-speaking participants ($p < 0.001$). Chinese-speaking participants had fewer education years compared to German-speaking participants ($p < 0.001$). English-speaking participants had fewer education years compared to German-speaking participants ($p = 0.001$) and Italian-speaking participants ($p = 0.022$). Finally, Polish speaking participants had fewer education years

compared to German-speaking participants ($p < 0.001$) and Italian-speaking participants ($p = 0.001$).

As age and education differed among the seven language-samples, they were added as covariates in all the between-subject analyses.

Multilingual CID-5 Task

Separate analyses were conducted to evaluate global performance and recognizability of single stimuli.

Data analysis

Global performance. To evaluate global performance, for each language we calculated the percentage of participants who correctly responded to the intention classification and the intention identification questions, and we extracted Signal Detection Theory (SDT) parameters.

To evaluate how participants distinguished between communicative and individual action stimuli (intention classification), we calculated sensitivity (d') and criterion (c'). For each participant, we calculated the proportion of hits (arbitrarily defined as “communicative” responses when the action stimulus was communicative) and false alarms (“communicative” responses when the action stimulus was individual). Proportions of 0 were replaced with $0.5/N$, and proportions of 1 were replaced with $(N-0.5)/N$ (where N is the number of communicative and individual stimuli). d' and c were then submitted to single sample t -tests (test value = 0) to ascertain whether discrimination performance was above chance level, and to verify the presence of any systematic response bias. Furthermore, to ascertain whether d' and c varied across languages, they were submitted to separate ANCOVAs with Language (Chinese, Dutch, English, French, German, Italian, and Polish) and Gender (Male vs. Female) as between-subject factors, and Age and Education as covariates. Finally, in order to verify the presence of interactions between participants' gender and the gender of the actors in the ability to classify the actions as communicative vs. individual, the d' was submitted to a repeated measures ANOVA with Actor gender as within-subject factor, and Gender as between-subject factor.

To evaluate global performance on the intention identification question, we first recodified responses as communicative vs. individual to calculate sensitivity (d') and criterion (c). d' and c were submitted to single sample t -tests (test value = 0) to ascertain whether discrimination performance was above chance level, and to verify the presence of any systematic response bias. Second, to evaluate the ability to select the correct response alternative, following the standard SDT approach to mAFC (e.g., Macmillan and Creelman, 2005), we used the proportion correct responses as a measure of sensitivity. To compare performance across different languages, we submitted the mean proportion of correct responses to a repeated measures ANCOVA with Intention (Communicative vs. Individual) as within-subject factor, Language (Chinese, Dutch, English, French, German, Italian, and Polish) and Gender (Male vs. Female) as between-subject factors, and Age and Education as covariates. Finally, in order to verify the presence of interactions between participants'

gender and the gender of the actors in the intention identification ability, the proportion of correct responses was submitted to a repeated measures ANOVA with Actor gender as within-subject factor and Gender as between-subject factor.

Stimulus recognizability. To provide researchers with detailed data on the classification and identification of single stimuli across languages, for each action stimulus, we first calculated whether the proportion of correct responses differed from chance level – that is, from 0.5 for question 1 (corresponding to 50% of correct responses) and 0.2 for question 2 (corresponding to 20% of correct responses) – by employing binomial tests. Bonferroni corrections were applied to adjust for multiple comparisons ($\alpha = 0.05/21 = 0.0023$). Second, we verified whether the distribution of correct responses (0 for incorrect response, 1 for correct response) and the distribution of errors (communicative alternative 1, communicative alternative 2, individual alternative 1, and individual alternative 2) varied depending on the factor Language (Chinese, Dutch, English, French, German, Italian, and Polish) by means of Chi-square analyses (see **Supplementary Table S1**). Bonferroni corrections were applied ($p < 0.0023$).

Global performance: results

Descriptive statistics (mean and SD) for the percentage of correct responses for the intention classification and the intention identification questions for each language are reported in **Table 3**.

Intention classification. Sensitivity (d') for the full sample ($N = 140$) ranged from 0.27 to 3.23 ($M = 2.46$, $SD = 0.57$; see **Figure 2A**), and was significantly higher than zero [$t_{(139)} = 51.15$, $p < 0.001$], thus suggesting that participants, as a group, were able to discriminate communicative from individual action stimuli well above the chance level. The ANCOVA on d' with Language and Gender as between subject factors, and Age and Education as covariates revealed no statistically significant effects [corrected model, $F_{(15,124)} = 0.81$, $p = 0.664$]. Specifically, no significant effect of Language [$F_{(6,124)} = 0.83$, $p = 0.546$], Gender [$F_{(1,124)} = 0.19$, $p = 0.662$], Age [$F_{(1,124)} = 3.48$, $p = 0.064$] or Education [$F_{(1,124)} = 0.23$, $p = 0.633$] was found. Criterion c for the full sample ($N = 140$) ranged from -1.61 to 0.65 ($M = -0.15$, $SD = 0.36$; see **Figure 2B**), and was significantly lower than zero [$t_{(139)} = -4.84$, $p < 0.001$] thus suggesting that participants, as a group, had a tendency to rate stimuli as

communicative. The ANCOVA with Language and Gender as between subject factors, and Age and Education as covariates was statistically significant [corrected model, $F_{(15,124)} = 2.95$, $p < 0.001$]. Specifically, a main effect of Language was found [$F_{(6,124)} = 4.75$, $p < 0.001$]. *Post hoc* comparisons revealed that Chinese-speaking participants had a stronger tendency to rate stimuli as communicative compared to French-speaking ($p = 0.014$) and Italian-speaking participants ($p = 0.007$). Similarly, Polish-speaking participants had a stronger tendency to rate stimuli as communicative compared to French-speaking ($p = 0.004$) and Italian-speaking participants ($p = 0.001$). No significant effect of Gender [$F_{(1,124)} = 3.08$, $p = 0.082$], Age [$F_{(1,124)} = 1.54$, $p = 0.218$] or Education [$F_{(1,124)} = 0.03$, $p = 0.864$] on c was found.

The repeated measures ANOVA on d' with Actor gender as within-subject factor and Gender as between-subject factor revealed a significant main effect of Actor Gender [$F_{(1,138)} = 6.70$, $p = 0.011$], with the proportion of correct responses being significantly higher for the female actresses ($M = 2.56$) compared to the male actors ($M = 2.33$). No effect of Gender [$F_{(1,138)} = 0.92$, $p = 0.340$] and no significant interaction between Actor Gender and Gender [$F_{(1,138)} = 0.06$, $p = 0.815$] was found.

Intention identification. The d' calculated on the full sample ($N = 140$) after re-codifying the responses as communicative vs. individual ranged from 0.40 to 3.27 ($M = 2.60$, $SD = 0.53$), and was significantly higher than zero [$t_{(139)} = 58.18$, $p < 0.001$], thus suggesting that, also for the action identification question, participants were able to discriminate communicative from individual action stimuli well above the chance level. Criterion c calculated on the full sample ranged from -1.27 to 0.65 ($M = -0.04$, $SD = 0.28$), and was not significantly different from zero [$t_{(139)} = -1.56$, $p = 0.122$], thus suggesting that participants, contrary to what happened in the intention classification question, when asked to select the correct action description among several action alternatives, showed no response bias toward a communicative response.

The proportion of correct response alternatives for each language is reported in **Figure 3**. The repeated measures ANCOVA with Intention as within-subject factor, Language and Gender as between subject factors and Age and Education as covariates revealed no significant effect of Intention [$F_{(1,124)} = 0.76$, $p = 0.384$], Gender [$F_{(1,24)} = 0.27$, $p = 0.606$], Age [$F_{(1,124)} = 2.21$, $p = 0.139$], or Education [$F_{(1,124)} = 2.13$, $p = 0.147$] on the proportion of correct responses. However, a significant effect of Language was found [$F_{(6,124)} = 2.71$, $p = 0.017$]. *Post hoc* comparisons revealed that French-speaking participants performed significantly better compared to the Dutch-speaking participants ($p = 0.026$). No two-way or three-way interaction reached statistical significance (all $ps > 0.056$).

The repeated measures ANOVA with Actor gender as within-subject factor and Gender as between-subject factor revealed a significant main effect of Actor Gender [$F_{(1,138)} = 70.40$, $p < 0.001$], with the proportion of correct responses being significantly higher for the female actresses ($M = 0.83$) compared to the male actors ($M = 0.71$). No effect of Gender

TABLE 3 | Percentage of correct responses for the intention classification and identification questions.

Language	Intention classification Mean (SD)	Intention identification Mean (SD)
Chinese	88% (6%)	71% (7%)
Dutch	90% (6%)	72% (12%)
English	92% (4%)	78% (10%)
German	90% (9%)	80% (9%)
French	90% (8%)	81% (8%)
Italian	89% (8%)	77% (11%)
Polish	89% (9%)	74% (12%)

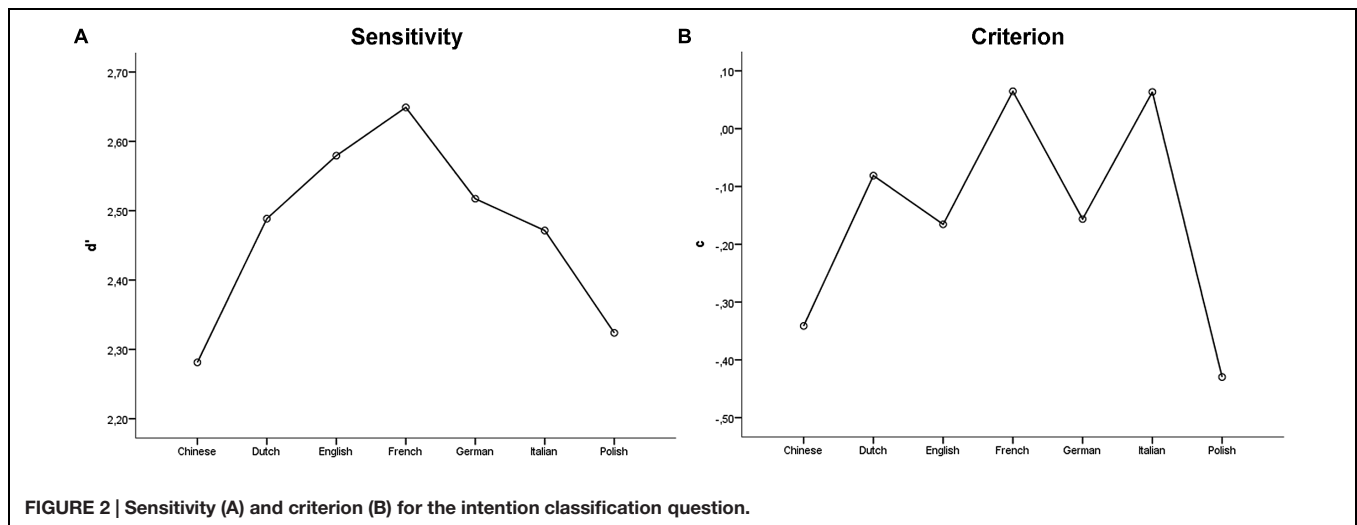


FIGURE 2 | Sensitivity (A) and criterion (B) for the intention classification question.

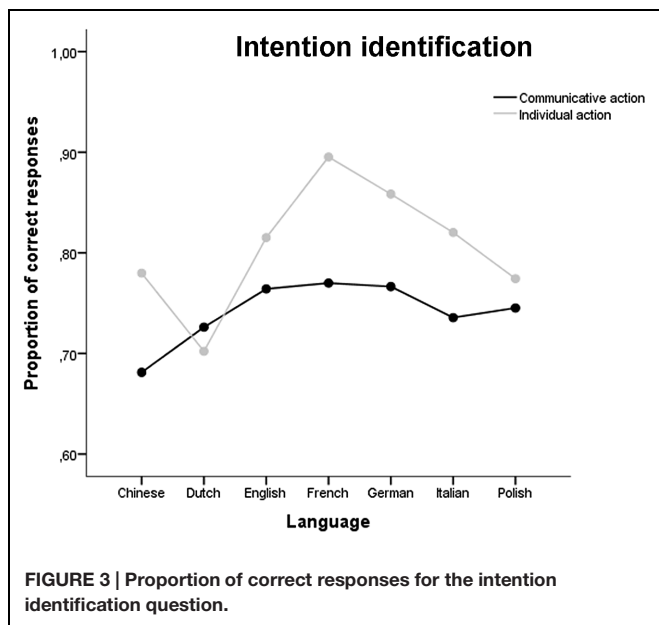


FIGURE 3 | Proportion of correct responses for the intention identification question.

$[F_{(1,138)} = 0.19, p = 0.664]$ and no significant interaction between Actor Gender and Gender $[F_{(1,138)} = 1.07, p = 0.302]$ was found.

Stimulus recognizability: results

For each stimulus, the percentage of participants who correctly responded to the classification question, and the percentage of participants who reported each of the alternatives in the identification question are reported in **Supplementary Table S1**.

Intention classification. On average, participants correctly classified 90% of the action stimuli as communicative vs. individual (range = 72–99%; $SD = 8\%$; Communicative stimuli, $M = 91\%$, $SD = 9\%$; Individual stimuli, $M = 87\%$, $SD = 8\%$). The actions that were less consistently recognized were “Look at the ceiling” for the communicative condition (correctly classified as communicative by 74% of the

participants) and “Sneeze” for the individual condition (correctly classified as individual by 72% of the participants). Bonferroni corrected binomial tests conducted on the full sample revealed that action classification was above chance level (proportion of correct responses of 0.50) for all the action stimuli (all $ps < 0.001$). Chi-square tests performed on the single action stimuli (intention classification \times Language) revealed that intention classification did not differ between languages for any of the 21 actions (see **Supplementary Table S1**).

Intention identification. On average, participants correctly described 76% of the action stimuli (range = 36–96%; $SD = 17\%$; Communicative stimuli, $M = 74\%$, $SD = 18\%$; Individual stimuli, $M = 80\%$, $SD = 14\%$). Examples of very well recognized stimuli are “Stop” and “Imitate me” for the communicative stimuli, and “Jump” and “Look under the foot” for the individual stimuli. Bonferroni corrected binomial tests conducted on the full sample revealed that action identification was above chance level (proportion of correct responses of 0.20) for all the action stimuli (all $ps < 0.001$). Bonferroni corrected Chi-square tests performed on the single action stimuli (intention identification \times Language) revealed that intention identification varied by Language only for the following two actions: “Go out of the way” ($p < 0.001$), and “No” ($p < 0.001$; see **Supplementary Table S1**). Bonferroni corrected Chi-square tests performed on the errors revealed no significant effect of Language on any of the action stimuli (all $ps > 0.018$), thus suggesting that, when the intention identification was incorrect, participants in the different language samples tended to select the same wrong action alternatives. Some response alternatives were thus more misleading than others in all languages.

DISCUSSION

In the present paper we describe the *multilingual CID-5*, a database of 21 full-body point-light stimuli depicting two agents

engaged in communicative interactions ($N = 14$) or performing non-communicative individual actions ($N = 7$) as seen from different viewpoints. For each stimulus, we provide five plausible response alternatives (only one being correct) translated into seven different languages (Chinese, Dutch, English, French, German, Italian, and Polish). Normative data collected from 140 naive participants (20 participants per language) confirmed that all the stimuli included in the multilingual CID-5 were classified as communicative vs. individual and recognized well above chance level from participants of all the seven language samples. Comparisons of global performance across different languages revealed no difference across samples in the ability to classify actions as communicative vs. individual, as indexed by the SDT parameter d' calculated on the action classification question. Similarly, analyses on the proportion of correct responses divided by action stimulus revealed that all the 21 action stimuli were classified as communicative vs. individual in a comparable way in all language samples, with some actions being consistently very easy (e.g., 'Stop') and some others more difficult (e.g., 'Look at the ceiling') to classify. Overall, in the intention classification question participants showed a liberal criterion (negative c), that is a bias towards reporting the presence of a communicative interaction. This bias may be partially explained by the presence of a greater number of communicative action stimuli. The response bias showed a significant variation across language samples, and was especially evident in the Chinese-speaking and Polish-speaking samples. However, no bias towards reporting a communicative response alternative was found in the intention identification question, when participants were asked to select among multiple response alternatives. Thus, researchers interested in an unbiased measure of the ability to classify stimuli as communicative vs. individual may decide to rely on the intention identification question, after re-coding the response alternatives as communicative vs. individual.

For intention identification (selection of the correct response alternative), we found some individual variations in the proportion of correct responses across language samples, with French-speaking participants performing better compared to Dutch-speaking participants. Analyses divided by action stimulus revealed that 19 out of the 21 stimuli were identified in a comparable way across languages, while only two stimuli ("Go out of the way" and "No") showed language-dependent variations. Furthermore, the error analysis showed that when the intention identification was incorrect, participants in the different language samples tended to select the same wrong alternative, suggesting that, for most of action stimuli, some response alternatives were more misleading than others in all languages.

REFERENCES

- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., and Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 717–746. doi: 10.1068/p5096
- Autodesk Inc (1998). *Character Studio Max (Release 2) [Computer Software]*. Sausalito, CA: Author.
- Barrett, H. C., Todd, P. M., Miller, G. F., and Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: a cross-cultural study. *Evol. Hum. Behav.* 26, 313–331. doi: 10.1016/j.evolhumbehav.2004.08.015
- Brooks, A., Schouten, B., Troje, N., Verfaillie, K., Blanke, O., and van der Zwan, R. (2008). Correlated changes in perceptions of the gender and orientation of ambiguous biological motion figures. *Curr. Biol.* 18, R728–R729. doi: 10.1016/j.cub.2008.06.054
- These results provide evidence of instrument validity of the multilingual CID-5 as a new tool for the investigation of non-conventional communicative gestures in different languages. It is important to note that our data collection was designed to validate the alternatives in the different languages, and not to explore systematically cultural differences. Thus, from the present results, we cannot conclude that classification and identification of communicative gestures does not vary across cultures. First, participants in the different language samples were not balanced for age and education. Second, Chinese college students were tested in Italy and experienced thus mostly the same environment as Italian participants – a circumstance which might have well influenced their familiarity with the stimulus material. Third, and more importantly, the selected sample groups were not very distant in terms of shared cultural heritage. Future investigations should therefore remain open to the possibilities of systematic difference in nonverbal behavior across distant cultures. A final limitation relates to potential differences in social cognition and visuo-spatial abilities, not assessed in the present study. As individual differences in these abilities have been shown to correlate with recognition of social information from point-light stimuli (Okruszek et al., 2015), taking these variables into account may help to clarify the true nature of cross-linguistic and cross-cultural differences, if any, in intention-from-motion understanding.

FUNDING

This work received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n. 312919.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01724>

TABLE S1 | Percentage of participants who correctly responded to the classification question, and who reported each of the alternatives in the identification question. The "Com vs. Ind" column indicates the percentage of correct responses to the intention classification question (classification of the action as communicative vs. individual). The column "Action" indicates the percentage of responses provided for each of the five response alternatives. The first action alternative (in bold) reports the correct description. The column "Chi Square" reports the Chi Square values calculated on the proportion of correct vs. incorrect responses (intention classification and intention identification by Language). Values indicated in bold are statistically significant ($***p < 0.001$).

- Casasanto, D. (2013). "Gesture and language processing," in *Encyclopedia of the Mind*, ed. H. Pashler (Thousand Oaks, CA: Sage Publications), 372–374.
- Centelles, L., Assaiante, C., Etchegoyhen, K., Bouvard, M., and Schmitz, C. (2013). From action to interaction: exploring the contribution of body motion cues to social understanding in typical development and in autism spectrum disorders. *J. Autism Dev. Disord.* 43, 1140–1150. doi: 10.1007/s10803-012-1655-0
- Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., and Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception* 34, 1171–1180. doi: 10.1068/p5203
- Dekeyser, M., Verfaillie, K., and Vanrie, J. (2002). Creating stimuli for the study of biological-motion perception. *Behav. Res. Methods Instrum. Comput.* 34, 375–382. doi: 10.3758/BF03195465
- Heberlein, A. S., Adolphs, R., Tranel, D., and Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. *J. Cogn. Neurosci.* 16, 1143–1158. doi: 10.1162/0898929041920423
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 201–211. doi: 10.3758/BF03212378
- Kozlowski, L. T., and Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Percept. Psychophys.* 21, 575–580. doi: 10.3758/BF03198740
- Loula, F., Prasad, S., Harber, K., and Shiffrar, M. (2005). Recognizing people from their movement. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 210–220. doi: 10.1037/0096-1523.31.1.210
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd Edn. Mahwah, NJ: Lawrence Erlbaum Associates.
- Manera, V., Becchio, C., Schouten, B., Bara, B. G., and Verfaillie, K. (2011a). Communicative interactions improve visual detection of biological motion. *PLoS ONE* 6:e14594. doi: 10.1371/journal.pone.0014594
- Manera, V., Del Giudice, M., Bara, B. G., Verfaillie, K., and Becchio, C. (2011b). The second agent effect: communicative gestures increase the likelihood of perceiving a second agent. *PLoS ONE* 6:e22650. doi: 10.1371/journal.pone.0022650
- Manera, V., Schouten, B., Becchio, C., Bara, B. G., and Verfaillie, K. (2010). Inferring intentions from biological motion: a stimulus set of point-light communicative interactions. *Behav. Res. Methods* 42, 168–178. doi: 10.3758/BRM.42.1.168
- Manera, V., Schouten, B., Verfaillie, K., and Becchio, C. (2013). Time will show: real time predictions during interpersonal action perception. *PLoS ONE* 8:e54949. doi: 10.1371/journal.pone.0054949
- Manera, V., von der Luhe, T., Schilbach, L., Verfaillie, K., and Becchio, C. (2015). Communicative interactions in point-light displays: choosing among multiple response alternatives. *Behav. Res. Methods* doi: 10.3758/s13428-015-0669-x [Epub ahead of print].
- Marsh, A. A., Elfenbein, H. A., and Ambady, N. (2007). Separated by a common language: nonverbal accents and cultural stereotypes about Americans and Australians. *J. Cross Cult. Psychol.* 38, 284–301. doi: 10.1177/0022022107300275
- Montpare, J. M., and Zebrowitz-McArthur, L. (1988). Impressions of people created by age-related qualities of their gaits. *J. Pers. Soc. Psychol.* 55, 547–556. doi: 10.1037/0022-3514.55.4.547
- Okruszek, Ł., Haman, M., Kalinowski, K., Talarowska, M., Becchio, C., and Manera, V. (2015). Impaired recognition of communicative interactions from biological motion in schizophrenia. *PLoS ONE* 10:e0116793. doi: 10.1371/journal.pone.0116793
- Pica, P., Jackson, S., Blake, R., and Troje, N. F. (2011). Comparing biological motion perception in two distinct human societies. *PLoS ONE* 6:e28391. doi: 10.1371/journal.pone.0028391
- Pollick, F. E., Kay, J. W., Heim, K., and Stringer, R. (2005). Gender recognition from point-light walkers. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 1247–1265. doi: 10.1037/0096-1523.31.6.1247
- Pollick, F. E., Paterson, H. M., Bruderlin, A., and Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition* 82, B51–B61. doi: 10.1016/S0010-0277(01)00147-0
- Simion, F., Regolin, L., and Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proc. Natl. Acad. Sci. U.S.A.* 105, 809–813. doi: 10.1073/pnas.0707021105
- Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Vanrie, J., and Verfaillie, K. (2004). Perception of biological motion: a stimulus set of human point-light actions. *Behav. Res. Methods Instrum. Comput.* 36, 625–629. doi: 10.3758/BF03206542

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Manera, Iani, Bourgeois, Haman, Okruszek, Rivera, Robert, Schilbach, Sievers, Verfaillie, Vogeley, von der Lühe, Willems and Becchio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Fifteen-month-old infants use velocity information to predict others' action targets

Janny C. Stapel*, Sabine Hunnius and Harold Bekkering

Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Evin Aktar,
University of Amsterdam, Netherlands
Anne Scheel,
Ludwig Maximilian University
of Munich, Germany
Laura Sparaci,
Institute of Cognitive Sciences
and Technologies, Italy

*Correspondence:

Janny C. Stapel,
Donders Institute for Brain, Cognition
and Behaviour, Radboud University
Nijmegen, P.O. Box 9104, 6500 HE
Nijmegen, Netherlands
j.c.stapel@donders.ru.nl

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 30 January 2015

Accepted: 16 July 2015

Published: 04 August 2015

Citation:

Stapel JC, Hunnius S
and Bekkering H (2015)
Fifteen-month-old infants use velocity
information to predict others' action
targets.
Front. Psychol. 6:1092.
doi: 10.3389/fpsyg.2015.01092

In a world full of objects, predicting which object a person is going to grasp is not easy for an onlooker. Among other cues, the characteristics of a reaching movement might be informative for predicting its target, as approach movements are slower when more accuracy is required. The current study examined whether observers can predict the target of an action based on the movement velocity while the action is still unfolding, and if so, whether these predictions are likely the result of motor simulation. We investigated the role of motor processes for velocity-based predictions by studying participants who based on their age differed in motor experience with the task at hand, namely reaching. To that end, 9-, 12-, and 15-month-old infants and a group of adults participated in an eye-tracking experiment which assessed action prediction accuracy. Participants observed a hand repeatedly moving toward and pressing a button on a panel, one of which was small, the other one large. The velocity of the reaching hand was the central cue for predicting which button would be the target of the observed action as the velocity was lower when reaching for the small compared to the large button. Adults and 15-month-old infants made more frequent visual anticipations to the close button when it was the target than when it was not and were thus able to use the information in the speed of the approach movement for the prediction of the action target. The 9- and 12-month-olds, however, did not display this difference. After the eye-tracking experiment, infants' ability to aim for and press buttons of different sizes was evaluated. Results showed that the 15-month-olds were more proficient than the 9- and 12-month-olds in performing the reaching actions. The developmental time line of velocity-based action predictions thus corresponds to the development of performing that motor act yourself. Taken together, these data suggest that motor simulation may underlie velocity-based predictions.

Keywords: action prediction, infancy, speed-accuracy trade-off, motor system, predictive eye-movements

Introduction

Predicting others' actions is crucial for social interactions to run smoothly (Bekkering et al., 2009; Sebanz and Knoblich, 2009). Anticipating which goal object an action partner will grasp, however, is complicated in a world full of objects. How do observers predict which object another person is reaching for? And how does the ability to predict others' actions develop early in life? Motor theories of action perception suggest that the motor system is used to predict others' actions

the same way it is used to predict the outcomes of one's own motor acts (Wolpert et al., 2003; Oztop et al., 2005; Prinz, 2006; Kilner et al., 2007a). In accordance with this notion, a large body of literature shows that the motor system is not only active during action execution but also during the observation of others' actions (Rizzolatti et al., 1996; Hari et al., 1998; Cattaneo et al., 2007), suggesting that similar processes are at work during observation and execution. Consequently, laws governing action production can be expected to also affect action perception. One of these laws is Fitts's law (Fitts, 1954), which describes that actions directed at small targets require more time to perform. Recent empirical findings illustrate that observers have expectations about the speed of an observed movement depending on the size of the target (Grosjean et al., 2007) and that these expectations follow Fitts's law. However, it is yet unclear whether this law is used to predict ongoing observed actions. If so, this would allow people to predict the target of a partner's actions when many potential targets are present. The first question of the current research was whether observers indeed can use the velocity of an action to *predict* whether an action is directed at a small or large object. The key advantage of action prediction over mere processing of completed actions is that prediction allows for smooth and timely social interaction (Bekkering et al., 2009; Sebanz and Knoblich, 2009). A second aim of the study was to investigate which mechanism underlies velocity-based predictions by taking a developmental approach. Given the large body of literature suggesting that the motor system is involved in action prediction (Wolpert et al., 2003; Oztop et al., 2005; Prinz, 2006; Kilner et al., 2007a) and prior empirical evidence that Fitts's law affects action observation (Grosjean et al., 2007; Eskenazi et al., 2009), it is plausible that motor simulations bring about velocity-based predictions. As a second question we therefore examined whether motor development goes hand in hand with the development of velocity-based predictions, by employing a cross-sectional design.

When acquiring a novel motor skill, the actor builds associations between the motor commands utilized and the effects of these motor commands as experienced via the sensory modalities (Miall and Wolpert, 1996; Kawato, 1999). At first, gaze is directed at the effectors (hands, fingers, feet) to monitor the results of the new motor commands (White et al., 1964; Sailer et al., 2005). With action proficiency, gaze will no longer be directed at the effectors, but at the target of the action (Sailer et al., 2005) and hence reveals the target of the ongoing action. Based on associations formed during the acquisition phase, a forward model of the action can be constructed, which allows the actor to predict the sensory consequences of an intended action ahead of time (Wolpert, 1997). The forward model becomes more fine-grained with increasing motor experience. In this way, motor experience leads to a precise forward model of the action and to precise predictions of future sensory states.

Motor theories of action perception assume that similar processes are active during action perception as during action production (e.g., Oztop et al., 2005). Numerous studies have demonstrated that brain areas responsible for action production are activated during action perception as well (Hari et al., 1998; Buccino et al., 2001; Cattaneo et al., 2010). The observers' motor

system of both adults (Calvo-Merino et al., 2005, 2006) and infants (van Elk et al., 2008) appears to be more activated during observation of acts that are firmly established in the observers' motor repertoire compared to more novel motor acts. On a behavioral level, goal-directed eye movements have been shown to be predictive and follow the same time course for action execution and action observation (Flanagan and Johansson, 2003), and blocking the motor system by means of Transcranial Magnetic Stimulation (TMS) disrupts these predictive eye movements (Elsner et al., 2013). Eye-tracking studies investigating the development of action prediction indicate that motor experience is crucial for predicting these actions in others (Falck-Ytter et al., 2006; Kanakogi and Itakura, 2011; Ambrosini et al., 2013; Stapel et al., submitted). Participants with difficulties in planning their own action sequences, namely children with autism, show also less indications that they predict others' actions, whereas typically developing children anticipate their own next action, and a similar predictive muscle activation is found when they observe the same action in others (Cattaneo et al., 2007; Fabbri-Destro et al., 2009). Based on these findings, it is therefore likely to assume that velocity-based predictions become more accurate as a consequence of motor development.

In action performance, speed depends on the accuracy required for successful completion of the action. That is, the more accurate one has to be, the slower the movements become. Fitts (1954) formalized and quantified this relation based on data he collected, and the relation he found was shown to hold for many movements (see for an overview Plamondon and Alimi, 1997). Fitts's law states that the time needed to move between two targets is based on the distance between the targets and the width of the target (Fitts, 1954). Hence, average velocity can be higher between large compared to small target objects, and bridging small distances can be done quicker than bridging large distances. For example, reaching and grasping a small object requires more accuracy, and has been shown to take more time (Bootsma et al., 1994; Zaai and Thelen, 2005).

Empirical research shows that in adults, not only action production follows Fitts's law; also action perception is influenced by it. For instance, adults were capable of dissociating whether an observed, artificial reaching movement was physically possible or impossible in reality given the average velocity, adhering in their judgments to Fitts's law (Grosjean et al., 2007). Also, a neurophysiological patient violating Fitts's law in his action production by not adjusting movement speed for smaller targets displayed similar violations in action perception (Eskenazi et al., 2009). This indicates that determining whether observed actions have an appropriate velocity might be grounded in the action production capabilities of the observer. Presumably, the neural motor system is recruited during action perception to simulate the observed action. These simulations during action observation may enable the observer to predict future states of the action (cf. Wilson and Knoblich, 2005). An fMRI study by Eskenazi et al. (2011) revealed that activity in motor areas of the brain during the observation of movements was related to the difficulty of performing these movements as formalized in Fitts's law. In sum, the speed-accuracy trade-off not only constrains action production, it also affects

action observation, and these constraints influence activity in motor cortical areas of the brain during observation and execution. The speed-accuracy trade-off has primarily been studied in adults; little is known about the development of the perception of actions that differ as a consequence the speed-accuracy trade-off. The current study takes a novel approach by investigating the mechanisms underlying processing of the speed-accuracy trade-off from a developmental perspective.

The study was set out to investigate whether observers not only use the speed-accuracy trade-off to dissociate possible from impossible actions, but whether they also use this principle to predict the targets of actions they observe. Furthermore, if the motor system generates target predictions based on the velocity of the observed movements, then these predictions can only be made by observers capable of performing the observed action herself, because before skill acquisition, the observer most probably lacks the necessary forward model to predict the action outcome. We therefore adopted a developmental approach: 9-, 12-, and 15-month-old infants participated together with adults in an eye-tracking experiment during which they observed an actor moving her hand toward and pressing a large or a small button. In all stimulus videos, there were two buttons, a large and a small one, at the end of a table. A hand started moving from the side of the table and to the other to press either the large or the small button. Natural movements were used in the stimuli, resulting in slower movements toward the small button than to the large button. If participants made more correct visual anticipations than incorrect anticipations, then that would form an indication that the observers used the velocity of the hand to predict whether a specific button would be pressed or not. We hypothesized the ability to predict others' aiming and pressing actions to develop in parallel with their own ability to accurately aim their hand and finger at a small target in order to press it. Pressing a small button requires the use of the index finger independently from the other fingers. This ability is also needed to grasp small objects with the pincer grasp. At 8 months of age, typically half of the infants is capable of performing the pincer grasp (van der Meulen et al., 2002). Infants begin to use the pincer grasp more frequently and more precisely as they get older. These developmental changes occur mainly until 15 months of age, as the use of the power (whole hand) grip decreases (Butterworth et al., 1997). Young infants might thus be able to successfully aim with their hand for a large button, but they might base their movements on a relatively inaccurate forward model, which prevents them from smoothly reaching for and pressing a small button. Having a coarse-grained forward model might necessitate them to make corrections in their movements if they would try to aim for and press a small button. At the same time, this coarse-grained forward model might not allow them to make accurate predictions of other's actions. To further clarify the role of motor expertise for velocity-based action prediction, the infant groups were tested for their ability to aim at a small button. This allowed us to disentangle whether potential developments in predicting targets based on speed arise specifically from the development of the motor skill at hand or rather reflect other age-related changes.

Materials and Methods

In the following section, we report the way sample size was determined, all data exclusions, all manipulations, and all measures of interest for the study.

Participants

Due to the innovative nature of the study, it was impossible to perform a reliable effect size estimation based on previous studies, rendering the study exploratory. We aimed to gather data of at least 24 infants per age group. As adults provide more stable gaze data and are better capable of attending for longer durations, we aimed for testing 18 adults. In this type of study, there are two forms of drop out: immediate drop-outs due to insufficient gaze calibration or infant distress and failure to collect enough gaze data. The first form can be noticed during testing and hence this type of drop out can immediately be replaced. The second form can only be discovered during the analyses. Drop-outs that occurred during the analyses were not replaced. For that reason, sample sizes vary slightly between the groups.

Twenty-seven infants (eight girls) with a mean age of 8.8 months ($SD = 0.3$), 28 infants (16 girls) with a mean age of 12.2 months ($SD = 0.3$), and 28 infants (11 girls) with a mean age of 15.0 months ($SD = 0.2$) participated in the study. Furthermore, 18 adults (12 women, mean age = 24.9 years, $SD = 5.2$) took part in a longer version of the experiment. Eight additional infants (three 9-month-olds, five 12-month-olds) and one additional adult were tested but excluded from the analyses because they did not meet the eye-tracking calibration criteria (seven infants) or because they produced an insufficient amount of gaze data (gaze data for only three or less trials: one infant, one adult). The production task of 12 infants (six 9-month-olds and six 12-month-olds) could not be analyzed as it turned out to be difficult to videotape the action execution task from an angle at which both the infant, the infant's hand and the device was visible at all times. In nine cases, (part of) the action was not visible in the video, rendering it impossible to code the behavior later on. In three other cases, the action was not recorded due to experimenter error. All infant groups were recruited via the Baby Research Center in Nijmegen. The adults were recruited via a participant database of Radboud University Nijmegen. Written informed consent of the participants or the participants' parents was obtained prior to participation. Participation in the study was rewarded with a small gift (an infant book or 10 Euros for the participating infants, 5-Euro-gift vouchers or credit points for the adults). The study was approved by the ethical committee of behavioral science at the Faculty of Social Sciences in Nijmegen (approval number ECG2012-1301-006 for the infant participants and approval number ECG2012-0910-058 for the adult participants), and was conducted in conformity to the ethical standards of (developmental) psychology.

Stimuli

Four different short video clips (duration: 3.1–3.6 s) were used as stimulus material. The videos showed a table with a large (4 by 4 cm) and a small (1 by 1 cm) button on one side (see **Figure 1**). Velocity of natural movements directly impacts the

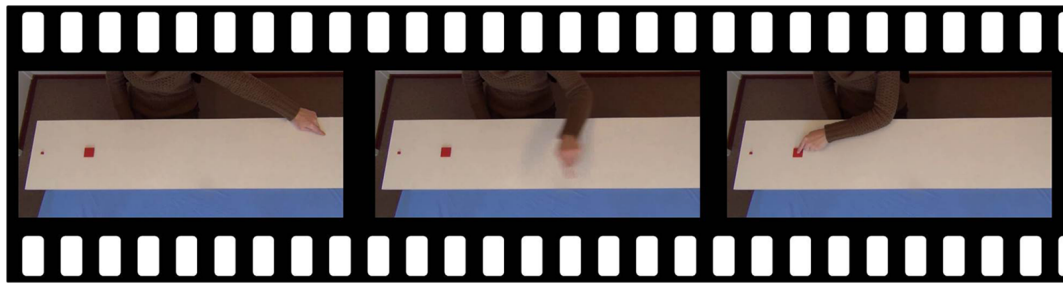


FIGURE 1 | Example frames from a stimulus video with the small button near the edge of the table and the hand moving to the large button.

height of the movement trajectory: slow movements allow for a stronger curvature than fast movements (Lacquaniti et al., 1983). To minimize potential effects of movement trajectory, the actions were filmed from a near top view. An actor was sitting behind the table. One of the buttons was relatively near the edge of the table, and the other one was a bit further away from the edge toward the middle (center-to-center distance between the buttons was 20 cm). In half of the videos, the small button was the one closer to the edge of the table, whereas it was the large button in the other half of the videos. The stimulus videos started with a still frame in which the actor's hand was shown on the far side of the table. To create a balanced stimulus set, also horizontally flipped versions of the videos were made by means of editing the original video material in VirtualDub (www.virtualdub.org). After 1 s, the hand started moving toward the buttons, and the action ended with the hand pressing one of the buttons with the index finger. The video ended with 1 s of still frame of the hand in its end position with the index finger pressing the button. This could be either a small and far, small and close, large and far, or a large and close button. Natural reaching movements were used because biological motion processing is thought to be disrupted by artificial compared to natural movements (Servos et al., 2002; Kilner et al., 2007b). The actress was instructed to fixate at the target throughout the trial and to direct her head to a fixed point in space on a line intersecting the midpoint between the two buttons, thereby avoiding potential cues of shoulder direction to influence the predictions of observers. The index finger was already stretched out during the start of the movie, such that during movement the fingers did not move with respect to the hand. As expected based on Fitts's law, movements toward the small buttons took more time than movements toward the large buttons (300 ms difference), and pressing the distal button required more time than pressing the proximal button (20 ms difference). The resulting average velocity of the hand until it reached the area of interest around the closest button ranged between 988 and 1522 px/s. The average velocity of the hand was 1222 px/s moving toward the large close button, 1522 px/s toward the large far button, 988 px/s toward the small close button, and 1240 px/s toward the small far button. Hence, the average velocity of the natural movements was manipulated by means of manipulating the size of the target button as well as the distance to the target button.

The motion paths of the actions have been visualized to give more insight into the variability between the stimuli, see **Figure 2**. The image was constructed in the following way: (1) The frames from the period of interest per stimulus video were saved as bitmaps, (2) The location of the tip of the index finger was marked per frame with a colored dot. (3) The images were read in frame by frame using Matlab R2014b (MathWorks Inc.) and the locations of the colored dots were stored into a matrix per video, (4). The four matrices were added and plotted. The figure illustrates that natural reaching movements indeed are variable, but no clear pattern is visible revealing that the one type of paths leads to the far and another type to the close button: the blue paths are not very similar to each other, nor are the red paths. The red paths continue further to the left, which illustrates that these actions decelerated at this point, whereas the other two actions continue on full speed at this point.

Button Press Device

To assess the infants' proficiency of aiming at and pressing large and small buttons, a button press device was constructed (see **Figure 3**). The device consisted of a wooden frame, in which boards with a single, red button could be fitted. Two boards were used, one with a small (1 by 1 cm) button, and one with a large button (4 by 4 cm) in the middle of the board. As the initial starting position of a reaching infant's hand is relatively difficult to control, manipulating distance was expected to be difficult. Therefore, only button size was manipulated in the execution task. To ensure that infants would aim precisely at the button instead of pushing it with their whole hand, the buttons were inlaid into the surface, with a black edge around them. Pressing elicited a sound to enhance infants' motivation to try to succeed in pressing the button.

Procedure

The procedure for data collection was kept as similar as possible across age groups. Participating infants were seated in a car chair resting on the lap of their caregiver in front of a computer monitor. Participating adults were seated on an office chair adjusted to their height. Infants' gaze was recorded using a Tobii 1750 (Tobii Technology, Sweden). Adults' gaze was recorded with a different, but comparable eye-tracker (Tobii T120; Tobii Technology, Sweden), as adults were tested for a different, unrelated study at the same occasion. All participants first

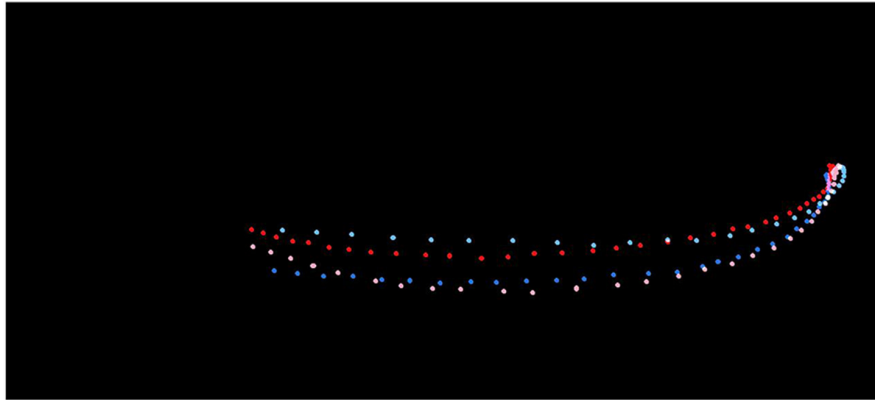


FIGURE 2 | Illustration of the motion paths used in the stimuli. The pink dots represent the motion path of the action toward the large close target, the red dots represent the action toward the small close target, the light blue represents the action targeted at the large far button, and the dark blue represents the action targeted at the small far button.

underwent a calibration procedure in which a contracting and expanding circle accompanied by a sound was shown on nine locations on the screen, forming a 3-by-3 grid. The calibration was accepted, if data was available for seven or more calibration points. The calibration procedure took between 2 and 5 min time. Immediately after calibration, the experiment started, which consisted of 96 (adults) or 48 (infants) trials. Trials were presented in pseudo-random order and were interleaved with brief attractive audiovisual clips to maintain the attention of the participants to the screen (16 for the infants, 3 for the adults). Stimulus presentation took 7 min for the adults and between 3 and 4 min for the infants (some infants were very attentive and in these cases some of the attention getters could be omitted). Trials were randomized within blocks, such that each block consisted of a random sequence of all eight unique stimulus videos. Infants thus observed 6 blocks and adults 12 blocks.

After the eye-tracking experiment, infants who had been sitting in the car seat were put on their parents lap. They were presented with the button pressing device, which stood on the table in front of them. Their actions were recorded with a video camera (Sony handycam DCR-SR190, frame rate: 25 Hz). They

were first asked to try to press the large button, then the small button, followed by again the large and then the small button. The large button was presented first to maximize the chances that infants would try out both buttons. Had first the small button been presented, some infants might have started with a failure, diminishing the chances that they would continue with the other button. Presenting the small button first might have caused a selective drop-out as the younger infants were expected to have problems pressing the small button. The experimenter demonstrated how to press the button and encouraged the infant to follow her example in case infants were hesitant to press the button themselves. Infants were tested until they lost interest or for maximally 1.5 min per button type (large or small). On average, infants explored the large button for 56 s in case of the large button, and 55 s in case of the small button. One 15-month-old did not attempt pressing any of the buttons. In addition, one 9-month-old and two 15-month-olds did not show clear attempts pressing the small button but did demonstrate attempts pressing the large button (see **Table 1**).

The eye-tracking task always preceded the button press task as infants tend to become restless over time during a testing session,

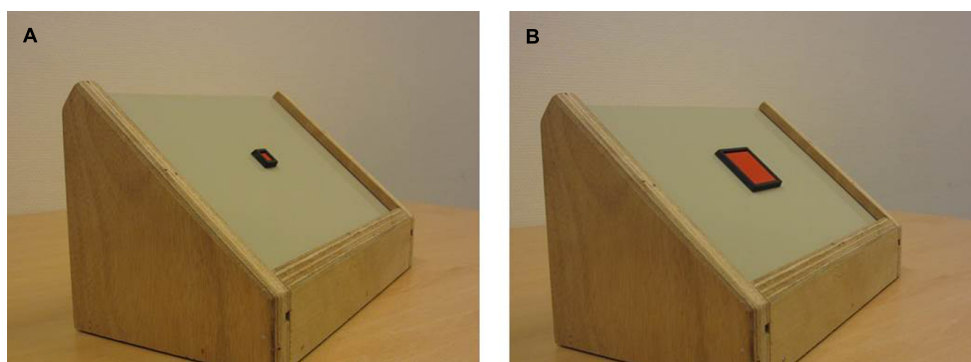


FIGURE 3 | The button press device. The small button is presented at the left (A), and the large button at the right (B).

TABLE 1 | Minimum, maximum, SD, and average number of valid attempts to aim and press the button per age group and button size.

Age group	Button size	Minimum	Mean (SD)	Maximum
9-month-olds	Large	3	12 (7)	27
	Small	1	9 (7)	25
12-month-olds	Large	2	8 (5)	21
	Small	0	5 (3)	9
15-month-olds	Large	0	9 (6)	21
	Small	0	6 (4)	20

and the button press task allowed for more movement of the infant than the eye-tracking task. Previous research has shown that only motor training but not observational training affects later perception of the trained action (Gerson and Woodward, 2014a,b; Gerson et al., 2015), and therefore no carry-over effects were expected.

Gaze Data Analyses

Square-shaped areas of interest (AoIs) of equal size (100 by 100 pixels) were defined around the buttons in the stimulus displays, and in addition, an AoI was defined containing the full display of the stimulus movie (1280 by 580 pixels). First, the stimuli that were attended to were counted per participant and per condition. A stimulus was considered “watched” if at least one fixation fell on the full stimulus AoI while the stimulus video was playing. Second, per condition, trials were counted in which the participants fixated at one of the two button AoIs after onset of the hand movement and before the hand reached the AoI of the close button. These target fixations are subsequently referred to as “anticipatory looks.” A percentage of trials in which participants showed an anticipatory look to one of the buttons was calculated based on the total number of watched trials in that condition. In trials in which participants looked at both buttons during the anticipation interval, the trial would count both as a target and a non-target anticipation. Repeated measures ANOVAs were used to investigate whether participants correctly predicted whether a button served as the target of the action or not.

Video Coding of Button Presses

Infants’ attempts to press the large and small buttons were coded from the video-recordings. Per type of button, the attempts to press the button were counted. Behavior was considered as an attempt to press the button if the infant’s hand touched the board in which the button was embedded while the infant looked at the button. Button press attempts were considered successful if the infant touched the button while looking at it. Attempts in which the infant was being moved or helped by their caregiver were excluded from the analyses. Beside success on the task, we were interested in the quality of the infant’s aiming. A well-aimed button press needs no correction in the movements, such that the aiming hand or finger lands directly on the button instead of first on the surroundings of the button. Movement correction was quantified as the time between the first moment the device was touched and the first moment the button was touched. Accurate initial aims would result in short (down to 0 s)

movement correction times. If an infant had no successful button press attempts for one of the buttons, no data was available for the movement correction time of that button.

Results

Action Perception

The action in the stimulus display became disambiguated once the hand reached the close button, as then either the hand stayed on the close button, or continued to the far button. Thus, importantly, only anticipatory fixations initiated during this first ambiguous phase of the action were analyzed (the duration of the ambiguous phase ranged from 1.58 to 1.88 s after stimulus onset). An implication of this analysis choice was that fixations to the close button would likely occur more frequently compared to fixations to the far button, because for the latter, gaze needed to be more ahead of the action in space and time to reach the button during that period. Inspection of the data substantiated this assumption. **Figure 4** and **Table 2** display the mean percentage of fixations to the close button (closest to the initial position of the hand) and the far button (further from the initial position of the hand) during the analysis window collapsed over conditions. Given that participants tended to anticipate only to the close button and appeared to exhibit hardly any anticipations to the far button, the subsequent conditional analyses will focus on anticipations to the close button, which was either the target of the action, or not.

A repeated measures ANOVA was conducted to analyze the frequency of anticipatory looks to the close button with button function as a within-subjects factor (target, non-target) and age group (9-month-olds, 12-month-olds, 15-month-olds, adults) as a between-subjects factor. There was a main effect of age on the percentage of anticipatory looks [$F(3,97) = 50.33$, $p < 0.001$, $\eta_p^2 = 0.61$]. *Post hoc* independent samples *t*-tests

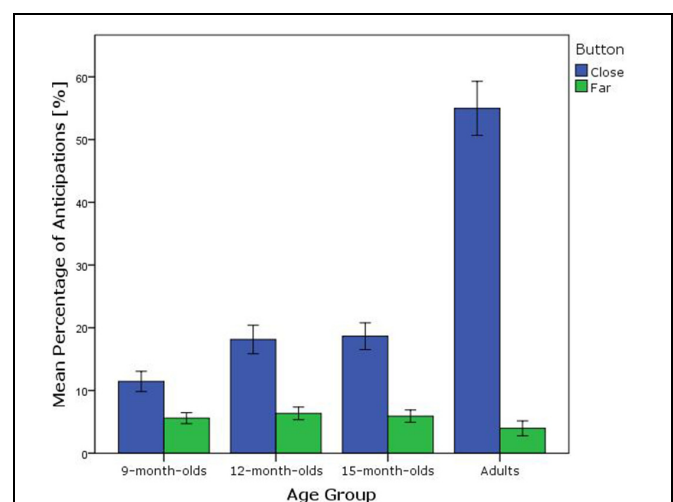


FIGURE 4 | Mean percentage of visual anticipations to the button close or far from the initial position of the hand, regardless of condition, split by age group. Error bars represent 1 SEM.

TABLE 2 | Minimum, maximum, SD, and average number of observed trials per age group and button location.

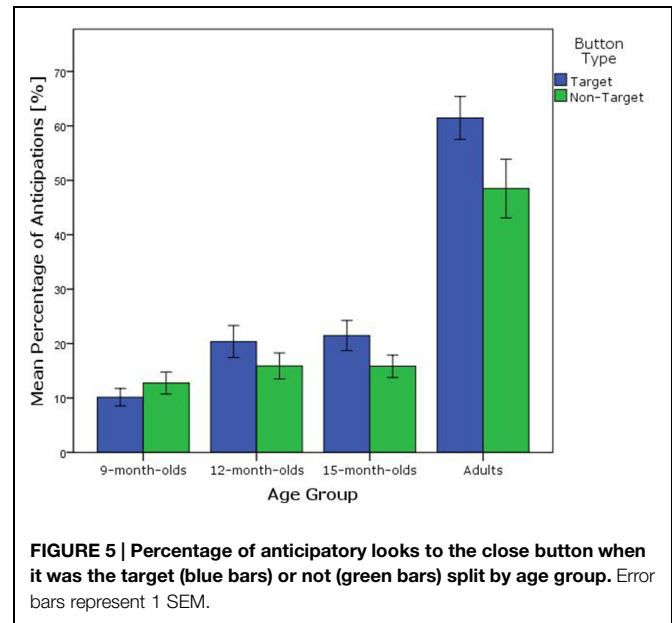
Age group	Button location	Minimum	Mean (SD)	Maximum
9-month-olds	Close	9	20.1 (4.5)	24
	Far	7	19.4 (5.0)	24
12-month-olds	Close	7	20.1 (4.8)	24
	Far	10	20.8 (4.4)	24
15-month-olds	Close	17	22.3 (1.8)	24
	Far	18	22.7 (1.5)	24
Adults	Close	40	47.6 (1.9)	48
	Far	40	47.6 (1.9)	48

showed that adults displayed a higher percentage of anticipatory looks ($M = 55\%$, $SD = 18$) than the 15-month-olds [$M = 19\%$, $SD = 11$, $t(25.4) = 7.55$, $p < 0.001$]¹ and the 12-month-olds [$M = 18\%$, $SD = 12$, $t(26.5) = 7.56$, $p < 0.001$]. No difference was found in anticipatory looks between the 15- and 12-month-olds [$t(54) = 0.17$, $p = 0.867$]. The 9-month-olds showed less frequent anticipatory looks ($M = 11\%$, $SD = 8$) than the 12- [$t(53) = 2.38$, $p = 0.021$] and 15-month-olds [$t(53) = 2.68$, $p = 0.010$].

A main effect of button function was observed [$F(1,97) = 14.56$, $p < 0.001$, $\eta_p^2 = 0.13$], indicating that across age groups, participants anticipated more frequently to the close button when it was the target ($M = 25\%$; $SD = 22$) compared to when it was not the target button ($M = 21\%$; $SD = 19$). A significant interaction effect was found [$F(3,97) = 5.09$, $p = 0.003$, $\eta_p^2 = 0.14$], indicating that the age groups differed in the frequency of anticipatory looks to the target compared to the non-target button. To further verify that the interaction effect was not solely due to the difference between adult and infant performance, an ANOVA was run without the adult data. A marginally significant main effect of button function was found [$F(1,80) = 3.38$, $p = 0.070$, $\eta_p^2 = 0.04$], together with a significant interaction effect of age group and button function [$F(2,80) = 3.51$, $p = 0.035$, $\eta_p^2 = 0.08$]. Planned paired comparisons for the separate age groups revealed that adults anticipated more frequently to the button when it was the target compared to when it was not [$t(17) = 3.32$, $p = 0.004$]. The same was the case for the 15-month-olds [$t(27) = 2.37$, $p = 0.025$], whereas the 12- and 9-month-olds did not look more frequently at the close button when it was the target compared to when it was not [12-month-olds: $t(27) = 1.59$, $p = 0.125$, 9-month-olds: $t(26) = -1.45$, $p = 0.141$; see **Figure 5** and **Table 3**].

Action Production

A repeated measures ANOVA was used to analyze the effect of button size (small, large) and age group (9-month-olds, 12-month-olds, 15-month-olds) on the percentage of successful button press attempts out of all attempts. A main effect of button size on the percentage of successful button presses was found [$F(1,68) = 28.05$, $p < 0.001$, $\eta_p^2 = 0.29$], indicating that the infants were more successful in pressing the large ($M_{\text{large}} = 88\%$, $SD_{\text{large}} = 22$) compared to the small button ($M_{\text{small}} = 69\%$,

**FIGURE 5 | Percentage of anticipatory looks to the close button when it was the target (blue bars) or not (green bars) split by age group. Error bars represent 1 SEM.****TABLE 3 | Minimum, maximum, SD, and average frequency of anticipation (%) per age group and button function.**

Age group	Button function	Minimum	Mean (SD)	Maximum
9-month-olds	Target	0	10.2 (8.3)	33.3
	Non-target	0	12.8 (10.6)	36.1
12-month-olds	Target	0	20.4 (15.6)	62.5
	Non-target	0	15.9 (12.7)	41.7
15-month-olds	Target	0	21.5 (14.7)	56.0
	Non-target	0	15.8 (10.9)	37.5
Adults	Target	37.5	61.5 (16.8)	89.6
	Non-Target	12.5	48.5 (22.9)	87.5

$SD_{\text{small}} = 37$). Furthermore, the interaction between age group and button size was found to be significant [$F(2,68) = 15.18$, $p < 0.001$, $\eta_p^2 = 0.31$]. Independent samples t -tests showed that the 12-month-olds were more successful than the 9-month-olds when trying to press the small button [$t(32.9) = 5.79$, $p < 0.001$], but no significant differences were found between these groups when trying to press the large button [$t(42) = 0.51$, $p = 0.611$]. In addition, the percentage of successful button presses was found to depend on age [$F(2,68) = 15.18$, $p < 0.001$, $\eta_p^2 = 0.31$], with the 12-month-olds showing more successful button presses than the 9-month-olds [9-month-olds: $M_{\text{small}} = 38\%$, $SD_{\text{small}} = 34$, $M_{\text{large}} = 85\%$, $SD_{\text{large}} = 12$; $t(42) = 4.73$, $p < 0.001$]. The success rates of the 12-month-olds for the small and large button ($M_{\text{small}} = 86\%$, $SD_{\text{small}} = 19$, $M_{\text{large}} = 88\%$, $SD_{\text{large}} = 24$) were not different from the 15-month-olds [$M_{\text{small}} = 81\%$, $SD_{\text{small}} = 35$, $t(47) = 0.65$, $p = 0.522$; $M_{\text{large}} = 90\%$, $SD_{\text{large}} = 26$, $t(48) = -0.29$, $p = 0.771$].

An identical repeated measures ANOVA was conducted on the movement correction time data. A main effect of button size was observed [$F(1,63) = 53.81$, $p < 0.001$, $\eta_p^2 = 0.46$], as significantly more time was needed to correct the aiming movement to a

¹In case equal variances could not be assumed as indicated by a significant outcome of Levene's test for equality of variances, adjusted dfs are reported.

small ($M_{\text{small}} = 0.52$ s, $SD_{\text{small}} = 0.54$) than to a large button ($M_{\text{large}} = 0.08$ s, $SD_{\text{large}} = 0.14$). The interaction between age group and button size had a significant effect on the movement correction times [$F(2,63) = 6.69$, $p = 0.002$, $\eta_p^2 = 0.18$]. The three age groups were equally fast in pressing the large button ($M_{9 \text{ months}} = 0.09$ s, $SD_{9 \text{ months}} = 0.12$, $M_{12 \text{ months}} = 0.10$, $SD_{12 \text{ months}} = 0.12$, $M_{15 \text{ months}} = 0.06$, $SD_{15 \text{ months}} = 0.18$, all $ts < 1.0$, all $ps > 0.308$). However, the 15-month-olds needed less time for correcting their movements than the other two groups when aiming for the small button ($M_{9 \text{ months}} = 0.82$ s, $SD_{9 \text{ months}} = 0.79$, $M_{12 \text{ months}} = 0.50$ s, $SD_{12 \text{ months}} = 0.32$, $M_{15 \text{ months}} = 0.27$ s, $SD_{15 \text{ months}} = 0.15$; $ts > 3.0$, $ps \leq 0.006$), whereas the 9- and 12-month-olds differed only marginally in this respect [$t(26.8) = 1.71$, $p = 0.099$]. Furthermore, movement correction time was dependent on age [$F(2,63) = 6.93$, $p = 0.002$, $\eta_p^2 = 0.18$], which was caused by differences in aiming for the small button.

Learning Effects

The set of video stimuli consisted of eight unique movies which were repeated six times for the infants and 12 times for the adults. Potentially, the found effects might hence be due to learning during the experiment. To investigate whether learning had occurred, the average anticipation frequency was calculated per block, per individual and split by condition. The anticipation frequencies were subjected to a six (blocks) by two (button function) by four (age group) mixed ANOVA. There are two results of relevance for the question of learning effects. First, an interaction between block and button function could indicate learning throughout the age groups. This interaction was found to be not significant [$F(5,420) = 1.09$, $p = 0.364$]. The second relevant result is the three-way interaction between block, button function, and age group. A significant interaction might indicate that the younger two groups did not show learning within the experiment whereas the other two groups did display learning effects. This three-way interaction was found to be marginally significant [$F(15,420) = 1.57$, $p = 0.078$]. To verify whether this indeed indicates that the older two age groups learnt when the close button was the target and when not, a follow-up six by two by two ANOVA was conducted only including the data of the 15-month-olds and the adults. If learning is to explain the differences found between the younger two groups and the older groups, then this ANOVA should yield a significant interaction between block and button function. This interaction was not found to be significant [$F(5,215) = 0.64$, $p = 0.673$], which shows that learning during the experiment cannot explain the differences found in predictions between the 9- and 12-month-olds on the one hand, and the 15-month-olds and adults on the other hand. More details on the analyses of potential learning effects can be found in the supplementary materials.

Relation between Action Observation and Action Production

The results presented above show that success rates in aiming at the small button improved between 9 and 12 months of age and movement correction times decreased between 12 and

15 months of age. The ability to make velocity-based predictions develops in parallel, as 15-month-olds displayed velocity-based predictions, whereas 9- and 12-month-olds did not. To study the relation between action observation and action performance more closely, we examined the group of 12-month-olds, as this was the transitional group consisting of infants who were at the verge of learning to use velocity of natural movements to predict actions. A correlation analysis was performed to investigate whether action production and action prediction skills were related at the level of the individual infants. In the correlation analyses, proficiency in aiming at the small button was used as the measure of interest, as this reflects the ability to aim with high precision best. The time needed to correct the aiming movements to the small button was not found to be related to the prediction accuracy, expressed as the difference between the percentage of target and non-target anticipations ($p = 0.654$, controlling for age in days). Likewise, the relation between the success rate of aiming at the small button was not found to be related with action prediction accuracy ($p = 0.902$, controlling for age in days).

Discussion

The aim of the current study was to examine whether the velocity of a natural movement, as manipulated through manipulating the size of and the distance to the targets, is used by an observer to predict whether an object will be the target of the observed action, and if so, whether motor development and hence the motor system is crucial for these predictions to emerge. Gaze data showed that adults and 15-month-old infants more frequently displayed visual anticipations to a button when it was the target compared to when it was not. No learning over trials was observed. The speed-accuracy trade-off, slower movements toward smaller targets, and the two-thirds power law expressing a related velocity dependent phenomenon, namely slower movements allow higher bell-shaped movement trajectories (Lacquaniti et al., 1983), are the only lawful relations that can have been acquired prior to the study. The results thus indicate that 15-month-olds and adults based their predictions on the speed of an observed movement, as velocity was the central cue for distinguishing targets from non-targets. In contrast, infants of 9 and 12 months of age did not show any indications that they used the speed information of the observed movement for their action predictions. This was congruent with the development of producing this action: 15-month-olds were more proficient in aiming at and pressing a button accurately than the 12- and 9-month-olds. This suggests that the motor system underlies velocity-based predictions.

Three factors influenced how frequently the observers looked at the buttons while the action was unfolding. First, many more anticipatory looks were made to the button nearest to the initial position of the hand than to the button located further away, when the hand had not yet passed the nearest button. However, our analysis period ended when the hand was at the point of passing the nearest button, because once the hand had passed the close button, it was obvious that the far button

was the target. As a consequence, to be counted as a predictive look, observers had to be more ahead of the action when predicting the far button than when predicting the close button. Due to the low base rate of predictions to the far button, only the predictions to the close button could be analyzed. Future studies could overcome this distance problem by using 3D environments such as virtual reality to, for instance, create an ambiguous situation in which the targets create an equally sized image on the retina but differ in distance to the observer. The speed of the movement might then disambiguate the situation.

The second factor that influenced anticipatory looks was the velocity of the natural movement, which was the main factor in the current study which was manipulated by means of using differently sized targets placed at two distances. The results showed that participants looked more frequently at the close button when it was the target compared to when it was not, which indicates that the participants made use of the velocity information of the hand to predict which button would be pressed.

The third factor that affected the frequency of anticipatory looks was age. Whereas adults and 15-month-old infants looked more frequently to the close button when it was the target compared to when it was not, 9- and 12-month-old infants did not show this difference.

Velocity-based predictions may result from action simulation in the motor system of the observer. The motor system has been shown to respond stronger to the observation of actions that have to be performed with more accuracy (Eskenazi et al., 2011). The speed people expect to see during an observed action matches the actual speed of the performed action (Grosjean et al., 2007; Eskenazi et al., 2009), which illustrates that the action-perception link also plays a role in the speed-accuracy trade-off (cf. Rizzolatti et al., 1996; Hari et al., 1998; Flanagan and Johansson, 2003; Cattaneo et al., 2007). However, thus far, observation of the speed-accuracy trade-off has primarily been studied in adults, which left the question unanswered how the perception of the speed-accuracy trade-off develops. Given these prior findings, the hypothesis of the current study was that the motor system not only underlies *post hoc* judgments of the observed velocity of movements, but also facilitates on-line predictions made while the action still unfolds. Our results are in line with this hypothesis: the action prediction performance of the 15-month-old infants suggested that they use velocity information in action prediction, whereas the 9- and 12-month-olds seemed not to integrate the observed velocity in their predictions of the observed actions. The tested 15-month-old infants were also better at pressing buttons than the 9- and 12-month-olds. Using velocity information to predict which button will be pressed thus follows – at least by and large – the same developmental time course as the ability to press buttons. This is in line with previous infant research showing that motor ability affects action perception (van Elk et al., 2008; Kanakogi and Itakura, 2011; Ambrosini et al., 2013; Gerson and Woodward, 2014a,b; Gerson et al., 2015). However, within the group of 12-month-old infants, the individual button pressing proficiency was not found to be related to the ability to use speed for action prediction. It might well be that our motor measurement was

not sensitive enough to correlate motor performance with action prediction performance at an individual level. Nevertheless, it is interesting that the differences in motor performance at the group level overlap with the anticipatory eye capacities in the observation task. However, at least two alternatives can be given for the suggested improvement in terms of motor simulation. First, visual experience acquired between 12 and 15 months of age may contribute to velocity-based predictions as well (Hunnius and Bekkering, 2014). Second, the effects observed could also be related to a general maturation pattern of the brain that enables both action execution as well as action observation. The importance of visual experience and brain maturation in the development of velocity-based predictions can be tested in future research by investigating whether 15-month-olds can use velocity information for the prediction of actions that are not yet part of their motor repertoire. Furthermore, it would be interesting to study groups with delays in motor development to gain more knowledge about whether or not motor experience is necessary for velocity-based predictions.

Conclusion

We found empirical evidence that observers can predict whether an object will be the target of an action based on the velocity of the observed natural movement, which was manipulated through manipulating the size of and the distance to the target objects. In the current study, the action target was a button. Fifteen-month-old, but not 9- and 12-month-old infants showed an adult-like prediction pattern, suggesting that at 15 months of age, infants are beginning to use velocity to inform their predictions of other's button pressing actions. The 15-month-olds were more proficient in performing this type of action compared to the 9- and 12-month-olds. Together, this indicates that the development of velocity-based predictions follows a time line corresponding to the development of motor skill of the predicted action. Future research should parse out the roles of visual and motor experience for action prediction. Being a proficient actor may turn out to be necessary in order to accurately predict what other people are planning to do.

Acknowledgments

We would like to thank the participants and the parents of the participating children. Furthermore, we would like to thank Lieke Zomer, Eline Koster, Amber Joosen, Birgit Knudsen and Angela Khadar for their support in recruiting participants, data collection and video coding. We are grateful to Mark van de Hei for creating the infant button box.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01092>

References

- Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., and Sinigaglia, C. (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PLoS ONE* 8:e67916. doi: 10.1371/journal.pone.0067916
- Bekkering, H., De Bruijn, E. R., Cuijpers, R. H., Newman-Norlund, R., Van Schie, H. T., and Meulenbroek, R. (2009). Joint action: neurocognitive mechanisms supporting human interaction. *Topics Cogn. Sci.* 1, 340–352. doi: 10.1111/j.1756-8765.2009.01023.x
- Bootsma, R. J., Marteniuk, R. G., MacKenzie, C. L., and Zaal, F. T. (1994). The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Exp. Brain Res.* 98, 535–541. doi: 10.1007/BF00233990
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Euro. J. Neurosci.* 13, 400–404. doi: 10.1046/j.1460-9568.2001.01385.x
- Butterworth, G., Verweij, E., and Hopkins, B. (1997). The development of prehension in infants: Halverson revisited. *Br. J. Dev. Psychol.* 15, 223–236. doi: 10.1111/j.2044-835X.1997.tb00736.x
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007
- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr. Biol.* 16, 1905–1910. doi: 10.1016/j.cub.2006.07.065
- Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., et al. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17825–17830. doi: 10.1073/pnas.0706273104
- Cattaneo, L., Sandrini, M., and Schwarzbach, J. (2010). State-dependent tms reveals a hierarchical representation of observed acts in the temporal, parietal, and premotor cortices. *Cereb. Cortex* 20, 2252–2258. doi: 10.1093/cercor/bhp291
- Elsner, C., D'Ausilio, A., Gredebäck, G., Falck-Ytter, T., and Fadiga, L. (2013). The motor cortex is causally related to predictive eye movements during action observation. *Neuropsychologia* 51, 488–492. doi: 10.1016/j.neuropsychologia.2012.12.007
- Eskenazi, T., Grosjean, M., Humphreys, G., and Knoblich, G. (2009). The role of motor simulation in action perception: a neuropsychological case study. *Psychol. Res.* 73, 477–485. doi: 10.1007/s00426-009-0231-5
- Eskenazi, T., Rotshtein, P., Grosjean, M., and Knoblich, G. (2011). The neural correlates of Fitts's law in action observation: an fMRI study. *Soc. Neurosci.* 7, 30–41. doi: 10.1080/17470919.2011.576871
- Fabbri-Destro, M., Cattaneo, L., Boria, S., and Rizzolatti, G. (2009). Planning actions in autism. *Exp. Brain Res.* 192, 521–525. doi: 10.1007/s00221-008-1578-3
- Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci.* 9, 878–879. doi: 10.1038/nn1729
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 381–391. doi: 10.1037/h0055392
- Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771. doi: 10.1038/nature01861
- Gerson, S. A., Bekkering, H., and Hunnius, S. (2015). Short-term motor training, but not observational training, alters neurocognitive mechanisms of action processing in infancy. *J. Cogn. Neurosci.* 27, 1207–1214. doi: 10.1162/jocn_a_00774
- Gerson, S. A., and Woodward, A. L. (2014a). Learning from their own actions: the unique effect of producing actions on infants' action understanding. *Child Dev.* 85, 264–277. doi: 10.1111/cdev.12115
- Gerson, S. A., and Woodward, A. L. (2014b). The joint role of trained, untrained, and observed actions at the origins of goal recognition. *Infant Behav. Dev.* 37, 94–104. doi: 10.1016/j.infbeh.2013.12.013
- Grosjean, M., Shiffrar, M., and Knoblich, G. (2007). Fitts's law holds for action perception. *Psychol. Sci.* 18, 95–99. doi: 10.1111/j.1467-9280.2007.01854.x
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15061–15065. doi: 10.1073/pnas.95.25.15061
- Hunnius, S., and Bekkering, H. (2014). What are you doing? How active and observational experience shape infants' action understanding. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130490. doi: 10.1098/rstb.2013.0490
- Kanakogi, Y., and Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nat. Commun.* 2, 1–6. doi: 10.1038/ncomms1342
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* 9, 718–727. doi: 10.1016/S0959-4388(99)00028-8
- Kilner, J. M., Friston, K. J., and Frith, C. D. (2007a). Predictive coding: an account of the mirror neuron system. *Cogn. Process.* 8, 159–166. doi: 10.1007/s10339-007-0170-2
- Kilner, J., Hamilton, A. F. D. C., and Blakemore, S. J. (2007b). Interference effect of observed human movement on action is due to velocity profile of biological motion. *Soc. Neurosci.* 2, 158–166. doi: 10.1080/17470910701428190
- Lacquaniti, F., Terzuolo, C., and Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychol.* 54, 115–130. doi: 10.1016/0001-6918(83)90027-6
- Miall, R. C., and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Netw.* 9, 1265–1279. doi: 10.1016/S0893-6080(96)00035-4
- Oztop, E., Wolpert, D., and Kawato, M. (2005). Mental state inference using visual control parameters. *Cogn. Brain Res.* 22, 129–151. doi: 10.1016/j.cogbrainres.2004.08.004
- Plamondon, R., and Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behav. Brain Sci.* 20, 279–303. doi: 10.1017/S0140525X97001441
- Prinz, W. (2006). What re-enactment earns us. *Cortex* 42, 515–517. doi: 10.1016/S0010-9452(08)70389-7
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0
- Sailer, U., Flanagan, J. R., and Johansson, R. S. (2005). Eye-hand coordination during learning of a novel visuomotor task. *J. Neurosci.* 25, 8833–8842. doi: 10.1523/JNEUROSCI.2658-05.2005
- Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Topics Cogn. Sci.* 1, 353–367. doi: 10.1111/j.1756-8765.2009.01024.x
- Servos, P., Osu, R., Santi, A., and Kawato, M. (2002). The neural substrates of biological motion perception: an fMRI study. *Cereb. Cortex* 12, 772–782. doi: 10.1093/cercor/12.7.772
- van der Meulen, B. F., Ruiter, S. A. J., Spelberg, H. L., and Smrkovsky, M. (2002). *Bayley Scales of Infant Development-II. Dutch version*. Lisse, NL: Swets.
- van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., and Bekkering, H. (2008). You'll never crawl alone: neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage* 43, 808–814. doi: 10.1016/j.neuroimage.2008.07.057
- White, B. L., Castle, P., and Held, R. (1964). Observations on the development of visually-directed reaching. *Child dev.* 35, 349–364. doi: 10.2307/1126701
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460
- Wolpert, D. M. (1997). Computational approaches to motor control. *Trends Cogn. Sci.* 1, 209–216. doi: 10.1016/S1364-6613(97)01070-X
- Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. Lon. Series B* 358, 593–602. doi: 10.1098/rstb.2002.1238
- Zaal, F. T. J. M., and Thelen, E. (2005). The developmental roots of the speed-accuracy trade-off. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 1266–1273. doi: 10.1037/0096-1523.31.6.1266

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Stapel, Hunnius and Bekkering. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Automatic imitation in a rich social context with virtual characters

Xueni Pan* and Antonia F. de C. Hamilton

Institute of Cognitive Neuroscience, University College London, London, UK

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Costantini Marcello,
University of Chieti, Italy
Matthew R. Longo,
Birkbeck, University of London, UK

*Correspondence:

Xueni Pan,
Institute of Cognitive Neuroscience,
University College London, Alexandra
House, 17-19 Queen Square, London
WC1N 3AR, UK
s.pan@cs.ucl.ac.uk

Specialty section:

This article was submitted to
Cognition, a section of the journal
Frontiers in Psychology

Received: 28 March 2015

Accepted: 26 May 2015

Published: 09 June 2015

Citation:

Pan X and Hamilton AFdeC (2015)
Automatic imitation in a rich social
context with virtual characters.
Front. Psychol. 6:790.
doi: 10.3389/fpsyg.2015.00790

It has been well established that people respond faster when they perform an action that is congruent with an observed action than when they respond with an incongruent action. Here we propose a new method of using interactive Virtual Characters (VCs) to test if social congruency effects can be obtained in a richer social context with sequential hand-arm actions. Two separate experiments were conducted, exploring if it is feasible to measure spatial congruency (Experiment 1) and anatomical congruency (Experiment 2) in response to a VC, compared to the same action sequence indicated by three virtual balls. In Experiment 1, we found a robust spatial congruency effect for both VC and virtual balls, modulated by a social facilitation effect for participants who felt the VC was human. In Experiment 2 which allowed for anatomical congruency, a form by congruency interaction provided evidence that participants automatically imitate the actions of the VC but do not imitate the balls. Our method and results build a bridge between studies using minimal stimuli in automatic interaction and studies of mimicry in a rich social interaction, and open new research venue for future research in the area of automatic imitation with a more ecologically valid social interaction.

Keywords: automatic imitation, virtual reality, social facilitation effect, action sequencing, virtual characters

Introduction

Mapping one's own body and actions to the body and actions of others is a core mechanism for social cognition. Multiple studies have shown that people respond faster and more accurately when they have the chance to perform an action that is congruent with an observed action than when they respond with an incongruent action (Brass et al., 2000; Stürmer et al., 2000; Cook and Bird, 2011). However, the majority of these studies use very minimal stimuli (e.g., an image of an isolated hand). Here we test if social congruency effects can be obtained in a richer social context with sequential hand-arm actions performed by a virtual character (VC). We further explore if these effects are modulated by spatial congruency or by anatomical congruency. First, we review past studies on social action congruency effects, and on the use of VCs to explore social interaction.

Social Congruency Effects

Automatic imitation occurs when a participant responds faster in an imitative context than in a matched, non-imitative context, and provides a robust measure of how easily a participant maps actions between self and other. Two early papers developed automatic imitation paradigms which have been widely used in social neuroscience (Brass et al., 2000; Stürmer et al., 2000). In Brass et al.'s study, participants were instructed to respond to a symbolic number cue (1 or 2) while ignoring an irrelevant finger movement in the background. Reaction times were faster when the irrelevant finger movement on the screen was congruent with the participant's response than when it was

not (Brass et al., 2000). In Strümer's study, participants were instructed to make a pre-specified movement (either hand-opening or hand-closing) as soon as they saw a hand movement on the screen. Responses were faster when the instructed response was congruent with the stimulus than when it was incongruent (Stürmer et al., 2000).

A key requirement for an automatic imitation effect is that it is driven by a precise mapping between one's own body and the body of the actor, and not purely by the spatial locations of items in the field of view. It has been shown that Brass et al.'s finger-movement task and Stürmer et al.'s hand-opening task both measure a true imitation effect because both are robust to changes in the orientation of the stimuli (Heyes et al., 2005; Bertenthal et al., 2006; Cook and Bird, 2011). For instance in Heyes et al.'s version of the hand-opening task, the stimulus hand was vertically aligned and the responding hand (participant's hand) was horizontally aligned (Heyes et al., 2005) so congruent responses are anatomically but not spatially matched. In Bertenthal et al.'s version of the finger-tapping task, participants were instructed to perform finger-tapping toward both left hand and right hand as stimuli (Bertenthal et al., 2006). They found evidence for both spatial compatibility and automatic imitation effects, with the latter decreasing over the course of each individual experimental block. This suggested that both effect exist independently.

The automatic imitation effect can be modulated by the form of the actor: several studies have found that the effect is stronger for human than non-human hands, however it is still present for the latter (Press et al., 2005; Longo et al., 2008; Longo and Bertenthal, 2009; Liepelt and Brass, 2010). In Press et al.'s study, using the hand-opening task, participants were presented with both human and robot hands (Press et al., 2005). It was found that there was a congruency effect with both forms of hands, as well as an interaction between stimulus form and congruency, indicating that the congruency effect was greater with the human hand (27.9 ms) than with the robotic hand (8.8 ms). Similar results were also obtained in Liepelt and Brass's study using the finger tapping task, where participants were primed to believe that the video of a hand (covered with a glove) was either a real human hand or a wooden hand (Liepelt and Brass, 2010). Although the actual video stimuli were identical, participants in the wooden-hand group showed a reduced congruency effect as result of priming. In Longo, Kosobud, and Bertenthal's study, participants were presented with computer-generated realistic looking hand, animated with either biomechanically possible or impossible movements (Longo et al., 2008). The compatibility effect was present in both automatic (Experiment 1) and spatial (Experiment 3) imitation, and the results were similar regardless of the type of stimuli (biomechanically possible or impossible). However, in their second experiment, when participants were explicitly informed about the movements before the experiment, the compatibility effect disappeared with the biomechanically impossible movements (only automatic imitation was tested in this experiment). A follow up study found that automatic imitation of a virtual hand was reduced - but not eliminated—when participants were informed that they were going to see a virtual hand (Longo and Bertenthal, 2009). Overall, these studies

have shown that automatic imitation can be obtained for human, mechanical and computer-generated hands, with the magnitude of the effect dependent on participant's beliefs about the hand.

Richer Contexts

One limitation of current studies of automatic imitation is that they mostly used isolated hand stimuli and limited contexts. A few studies have explored larger social contexts by adding faces to moving hands (Wang et al., 2010; Grecucci et al., 2013). Grecucci et al. displayed faces with either neutral or negative emotion before each stimulus, and instructed both ASD children and health controls to perform finger-tapping presented with finger-tapping (compatible) or finger-lifting (incompatible) stimuli (Grecucci et al., 2013). It was found that both ASD and control groups had a compatibility effect, and that the control group had a significant faster response toward the stimuli following the display of negative faces, whereas this effect was not present with the ASD group. Wang, Newport, and Hamilton displayed faces with direct or averted gaze before a hand-opening/closing stimulus and measured congruency effects (Wang et al., 2010). They found that participants were faster at the congruent trials with the direct gaze than with the averted gaze in a hand-opening task, indicating that direct gaze enhances automatic imitation.

Others have added social priming before a hand action imitation task (Leighton et al., 2010; Wang and Hamilton, 2013). Using a scrambled-sentence paradigm, Leighton et al. found that pro-social priming elicited a larger automatic imitation effect in a hand-opening task, whereas anti-social priming elicited a reduced automatic imitation effect (Leighton et al., 2010). Wang and Hamilton further argued that such a pro- or anti-social priming effect is modulated by self-relatedness. They found that first-person prosocial and third-person antisocial primes both increased automatic imitation (Wang and Hamilton, 2013). A full review of the many factors modulating automatic imitation can be found in Heyes (2011) and Wang and Hamilton (2012).

The aim of the current paper was to test if automatic imitation effects can be obtained robustly in an even richer context, where participants perform actions in front of a life-size VC. VCs have been valuable in the study of human social interaction in various ways. Early studies in this area used a virtual ball tossing game with simple VCs to investigate perspective taking (David et al., 2006) and social exclusion (Eisenberger et al., 2003). Other studies use expressive VCs to study the social function of gaze (Georgescu et al., 2013), blushing (Pan et al., 2008), and mimicry (Bailenson and Yee, 2005). More recently, photo-realistic looking VCs animated with motion-captured data were used in studies of joint action (Sacheli et al., 2015), embodiment (Kiltner et al., 2013), and personality (Pan et al., 2015). The high level realism of both appearance and behavior in these studies provided a key element in achieving ecological validity.

In the present study, we used VCs to prime the performance of action sequences and test if automatic imitation can be obtained for sequential actions in a rich social context. On each trial, the VC performed a sequence of three actions, and then the participant was instructed to perform a sequence which could be congruent or incongruent with the actions of the character. As a control condition, participants saw three balls which touch

the same goal locations as the VC, without any human form or biological motion. We can identify three possible effects. First, we could find a main effect of congruency, with faster responses following congruent actions. Different configurations of the action goals can allow us to distinguish between spatial congruency and anatomical congruency (see below). Second, we could find a main effect of actor form, whereby participants are faster to respond when a VC is present. This would be a social facilitation effect, where responses are faster when participants are in the presence of a real (or virtual) human (Bond and Titus, 1983). Finally, we could find a congruency by form interaction, where form could be a virtual human or a non-human object (a moving ball). This is the signature of automatic imitation, because it indicates that participants are faster on congruent trials only when the actions are performed by a VC with a comparable body shape to the participant, and not when the same goals are indicated by a non-human object. Spatial effects can be ruled out.

There are two ways in which the actions of the participant could be congruent with the actions of the avatar. They could be directed to the same location in space (spatial congruency), or they could use the same arm movements (anatomical congruency). We test the former in Experiment 1, and the latter in Experiment 2. Based on previous findings that automatic imitation for simple finger movements is driven by anatomical effects, we predict that when movements are spatially (but not anatomically) congruent, we would find only a main effect of congruency (Experiment 1). We further predict that when movements are anatomically matched between participant and avatar, we would find a form by congruency interaction, indicating a true automatic imitation effect (Experiment 2).

Experiment 1—Spatial Congruency

Participants

A total of 22 participants were recruited from the ICN Subject Database (14 females; mean age = 22.5 years; $SD = \pm 4.3$ years).

All were right-handed (by self-report), had normal or corrected-to-normal vision, and were naïve to the purpose of the study. They received payment at the end of the study. The study was approved by the UCL graduate school ethics committee.

Materials

The experiments were conducted in our lab where VR graphics were displayed in 2D on a 90 cm × 160 cm projector screen. As shown in **Figure 1A**, the lab was prepared with a wooden stool in front of a wooden table with three plastic toy drums on top. Immediately beyond the table was a large projector screen, where the participant could see the virtual world. The virtual environment was modeled to match the real world with a virtual wooden table which looks like an extension of the real one, and three matching virtual drums modeled in 3D Studio Max (Autodesk). The drums on the desk and the drums in the virtual world were numbered 1, 2, or 3 as illustrated in **Figure 1B**. A female VC (Jessie) sat behind the virtual wooden table, facing the participant. Jessie was animated with pre-recorded motion captured data and was controlled by the VR application in real time. In front of Jessie the participant could see a virtual iPad, where the participant received instructions.

Jessie's motion was obtained by motion-capturing a single female actor performing the same task as participants at the same desk. The actor had four Polhemus Liberty magnetic motion trackers, placed on her head, chest, right hand side elbow, and middle finger. The Polhemus data was streamed into a machine running Motionbuilder (through the Polhemus plugin for Motionbuilder), which produces character animation after a small calibration session. Unlike from optical motion capture system, the magnetic trackers used here give both position and orientation data and therefore four trackers were enough to produce high quality human-like animation for our setting (upper body with one arm movement). Animation data was saved while the actor performed all possible sequences of taps on the three drums, instructed by number cues. The animation files

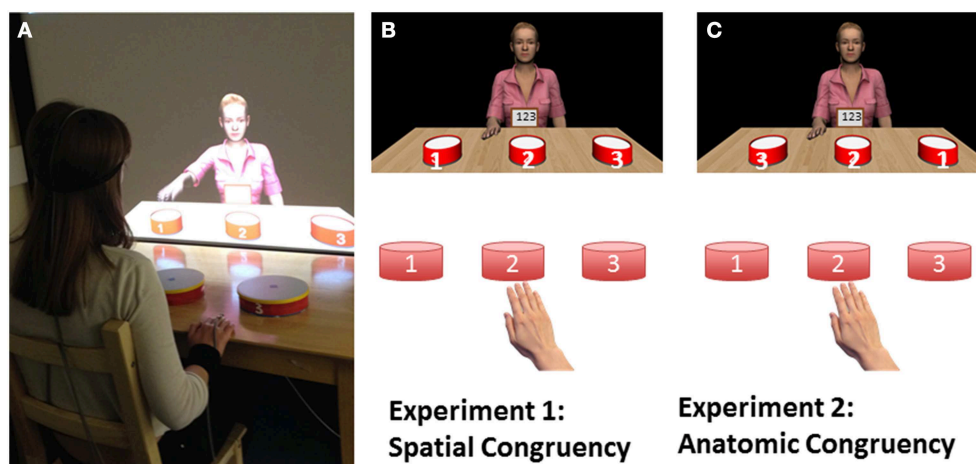


FIGURE 1 | (A) A photo of the study in progress. The participant has a Polhemus magnetic marker on her head and right hand to track her movements. **(B)** Configuration for Experiment 1. The drum numbering and

the participant's responses are spatially congruent between participant and VC. **(C)** Configuration for Experiment 2. The drum numbering and responses are anatomically congruent between participant and VC.

were stored the Cal3D format and applied to Jessie within our interactive VR application developed in Vizard (WorldViz Inc.).

Experimental Design

A 2×2 within participants design was used and the two factors were form (Jessie or balls) and congruency (congruent or incongruent). The number cue, displayed on the virtual iPad during the training and experiment session, consisted of a sequence of three numbers with all possible combinations of 1, 2, and 3, excluding only “1-1-1,” “2-2-2,” and “3-3-3.” This gives 24 possible combinations. Each participant completed 6 blocks (three Jessie, three balls, alternatively): half of them had Jessie as their first block and the other half the balls. Each block consists of 48 trials: 24 congruent trials and 24 incongruent, displayed in random order. Each block could last between 4 and 7 min depending on participants’ speed, and the full set of six blocks could be completed in less than 30 min.

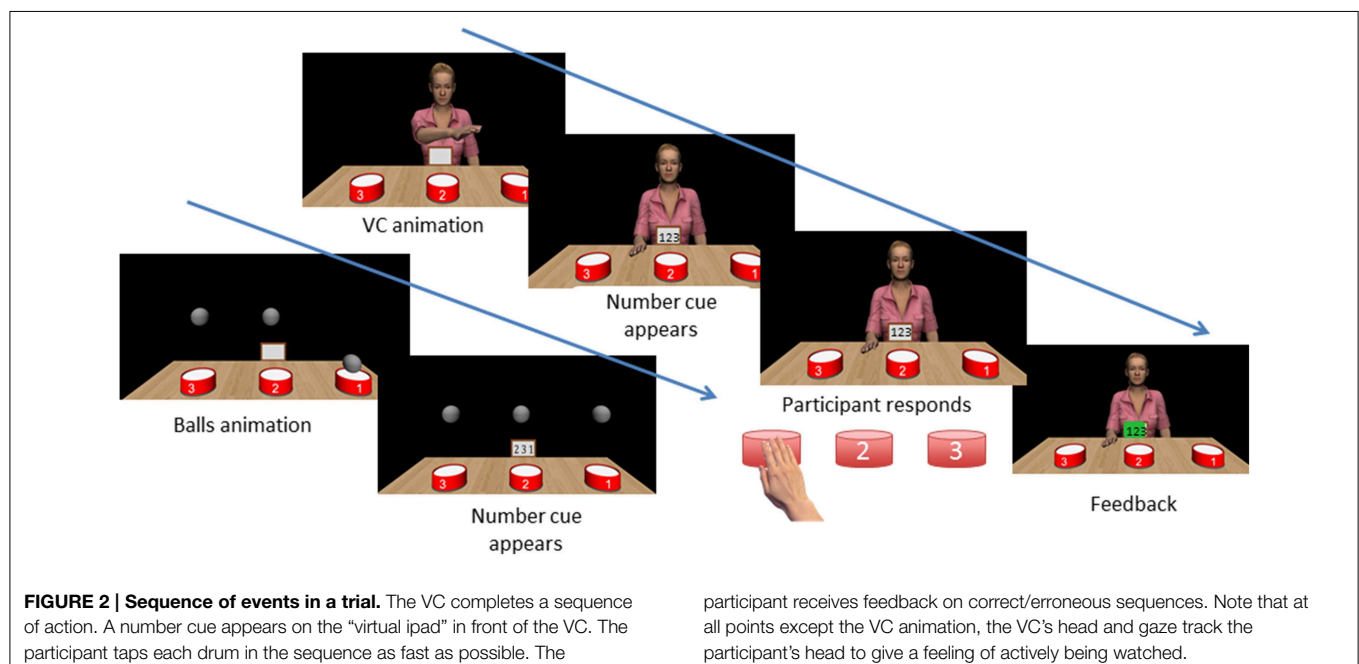
Procedure

On arrival at the lab, each participant was introduced to the VR setup and completed the consent form. Two Polhemus motion-tracking markers were fitted to the participant’s right index finger and forehead with medical tape and a headband. The participant then completed a 5-min calibration and training session for the drum tapping. They were instructed to tap each drum in order as soon as they saw the number cue on the virtual iPad, and should then return to the rest position. Participants practiced this for at least 5 successive correct trials before moving on to the main experiment.

For the main experiment, participants were instructed that they would perform the same drum tapping task, taking turns with Jessie or with some balls. For each trial, the participant first

saw either Jessie or the balls tap a three-beat sequence (e.g., 2-1-3) which lasted approximately 3 s. A drum sound effect played at each point when Jessie or the balls hit each drum. Then the virtual iPad provided a number cue instructing the participant to perform a three-beat sequence (Figure 2). Unbeknownst to the participant, these sequences can be congruent to the action of the VC (e.g., “2-1-3”) or incongruent (e.g., “3-1-1”). In the congruent trials the VC would tap the same spatial locations as the participant i.e., both the physical and virtual drum “1” was at left-hand side of the participant (spatial congruency). Both the participant and Jessie used their right hand, so a reach to drum 1 was a contralateral movement for the participant but an ipsilateral movement for Jessie. This means that the spatially congruent actions were not anatomically congruent. The incongruent animations were designed to be incongruent both anatomically and spatially. For instance, in an incongruent trial, when the participant was cued to tap “2-1-3,” the animation was neither “2-1-3” nor “2-3-1.”

During the participant’s response period, Jessie would “actively watch” the participant. This means that Jessie’s head rotation (left/right, up/down) were programmed so that Jessie was always looking at the participant—if the participant moved slightly left, Jessie looked slightly to the left. This was implemented using the “lookAt” function in Vizard, setting Jessie’s head to orient toward an invisible virtual object whose position was tied to the motion tracker on the participant’s head during participant’s response period, and was tied to the position of the middle virtual drum during Jessie’s tapping session. For transitions between the response period and Jessie’s drumming, the position of the virtual object was updated by linearly interpolating between the two possible positions over 0.5 s. This ensured that Jessie produced smooth, realistic and socially engaging head movements over the whole study. Participants did



not explicitly notice that Jessie was actively watching them during the response phase, but we found it increased the feeling of social engagement and realism.

The motion tracking data collected from the participant's hand was used to monitor performance online. The times when participants touched each drum were defined by the Vizard function "vizproximity" set to detect when the hand marker moved within approximately 1 cm of the center of the drum. The drum-tapping sound effect was played as the participant tapped each drum. Any errors (tapping the wrong drum) resulted in the virtual iPad turning red and a harsh beep sound. When a trial was correctly completed, the virtual iPad turned green. The end of a trial was triggered when the participant's hand returned to the resting location, and the next trial began immediately.

Blocks with ball stimuli were matched in all features, except that Jessie was not present and instead the participant sees three balls suspended above the three drums. To tap a sequence, one ball at a time moved downwards with a constant velocity, tapped the drum and returned to its place. This was implemented using the "moveTo" build-in function in Vizard (**Figure 2** and **Video 1** from Supplement Material).

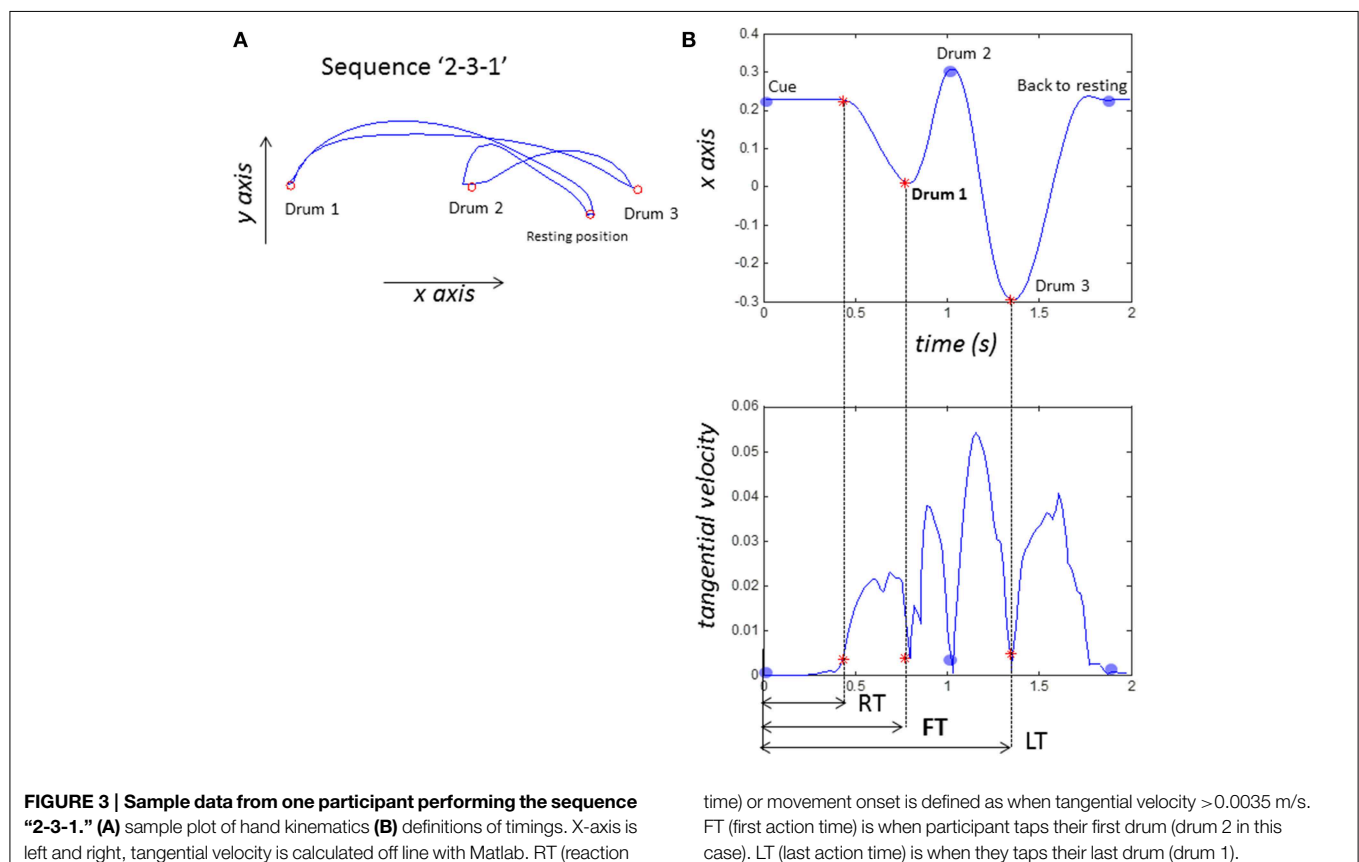
After participants completed all six blocks of the task (three blocks with Jessi and three blocks with balls), they filled in an online questionnaire concerning their subjective evaluation of the experience and of Jessie's personality (see Data Sheet 1 in

Supplementary Material). Participants gave their subjective social evaluation (*SE*) toward the VC through two questionnaires (co-presence and personal trait evaluation). These questions were adapted from previous Virtual Reality studies involving human-VC interactions (Pan et al., 2008, 2015), and here a Likert Scale of 1–7 (1: not at all; 7: very much so) was used. The average score across all 10 questions was used as a covariate in the data analysis. Finally, participants were debriefed and were paid for their time.

Data Analysis

Each participant completed 288 trials equally spread across the following four conditions, with 72 trials in each condition: *congruent-balls* (CB), *incongruent-balls* (IB), *congruent-VC* (CV), and *incongruent-VC* (IV). Two.csv files were produced in real-time with our Vizard application: (1) *event file* contains the time and type of events (e.g., number cue display, participant taps the first drum, and participants' action was correct or incorrect) (2) *tracking file* contains time and motion captured data (position and rotation). In our analysis only position data was used. The following features were extracted for each trial (see **Figure 3**):

- **Reaction time (RT):** The time from the onset of the number cue to the first hand movement, was calculated offline with Matlab. RT was defined as the point when tangential velocity of the finger marker surpassed 0.0035 m/s.



- First drum time (*FT*): The time from the onset of the number cue to the time when the participant tapped the first drum of the sequence of three. This is our primary outcome measure and was recorded in real-time with our Vizard application.
- Last drum time (*LT*): The time from the onset of the number cue to the time when the participant tapped their last drum in the sequence of three; this was recorded in real-time with our Vizard application.
- Errors (*ER*): correct or error in the response.

Each of the four features was averaged at condition level for each participant. For *RT*, *FT*, and *LT*, incorrect trials, or trials where *FT* (our primary measurement) is more than two standard deviations from the mean were excluded from the analysis (4.2%). Data for each of the four features was input to a repeated-measures ANOVA, both with and without mean SE scores as a covariate. Our primary outcome measure was the time to touch the first drum (*FT*) and we report this measure in the text and tables. Other measures (*RT*, *LT*, and *ER*) are presented in **Tables 1–4** only, for completeness.

Results

The mean error rate was 1.2%. SE scores had a mean of 3.45 (SD 1.29).

A repeated measure ANOVA revealed a congruency effect for first drum time [$F_{(1, 21)} = 25.62$, $p < 0.001$, $\eta^2 = 0.550$] indicating that participants were faster in the congruent trials than the incongruent trials. There was no form effect or interaction effect. See **Figure 4A**.

TABLE 1 | Experiment 1: repeated measure ANOVA ($n = 22$).

	Measure	<i>F</i>	<i>P</i>	Partial η^2
RT	cong	10.268	0.004	0.328
FT	cong	25.623	0.000	0.550
LT	cong	13.228	0.002	0.386
ER	cong	6.720	0.017	0.242

TABLE 2 | Experiment 1: Repeated measure ANOVA with SE as covariance ($n = 22$).

	Measure	<i>F</i>	<i>P</i>	Partial η^2
RT	form	5.510	0.029	0.216
	form*SE	10.251	0.004	0.339
FT	form	8.954	0.007 ^a	0.309
	form*SE	13.665	0.001 ^a	0.406
	cong	5.928	0.024	0.229
LT	form	5.558	0.029	0.217
	form*SE	7.355	0.013	0.269
	cong	12.913	0.002 ^a	0.392
	cong*SE	5.567	0.029	0.218

^a The effect is still preserved ($p < 0.05$) after we remove a potential outlier with $SE > 6$.

A repeated measure ANOVA taking participants' SE score as a covariate revealed a similar congruency effect [$F_{(1, 20)} = 5.93$, $p = 0.024$, $\eta^2 = 0.229$]; a form effect [$F_{(1, 20)} = 8.95$, $p = 0.007$, $\eta^2 = 0.309$], indicating that participants reached the first drum faster with the VC than drums; and form-SE interaction [$F_{(1, 20)} = 13.67$, $p = 0.001$, $\eta^2 = 0.406$]. To explore the direction of this effect, we calculated a form effect for each participant as the first-drum-time for the VC minus the first-drum-time for the balls. As shown in **Figure 4B**, the form effect was negatively correlated with SE ($R = -0.637$, $R^2 = 0.406$, $p = 0.001$). This means that the more a participant felt socially connected to the VC, the quicker they reacted to the VC compared to the balls.

Discussion

Our results from Experiment 1 show a main effect of congruency, but no other effect was significant. This can be accounted for by a purely spatial effect, where participants were faster to respond to a particular sequence when they had just viewed a sequence directed toward the same drum locations. This is in line with our prediction that spatial congruency between the participant's drums and the VCs drums should lead to purely spatial effects.

Furthermore, when taking into account participant's reported level of SE of the VC, we found a main effect of form and a form-SE interaction. These results suggest that a social facilitation effect can be obtained using our VC, where participants are faster to respond to a human-like VC than to non-human balls. This is compatible with previous reports of social facilitation to computer generated figures (Hoyt et al., 2003; Zanbaka et al., 2007).

TABLE 3 | Experiment 2: repeated measure ANOVA ($n = 32$).

	Measure	<i>F</i>	<i>P</i>	Partial η^2
RT	form	9.349	0.005	0.232
	cong	11.790	0.002	0.276
FT	form	14.930	0.001	0.325
	cong	13.018	0.001	0.296
	form*cong	5.246	0.029	0.145
LT	form	4.281	0.047	0.121
	cong	7.437	0.010	0.193

TABLE 4 | Experiment 2: repeated measure ANOVA with SE as a covariate ($n = 32$).

	Measure	<i>F</i>	<i>P</i>	Partial η^2
RT	cong	5.290	0.029	0.150
FT	cong	8.325	0.007	0.217
LT	cong	11.673	0.002	0.280
	cong*SE	6.270	0.018	0.173
ER	cong	5.294	0.029	0.150
	cong*SE	6.921	0.013	0.187

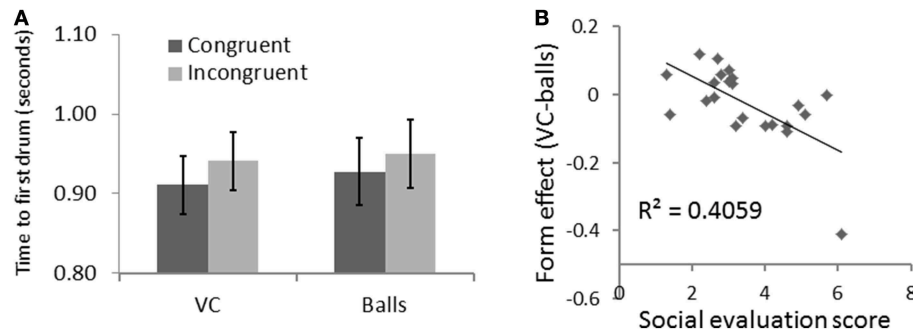


FIGURE 4 | Results for Experiment 1 with SE as a covariate. (A) Time to reach the first drum was faster for VC than for balls. **(B)** This effect correlated with the social evaluation score.

Experiment 2: Anatomic Congruency

Participants

A total number of 32 participants (24 females; mean age = 23.1 years; $SD = \pm 3.73$ years) attended Experiment 2. All were right-handed, had normal or corrected-to-normal vision, and were naïve to the purpose of the study. They received payment at the end of the study. The study was approved by the UCL graduate school ethics committee.

Experimental Design

The experimental design and trial arrangement closely matched Experiment 1. As shown in **Figure 1C**, the only difference was that the virtual drums were displayed in the opposite order as compared to Experiment 1. This means that the participant reaches contralaterally to drum 1, and the VC also reaches contralaterally to drum 1. These movements are anatomically congruent but not spatially congruent. All trials for this experiment were defined in terms of anatomical congruency (not spatial congruency). Note that there was no need to record new animation clips, because the animation clip of the VC playing “2-1-3” in Experiment 1 was the same as that of “2-3-1” in Experiment 2. Instructions, trial structure and trial numbers were identical to Experiment 1. Participant filled the same SE questionnaire and an SE score was calculated. As before, we report analysis over our main measurement (*FT*) in the text and figures, and present data from all measurements (*RT*, *FT*, *LT*, and *ER*) in **Tables 3, 4**.

Results and Discussion

The mean error rate was 1.5%. Again, for *RT*, *FT*, and *LT*, incorrect trials, or trials where *FT* (our primary measurement) is more than two standard deviations from the mean were excluded from the analysis (4.0%). SE scores had a mean of 2.88 (SD 1.16). A *t*-test directly comparing SE scores in Experiment 1 vs. Experiment 2 did not show a significant difference ($p = 0.095$). We speculate that in Experiment 1 it was easier to map spatially between the participant's action and the VC's action on congruent trials, leading the participant to feel similar to the VC. In contrast, in Experiment 2 the participant had to mentally rotate his/her body to the location of the VC to create a strong self-other correspondence on congruent trials. This difference in the ease

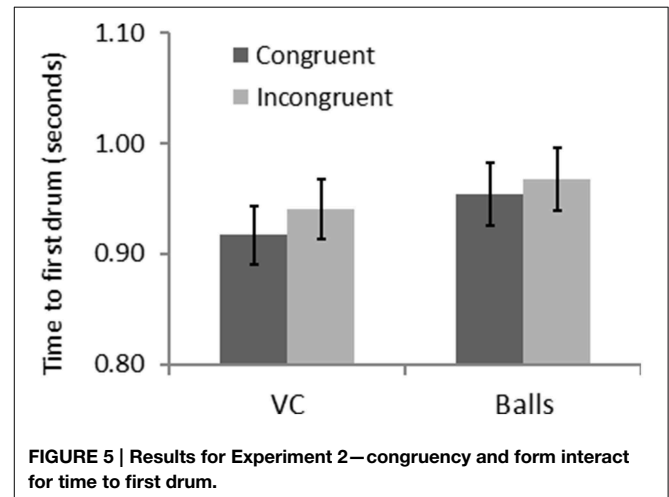


FIGURE 5 | Results for Experiment 2—congruency and form interact for time to first drum.

of self-other mapping between the studies might lead to a trend toward a difference in SE scores. This parallels previous reports that mimicry (mirroring) enhances liking even in VCs (Bailenson and Yee, 2005; Gratch et al., 2007).

A repeated measure ANOVA revealed a significant effect of form [$F_{(1, 31)} = 14.93, p = 0.001, \eta^2 = 0.325$] and congruency [$F_{(1, 31)} = 13.02, p = 0.001, \eta^2 = 0.296$] for *FT*. **Figure 5** shows that participants were faster in the congruent trials, and that they were also quicker with the VC than the balls. There was also an interaction of form and congruency [$F_{(1, 31)} = 5.25, p = 0.029, \eta^2 = 0.145$] for *FT*: the congruency effect is bigger with the VC than the balls. This is consistent with an automatic imitation effect.

A repeated measure ANOVA taking participants' SE score as a covariate preserved the congruency effect for *FT* [$F_{(1, 30)} = 8.33, p = 0.007, \eta^2 = 0.217$] but the form effect and interaction were no longer present. There were also no effects of SE or interactions with SE, suggesting that SE does not add explanatory value to our model but rather reduces power. Thus, we focus our discussion on the basic model without additional covariates.

To summarize, Experiment 2 revealed a congruency effect with faster responses to anatomically congruent trials, and a form

effect, suggesting that participants were faster with the VCs as compared to the balls. More importantly, the reliable form by congruency interaction indicates that participant's automatically imitate the actions of the VC but do not imitate the balls.

General Discussion

In this study, we test if automatic imitation can be obtained in a rich social context with a VC performing sequential actions. We find that spatial congruency effects can be obtained in a context where the virtual drums spatially match the participant's drums (Experiment 1) while automatic imitation can be obtained in a context where the VC's movements anatomically match the participant's movements. These results confirm that automatic imitation can be studied in a richer social context with sequential actions. We consider first the general implications of our novel task and then the specific spatial and anatomical versions of the task.

Measuring Social Congruency with a Drumming Task

In this study we developed a new drumming task for measuring and manipulating automatic imitation. Our task differs from previous tasks (Brass et al., 2000; Stürmer et al., 2000) in at least three ways—it involves sequential actions, it involves goal-directed actions and it is embedded in a rich social context. Sequential actions are an advantage because there are more action options available. This means that our control (incongruent) condition in the sequence task with 3 drums can have neither anatomical nor spatial congruency, thereby providing a better baseline. However, there is also a limitation that automatic imitation can only occur if participants remember the three item sequence from demonstration to the trial. Previous studies of automatic imitation have used simple, single actions (Brass et al., 2000; Stürmer et al., 2000), and it could be argued that the present study does not tap automatic imitation because the action sequences are too complex.

However, there are several reasons to believe that sequential actions can also drive imitation without awareness. Heyes' influential associative sequence learning model of imitation includes action sequences as well as simple stimulus-response associations (Brass and Heyes, 2005). Careful video analysis of natural human behavior also shows copying of action sequences (Grammer et al., 1998). Neuroimaging studies suggest that action sequences and simple actions are stored in a hierarchical format across the cortex (Hamilton and Grafton, 2007). The present study also provides an opportunity to test the hypothesis that action sequences can drive automatic imitation in the same way as simple actions, and provides a positive answer.

A possible limitation of the present study is that verbal encoding of the sequences (both the VCs sequence and the participants sequence) could interfere with the automatic imitation effect. However, verbal encoding would not lead to an advantage in performance that is specific to the human—congruent condition. The fact that we are able to demonstrate an automatic imitation effect (form X congruency interaction in Experiment 2) despite these limitations demonstrates that

observing an action can have a robust and enduring effect on subsequent performance. Future versions of our task may use color cues or other symbols to reduce the likelihood that participants verbally encode the number cues.

Unlike previous tasks, our drumming task is goal directed and each action leads to a noticeable effect (drum sound). This contrasts with the finger-lifting (Brass et al., 2000) and hand-opening tasks (Stürmer et al., 2000) which are not directed toward a particular object. Automatic imitation has also been shown in finger-tapping tasks (Wang and Hamilton, 2013), suggesting that the absence of a goal is not essential for this effect. The present data adds weight to this conclusion, suggesting that even sequential goal-directed actions can lead to an imitation effect. This is also consistent with data showing imitation of kinematic features of sequential pointing actions (Wild et al., 2010), and point to the generality of imitative behavior.

Finally, our new paradigm allows us to study automatic imitation in a very rich social context with an increase in social and ecological validity. It is socially plausible that sometimes you are required to take turns with other person to play drums. The set up of the study could be interpreted as a joint-action, where “two or more individuals coordinate their actions in space and time to bring about a change in the environment” (Sebanz et al., 2006). Similar actions also occur in the context of music (Keller, 2008). This paradigm can therefore offer more direct insights in interpreting automatic imitation or mimicry in everyday social activities and joint actions (Grammer et al., 1998), and provide a bridge between minimal automatic imitation tasks and real-world social psychology mimicry tasks.

To achieve a high level of ecological validity while retaining experimental control, we make use of virtual reality technology to create realistic and interactive VCs. Our experiment was conducted on a large projector screen with life-sized VCs sitting right in front of participants, and that our Virtual Environment was implemented to look like an extension of our real lab. This is key to allowing real-life like social interaction experience and enables participants' automatic social responses. Slater (Slater, 2009) proposed that the two orthogonal components contributing to participants' realistic response in Virtual Reality are *Place Illusion* and *Plausibility Illusion*. In our study, the *Place illusion* was achieved by matching the virtual world to the real world in physical setting the sizes of objects and people, such that participants could believe they were looking through a window into a virtual world. The *Plausibility Illusion* was achieved via realism and interactivity of the VC. Here realism is generated not only by using photo-realistic VC but also by using motion-captured data to animate its movement. Interactivity is achieved by ensuring that VC looks toward the participant during each response period, and that she reacts toward the participants' movement, always waiting for the participant to finish their trial before starting her own. The interactive behavior of our VC, though subtle, is a very important aspect which provides a feeling of social contingency between the participant and VC. The fact that participants' own action and movement could bring about a change in the behavior of another “person” make the whole experience more social and plausible. This is a first step toward “second person neuroscience” (Schilbach et al., 2013).

Spatial, Social, and Imitative Effects on Task Performance

Our two studies allow us to distinguish a number of specific effects on performance. Note that our key performance measure was the time to touch the first drum, which reflects both the planning and initial execution of the action sequence without being contaminated by differences in the movement path. In Experiment 1, participants could be primed by a VC performing a spatially (not anatomically) congruent sequence or by three balls performing a spatially congruent sequence. In this study, we found a clear spatial congruency effect (faster responses on congruent trials). When the participant's SE of the VC was included as a covariate, an effect of form emerged such that those participants who considered the VC to be more human also showed a social facilitation effect and responded faster in the presence of the VC. Social facilitation effects have been demonstrated before for VCs (Hoyt et al., 2003; Zambaka et al., 2004). Here we further show that not all participants react toward Virtual Reality to the same extent, or show the same degree of social facilitation. Individual differences in participant's response to the VC could be caused by many different elements including their personality, their prior experience with virtual reality, and whether the VC's appearance matches themselves. The SE questionnaire used here provides useful information in interpreting our results, and it should also be included in other VR study with VCs.

Our Experiment 2 provides the core test of automatic imitation effects. In congruent trials for this version of the task, the actions of the VC were anatomically congruent with those of the participant, but not spatially congruent. This means that if participants map the VC actions onto their own body, then they will have a performance advantage for the congruent VC trials only, and show a form by congruency interaction. This effect was found, and indicates that participants can automatically imitate the VC. Note that in both Experiments 1 and 2, the action goals (tapping drum number 1, 2, or 3) are congruent for both the VC and the ball trials. The anatomical congruency effect we show here occurs over-and-above any goal congruency effects, because it is present only when the VC performs the action and not when the balls indicate the goals. It is surprising to note that adding

SE as a covariate in the analysis for Experiment 2 did not help us interpret the results. This might imply that automatic imitation is not influenced by the same types of SE as social facilitation, but further studies would be needed to test this systematically.

Future Research Directions

At the present stage, our sequential social congruency task implemented in virtual reality provides a new method to explore automatic imitation in a rich, more ecologically valid setting. One of the advantages of using VC in our stimuli is that in future we can easily adapt our current VR application to test other aspects of social interaction and automatic imitation. For instance, we could test the effect of in-group and out-group by changing the appearance of the VC. We can precisely manipulate the social behavior and emotion of the VC to define how different factors modulate imitation behavior (Wang and Hamilton, 2012). Future studies can also implement this task in a fully immersive virtual world (for instance, with the Oculus Rift) to facilitate the place illusion and enhance the social interaction aspect of participants' experience, and can use VR in conjunction with neuroimaging techniques such as fMRI and functional near-infrared spectroscopy. Overall, we suggest that studying social imitation behavior in rich, well-controlled virtual reality settings is a valuable method for social neuroscience with great promise for the future.

Acknowledgments

This work is funded by the ERC Starting Grant: 313398-INTERACT.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00790/abstract>

Supplementary Video 1 | Participant taking turns with a VC or three balls playing a sequence of three drumming tapping. The VC's gaze is actively tracking the participant's head movement to give a feeling of actively being watched.

References

- Bailenson, J. N., and Yee, N. (2005). Digital chameleons: automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychol. Sci.* 16, 814–819. doi: 10.1111/j.1467-9280.2005.01619.x
- Bertenthal, B. I., Longo, M. R., and Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 210–225. doi: 10.1037/0096-1523.32.2.210
- Bond, C. F., and Titus, L. J. (1983). Social facilitation: a meta-analysis of 241 studies. *Psychol. Bull.* 94, 265–292. doi: 10.1037/0033-2909.94.2.265
- Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn.* 44, 124–143. doi: 10.1006/brcg.2000.1225
- Brass, M., and Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends Cogn. Sci.* 9, 489–495. doi: 10.1016/j.tics.2005.08.007
- Cook, J., and Bird, G. (2011). Social attitudes differentially modulate imitation in adolescents and adults. *Exp. Brain Res.* 211, 601–612. doi: 10.1007/s00221-011-2584-4
- David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., et al. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *J. Cogn. Neurosci.* 18, 898–910. doi: 10.1162/jocn.2006.18.6.898
- Eisenberger, N. I., Lieberman, M. D., and Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292. doi: 10.1126/science.1089134
- Georgescu, A. L., Kuzmanovic, B., Schilbach, L., Tepest, R., Kulbida, R., Bente, G., et al. (2013). Neural correlates of “social gaze” processing in high-functioning autism under systematic variation of gaze duration. *Neuroimage Clin.* 3, 340–351. doi: 10.1016/j.nicl.2013.08.014
- Grammer, K., Kruck, K. B., and Magnusson, M. S. (1998). The courtship dance: patterns of nonverbal synchronization in opposite-sex encounters. *J. Nonverbal Behav.* 22, 3–29. doi: 10.1023/A:1022986608835

- Gratch, J., Wang, N., Gerten, J., Fast, E., and Duffy, R. (2007). "Creating rapport with virtual agents," in *Intelligent Virtual Agents*, Vol. 4722, eds C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, and D. Pelé (Berlin; Heidelberg: Springer), 125–138. doi: 10.1007/978-3-540-74997-4_12
- Greccucci, A., Brambilla, P., Siugzdaitė, R., Londero, D., Fabbro, F., and Rumiati, R. I. (2013). Emotional resonance deficits in autistic children. *J. Autism Dev. Disord.* 43, 616–628. doi: 10.1007/s10803-012-1603-z
- Hamilton, A. F. de C., and Grafton, S. T. (2007). "The motor hierarchy: from kinematics to goals and intentions," in *Attention & Performance 22: Sensorimotor Foundations of Higher Cognition*, eds P. Haggard, Y. Rossetti, and M. Kawato (Oxford: Oxford University Press), 381.
- Heyes, C. (2011). Automatic imitation. *Psychol. Bull.* 137, 463–483. doi: 10.1037/a0022288
- Heyes, C., Bird, G., Johnson, H., and Haggard, P. (2005). Experience modulates automatic imitation. *Brain Res. Cogn. Brain Res.* 22, 233–240. doi: 10.1016/j.cogbrainres.2004.09.009
- Hoyt, C. L., Blascovich, J., and Swinth, K. R. (2003). Social inhibition in immersive virtual environments. *Presence* 12, 183–195. doi: 10.1162/105474603321640932
- Keller, P. E. (2008). "Joint action in music performance," in *Enacting Intersubjectivity: a Cognitive and Social Perspective to the Study of Interactions*, 205–221. Available online at: <http://pubman.mpdl.mpg.de/pubman/faces/viewItemOverviewPage.jsp?itemId=escidoc:723755:2>
- Kiltner, K., Bergstrom, I., and Slater, M. (2013). Drumming in immersive virtual reality: the body shapes the way we play. *IEEE Trans. Vis. Comput. Graph.* 19, 597–605. doi: 10.1109/TVCG.2013.29
- Leighton, J., Bird, G., Orsini, C., and Heyes, C. (2010). Social attitudes modulate automatic imitation. *J. Exp. Soc. Psychol.* 46, 905–910. doi: 10.1016/j.jesp.2010.07.001
- Liepert, R., and Brass, M. (2010). Top-down modulation of motor priming by belief about animacy. *Exp. Psychol.* 57, 221–227. doi: 10.1027/1618-3169/a000028
- Longo, M. R., and Bertenthal, B. I. (2009). Attention modulates the specificity of automatic imitation to human actors. *Exp. Brain Res.* 192, 739–744. doi: 10.1007/s00221-008-1649-5
- Longo, M. R., Kosobud, A., and Bertenthal, B. I. (2008). Automatic imitation of biomechanically possible and impossible actions: effects of priming movements versus goals. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 489–501. doi: 10.1037/0096-1523.34.2.489
- Pan, X., Gillies, M., and Slater, M. (2008). "The impact of avatar blushing on the duration of interaction between a real and virtual person," in *Presence 2008: The 11th Annual International Workshop on Presence*, 100–106. Available online at: <http://discovery.ucl.ac.uk/125118/>
- Pan, X., Gillies, M., and Slater, M. (2015). Virtual character personality influences participant attitudes and behavior – an interview with a virtual human character about her social anxiety. *Front. Robot. AI* 2:1. doi: 10.3389/frobt.2015.00001
- Press, C., Bird, G., Flach, R., and Heyes, C. (2005). Robotic movement elicits automatic imitation. *Brain Res. Cogn. Brain Res.* 25, 632–640. doi: 10.1016/j.cogbrainres.2005.08.020
- Sacheli, L. M., Christensen, A., Giese, M. A., Taubert, N., Pavone, E. F., Aglioti, S. M., et al. (2015). Prejudiced interactions: implicit racial bias reduces predictive simulation during joint action with an out-group avatar. *Sci. Rep.* 5, 8507. doi: 10.1038/srep08507
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/S0140525X12000660
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76. doi: 10.1016/j.tics.2005.12.009
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138
- Stürmer, B., Aschersleben, G., and Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1746–1759. doi: 10.1037/0096-1523.26.6.1746
- Wang, Y., and Hamilton, A. F. de C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Front. Hum. Neurosci.* 6:153. doi: 10.3389/fnhum.2012.00153
- Wang, Y., and Hamilton, A. F. de C. (2013). Understanding the role of the "self" in the social priming of mimicry. *PLoS ONE* 8:e60249. doi: 10.1371/journal.pone.0060249
- Wang, Y., Newport, R., and Hamilton, A. F. de C. (2010). Eye contact enhances mimicry of intransitive hand movements. *Biol. Lett.* 7, 7–10. doi: 10.1098/rsbl.2010.0279
- Wild, K. S., Poliakoff, E., Jerrison, A., and Gowen, E. (2010). The influence of goals on movement kinematics during imitation. *Exp. Brain Res.* 204, 353–360. doi: 10.1007/s00221-009-2034-8
- Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., and Hodges, L. F. (2004). "Effects of virtual human presence on task performance," in *Proceeding of the International Conference on Artificial Reality and Telexistence*, 174–181. Available online at: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.218.8664>
- Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., and Hodges, L. F. (2007). "Social responses to virtual humans: implications for future interface design," in *Proceedings of the SIGCHI Conference on Human factors in computing systems - CHI'07* (New York, NY: ACM Press), 1561–1570.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Pan and Hamilton. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Video stimuli reduce object-directed imitation accuracy: a novel two-person motion-tracking approach

Arran T. Reader¹ and Nicholas P. Holmes^{2*}

¹ School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK, ² School of Psychology, University of Nottingham, Nottingham, UK

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Emma Gowen,
University of Manchester, UK
Elisa De Stefani,
University of Parma, Italy

*Correspondence:

Nicholas P. Holmes,
School of Psychology, East Drive,
University Park Campus, University of
Nottingham, Nottingham,
Nottinghamshire NG7 2RD, UK
npholmes@neurobiology.info

Specialty section:

This article was submitted to
Cognition, a section of the journal
Frontiers in Psychology

Received: 29 January 2015

Accepted: 02 May 2015

Published: 19 May 2015

Citation:

Reader AT and Holmes NP (2015)
Video stimuli reduce object-directed
imitation accuracy: a novel
two-person motion-tracking
approach. *Front. Psychol.* 6:644.
doi: 10.3389/fpsyg.2015.00644

Imitation is an important form of social behavior, and research has aimed to discover and explain the neural and kinematic aspects of imitation. However, much of this research has featured single participants imitating in response to pre-recorded video stimuli. This is in spite of findings that show reduced neural activation to video vs. real life movement stimuli, particularly in the motor cortex. We investigated the degree to which video stimuli may affect the imitation process using a novel motion tracking paradigm with high spatial and temporal resolution. We recorded 14 positions on the hands, arms, and heads of two individuals in an imitation experiment. One individual freely moved within given parameters (moving balls across a series of pegs) and a second participant imitated. This task was performed with either simple (one ball) or complex (three balls) movement difficulty, and either face-to-face or via a live video projection. After an exploratory analysis, three dependent variables were chosen for examination: 3D grip position, joint angles in the arm, and grip aperture. A cross-correlation and multivariate analysis revealed that object-directed imitation task accuracy (as represented by grip position) was reduced in video compared to face-to-face feedback, and in complex compared to simple difficulty. This was most prevalent in the left-right and forward-back motions, relevant to the imitator sitting face-to-face with the actor or with a live projected video of the same actor. The results suggest that for tasks which require object-directed imitation, video stimuli may not be an ecologically valid way to present task materials. However, no similar effects were found in the joint angle and grip aperture variables, suggesting that there are limits to the influence of video stimuli on imitation. The implications of these results are discussed with regards to previous findings, and with suggestions for future experimentation.

Keywords: imitation, two-person, kinematics, grip aperture, joint angles, ecological methods

Introduction

To effectively imitate, visual information about an action must be combined or compared with a representation of the movements necessary to complete the action (Molenberghs et al., 2009). In relation to this, imitation research has often gone hand-in-hand with studies relating to the proposed human “mirror neuron system” (MNS). The MNS provides a potential basis for the ability

to combine visual information with an internal representation of the observed movement. Early research using single cell recording found that neurons in the macaque premotor cortex activated both during the performance of an action, and when the same action was observed in another individual (di Pellegrino et al., 1992; Gallese et al., 1996), hence the term “mirror neurons.” Observed actions are often related to those encoded in one’s own motor repertoire (Oztop et al., 2013), which may in turn provide an insight into the aims of the action and the potential intentions of the observed individual. Much research has aimed to establish the existence of a human MNS—a frontoparietal network that activates during both action observation and performance (Iacoboni et al., 1999), supported by neuroimaging (Molenberghs et al., 2009) and neurophysiological (Naish et al., 2014) evidence. The MNS likely plays a vital role in imitation, and it is possible that imitation relies in part on accurate, unconstrained observation of another’s actions. It follows that any methodology impeding the natural observation of actions is likely to result in less effective understanding of the action, and therefore less effective imitation.

Surprisingly, little research addresses the reliability of video stimuli for experiments on the MNS, social interaction, or imitation. Limb movement is complex and three dimensional, and its observation could be undermined by a 2D viewing set-up (i.e., as observed on flat computer monitors or projection screens). This is particularly worth consideration when much of the research into imitation has used video stimuli presented to a group of solitary observers. There are discrete differences between direct observation of a scene, and observing the same scene reconstructed on a 2D surface (e.g., a computer monitor or projected image). For example, information from binocular disparity in a 3D scene is lost when presented in 2D. The treatment of 2D and 3D stimuli by the visual system varies wildly (Patterson, 2009). Additionally, there is little understanding of how the motor system responds to video vs. real life scenes.

Järveläinen et al. (2001) suggested that video feedback may not be the most appropriate medium for studying social interaction, particularly in an object-directed context. They focused on one proposed element of the human MNS—the primary motor cortex (Hari et al., 1998). Using magnetoencephalography (MEG) they recorded magnetic field signals over participants’ scalps, in two observation conditions: observing a simple right-handed object manipulation performed either by a live actor, or on a pre-recorded video. In a third condition, participants performed the actions themselves. Järveläinen et al. (2001) found that the primary motor cortex showed corresponding activation during both observation and performance of actions. More importantly, they found that this activation was significantly reduced for the observation of video movements compared to live actions. Similar results have been observed in infants (Ruysschaert et al., 2013). Järveläinen et al. (2001) suggested that the difference between video and live feedback reflected the greater ecological validity of the latter and therefore greater participant interest in the 3D visual properties of the action. These results are particularly important considering recent findings suggesting that neural processes in interacting individuals may be “coupled” by contextual parameters (Schippers et al., 2010; Hasson et al.,

2012; Yun et al., 2012). Hasson et al. (2012, p. 115) stated that “the coordination of behavior between the sender and receiver enables specific mechanisms for brain-to-brain coupling unavailable during interactions with the inanimate world.” If we are to measure social interaction, it seems best that we do indeed measure *interaction*, and not just *observation*. If we accept the commentary presented by Hasson et al. (2012), then social interaction is a “live” process, in which both parties are necessary to adequately represent the phenomenon.

Furthermore, most imitation research has used keypress or electromyographic measures from single effectors to measure imitation accuracy. Since muscle activity is only indirectly related to movement kinematics (Knudson, 2007), the above methods may not capture all the information encoded in movement. Perhaps surprisingly, few experiments have used motion tracking to study imitation, and most research has focused on the behavior of the imitator, rather than that of the actor or the interaction between the two. However, movement kinematics may help to inform the observer about an actor’s intent (Becchio et al., 2008; Sartori et al., 2011; but see Naish et al., 2013), and the effect of movement observation on one’s own actions can be so strong as to bias the action toward one more closely representing the observed action, even if imitation is not required (Hardwick and Edwards, 2012).

High-resolution motion tracking might allow greater insights into imitation, so the few studies using this methodology warrant attention. Wild et al. (2010) asked participants to observe videos of actors performing goal-directed and non-goal directed actions at fast and slow speeds and then to imitate the movements. A motion sensor was attached to the index finger and tracked in 3D. The participant’s movement duration, peak velocity, and time to peak velocity, were then compared to the actor’s kinematics. Campione and Gentilucci (2011) also used motion tracking to study whether the automatic imitation of reaching actions is effector dependent. They recorded peak velocity and peak acceleration of the wrist, thumb, and index finger as measures of the effects of action observation on movement planning. These studies extracted relatively few kinematic landmarks from relatively few body positions. A better approach might be to use the whole time-series from as many body parts as possible. The correlation between the time-series data of the actor’s (the one performing the original actions) and the imitator’s movements must necessarily provide a valid measure of imitation effectiveness and therefore a more ecologically valid observation of the imitation process. This was taken into account when designing our experiment.

Also worth consideration is the “correspondence problem” (for a detailed commentary see Brass and Heyes, 2005). It is still unclear how the brain is able to transform the visual parameters of an observed action into a motor output that can match it. This has been put forward as one potential role of the MNS and there is much discussion regarding whether or not it is the intended goal of an action that is imitated, or the entire movement profile itself. In our experiment, the choice was to focus on goal-directed, transitive (object-directed) imitation for two reasons. Firstly, because it allowed us to make use of a more naturalistic, variable task (as explained below), that did not rely on a number of

pre-designated intransitive gestures. Secondly, this study was an attempt to explore the effects seen in experiments making use of object-directed imitation (e.g., Wild et al., 2010; Campione and Gentilucci, 2011; Braadbaart et al., 2012). This sort of imitation closely links to the learning of new motor skills, which occur throughout life, such as learning a new sport. Motion-tracking provides a reliable measure of this sort of imitation, since it can be used to test both object-directed task accuracy (the goal) and the kinematics as a whole.

The aim of the experiment reported here was twofold—first to understand what may be lost in typical transitive imitation paradigms using video feedback, and second to develop the use of motion tracking as a measure for examining imitation in pairs of people. By using face-to-face imitation we hoped to more closely measure imitation as it occurs relatively naturally. As such we developed an imitation game that allowed us to test participants on an object-directed task they are unlikely to have performed before. We recorded position data from 14 motion trackers distributed across the upper body and arms of six pairs of two participants, enabling us to greatly increase the number of tracked body locations compared to previous research. We then compared imitation accuracy in face-to-face feedback, and through a live video projection which prevented the imitator directly observing the actor. We hypothesized that video feedback would result in less accurate imitation than face-to-face feedback, and more complex imitation tasks would result in less accurate imitation than simple tasks. We developed analytic approaches to examine aspects of variable, dynamic time-series to look for correlations and their associated lags with regards to the movement and position of objects in the imitation task.

Materials and Methods

Participants

Twelve right-handed participants (mean \pm SE age = 29.4 ± 7.1 years, 2 male) were recruited from the University of Reading and the surrounding area. The experimental procedures were approved by the local ethics committee (refs: 2013_171_NH; UREC 11/11); participants gave written, informed consent; and the experiments were conducted in accordance with the Declaration of Helsinki. Each experiment required two participants, who took turns to perform as both actor and imitator.

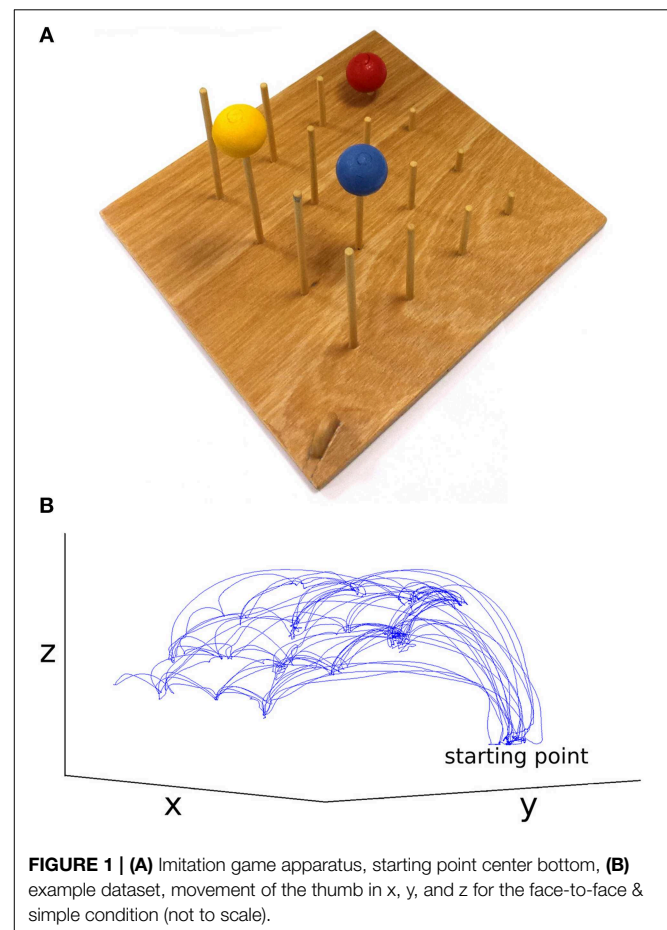
Apparatus and Stimuli

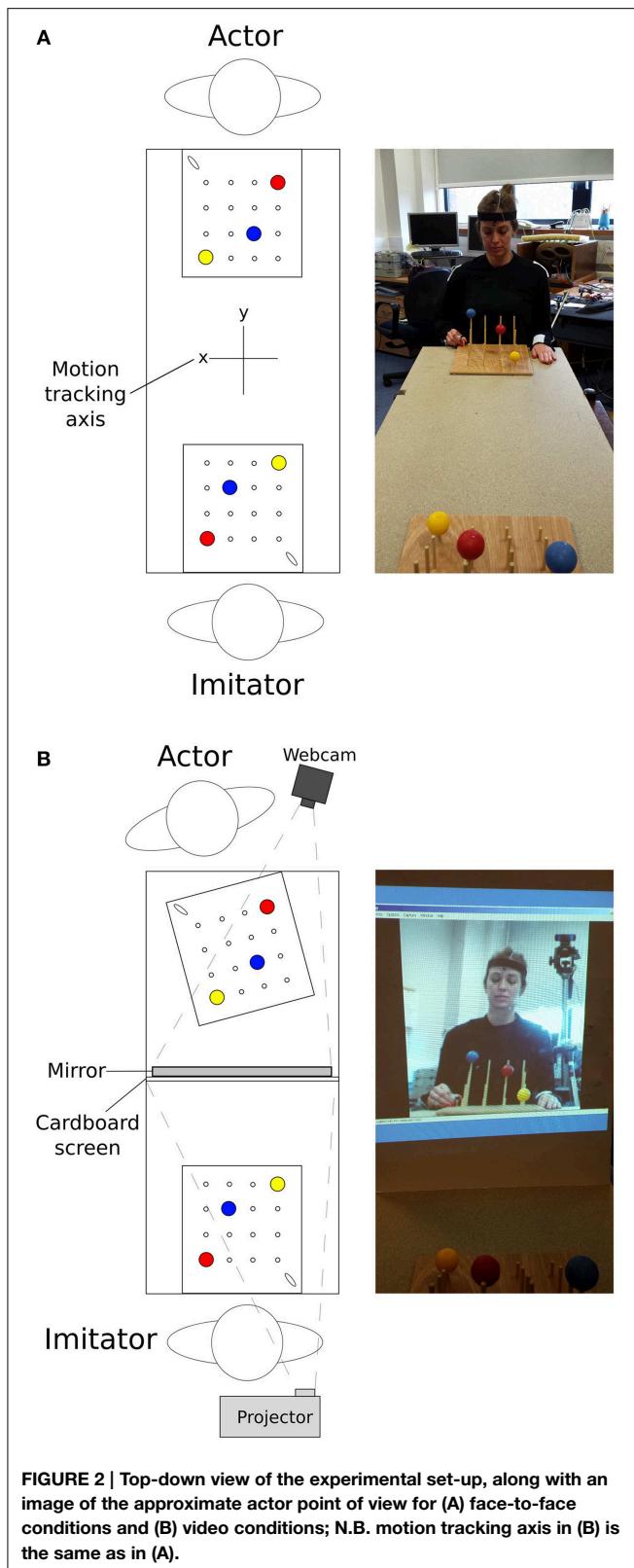
The position of participants' heads, right arms and right hands were recorded continuously using a wired Polhemus Liberty (Polhemus Inc., Colchester, VT, USA) 240 Hz, 14 channel (7 per participant) motion tracking system with 6 degrees of freedom (x, y, z, azimuth, elevation, and roll). Trackers were attached to the shoulder (acromial end of clavicle), elbow (olecranon), wrist (pisiform), thumb (tip), index finger (tip), little finger (tip), and central forehead. Tracking points were attached using adhesive medical tape or Velcro™. The experiment was controlled and data were acquired using custom software written in MATLAB 2014b (Mathworks, Inc.) and using the

ProkLiberty interface (<https://code.google.com/p/prok-liberty/>). We used LabMan and the HandLabToolbox to document and control experiments and analyze data. The associated repositories are freely available at <https://github.com/TheHandLaboratory>, whilst raw data are available from the Hand Laboratory's website (<http://neurobiography.info>) and/or on request.

The stimuli used were two identical custom-designed wooden imitation games consisting of a $300 \times 330 \times 10$ mm board with 4×4 vertical rods (diameter = 5 mm, 60 mm inter-rod spacing, **Figure 1**). The height of the 4 rods from front to back was 30, 70, 110, and 150 mm. On top of three of the rods were three colored (red, blue, yellow) solid cotton balls (diameter = 40 mm), with a 10 mm hole drilled into the center to allow rod placement. A curved wooden starting point of $30 \times 8 \times 25$ mm was situated on the lower right corner near the tallest pegs. These boards were placed facing each other at opposite ends of a table approximately 1370 mm in length, at a distance of 710 mm apart (**Figure 2**). In all conditions the imitation game boards were attached securely to the table using Blu-Tak®. The Polhemus motion tracking transmitter was placed underneath the center of the table (not pictured in **Figure 2**).

Video conditions used a high definition webcam (Logitech International S.A., Switzerland) with a recording resolution of 1080p (resolution of 1920×1080 , before zooming) and frame rate of approximately 30 FPS, to provide a live recording of the





actor. A mirror was placed in front of the actor, angled at 70° to be visible by the camera which was positioned overlooking the actor's shoulder (**Figure 2**). The angled mirror was used to

recreate a flat plane view of the actor in the video feed once the over-shoulder viewpoint was taken into account. A large white cardboard projection screen (840×590 mm) prevented the imitator viewing the actor. The webcam recorded the actor's movements from the mirrored image. This was then projected onto the cardboard screen (image size = 430×580 mm) for the imitator. The image was zoomed to the level that approximately represented the imitator's view of the actor in the face-to-face condition.

Design

A repeated measures design was used, with two independent variables, each with two levels: task difficulty condition (simple, complex) and feedback condition (face-to-face, video). The task difficulty condition was used in order to test whether any effects of feedback condition depended on the complexity of the imitated actions—it was of interest to test whether more complex tasks would be more greatly affected by video feedback. The simple and complex conditions were tested once for each of the video or face-to-face conditions. Each participant played the role of both actor and imitator, meaning that each individual took part in a total of 2 sessions (80 trials)—one as an actor and one as an imitator, to account for two repetitions of the crossed condition design. Whilst using a single individual as the actor may have reduced variability between participants, we wanted to maintain a more naturalistic task with naïve participants, rather than a potentially biased confederate. Each crossed condition lasted 250 s and consisted of ten 20 s trials with 5 s rest gaps between. The dependent variables were the 6 degrees of freedom across 14 motion tracking points.

Procedure

In each testing period, the two participants were assigned to either the role of actor or imitator, which were then reversed once 1 session (4 crossed conditions) was complete. Each testing session included a face-to-face and video feedback condition, and the order in which they occurred was randomized and counterbalanced (i.e., an imitator would observe and imitate in both the video and face-to-face conditions before swapping roles and becoming the actor). Both participants played both roles in order to maximize the data collected and ensure a balanced design. The simple difficulty conditions were always performed first in each feedback condition. This was done in place of a practice trial, in order to cut down testing time and maintain participant motivation and accuracy. Since we predicted that the simple task would be more accurately imitated anyway, we did not believe that this confound would be heavily altered by practice effects. The simple condition ensured that in each of the feedback conditions, the actor and imitator were quickly introduced to the constraints and demands of the task. Note that the main variable studied here is the feedback condition—face-to-face vs. video—the order of which was fully counterbalanced. A live video feed was used in the video feedback condition primarily to cut down on experimentation time, but also to reduce the variability between the feedback conditions to just the effects of video feedback.

In the face-to-face condition, participants sat opposite each other at either end of the table. The imitation boards were placed on the ends of the table in front of each participant who sat approximately 150 mm away. Both participants started with their right index finger and thumb gripping the starting point at the near right hand side of the board. The three balls were randomly distributed across the pegs on the actor's game board at the start of each condition, and the imitator's game board was matched to this. The actor was requested to move balls across the board in two different conditions, whilst the imitator copied the actions in an anatomical fashion (i.e., both participants used their right hand, and a move of the ball to the right by the actor corresponded to a move of the ball to the anatomical right for the imitator), as accurately as possible. Anatomical imitation was used to maintain a more naturalistic imitation task. This is akin to what may happen when one right-handed individual teaches another right-handed individual to perform a motor task, rather than in instances of spontaneous imitation where a mirrored response is more likely to be used by the observer (Pierpaoli et al., 2014).

In the simple condition, the actor freely moved a single ball along 10 consecutive and adjacent pegs moving left or right, or up or down, but not diagonally, touching each peg with the ball before placing it on the peg reached once 10 moves were complete. They then returned to the starting point, gripping it with thumb and index finger. The complex condition also required 10 moves across consecutive pegs, but in this case participants were required to use each of the three balls, in any order as long as a total of 10 moves were made. In each of the crossed conditions, the actor was permitted to move the balls freely within the given parameters of the task, and did not have to perform the same movement sequence across different conditions. Both the actor and the imitator were informed of the constraints of the actor's task. A beep played through the computer's speakers signaled the actor and imitator to begin and finish at the start and end of each 20 s trial. Participants were requested to make the most of the total 20 s, timing their 10 moves accordingly. Participants always moved back to the start point once their moves were complete. Example data are shown in **Figure 1**. Imitators were requested to copy the actor's movements as accurately as possible. They were asked to begin imitating the actor as soon as the actor started moving. No instructions were given to either participant regarding eye gaze.

The tasks in the video feedback condition were identical, except that the imitator observed the actor through a live video projection, and any natural vision of the actor was obscured by the cardboard screen (**Figure 2**). For the actor, the angle of the imitation game was shifted by 13° anticlockwise and the apex of the mirror was placed 570 mm from the edge of the table, with the reflective side facing the actor. The actor was then seated facing the game board at the same distance and orientation as in the video condition (i.e., directly facing the board, sat approximately 150 mm away). These changes allowed the webcam (angled appropriately) to record the actions of the actor, passing the video on to an image projected on to the card screen mounted on the back of the mirror, 640 mm away from

the imitator. The imitator could perform the required actions without direct observation of the actor.

At the start of each video or face-to-face condition, a brief calibration test was run. This required the actor to trace the outside of the imitation game board with their thumb and index finger, following a tone. The imitator was requested to copy this action. The calibration enabled the experimenter to ensure that all trackers were recording correctly and that there were no obvious distortions in the data prior to data collection.

Data Pre-Processing

Five pre-processing steps were performed in order to clean the data. First, single time-point spikes (>3 SD from the mean) in each variable's double-differentiated time-series (i.e., acceleration) were deemed electromagnetic artifacts and removed by interpolation across two adjacent samples either side. Second, the position data were filtered using a bidirectional low-pass 4th order Butterworth filter (cutoff frequency 15 Hz). Third, the position data for the actor in the video condition were rotated by 13° clockwise in the x ($x = x(\cos 13) - y(\sin 13)$) and y ($y = y(\cos 13) + x(\sin 13)$) axes in order to correct for the angled game board.

Fourth, the time-series for the imitator data in the video condition was shifted backwards by 111 ms to account for the latency between the recording and presentation of video stimuli, ensuring that any effects of the video condition were due to the condition itself rather than the delay in stimulus presentation. Latency was calculated by measuring the time difference on an independent PC using Chart 5 software to detect a flash of light presented to two light detecting diodes—one located at the webcam aperture, the second located on the cardboard screen used to project video stimuli. Diodes were connected via a custom interface to an AD Instruments data acquisition unit sampling at 2 kHz. Video latency (the time between light detection in each of the two diodes) was measured over 25 discrete tests (whilst the data collection script was running in the background to simulate the experimental condition), resulting in a mean \pm SD latency of 111 ± 25 ms.

Finally, since data collection was continuous during the entire length of the condition (including rests) and actors often finished their 10 movements before the end of the (20 s) trial time, the lengths of each trial were calculated independent of the total trial time. This was done by defining correct trials (i.e., ignoring false starts) as >100 mm movement of the index finger away from the start point for any period >5 s. This ensured that false starts were excluded from the analysis, and trial onsets were timed to the actors' movements. These variable trial times were also applied to each actor's associated imitator's data, since imitators were requested to begin movement at the same time as the actor.

Exploratory Data Analysis

Prior to full data processing, an exploratory analysis of one half of the data (3 pairs of participants) was performed. This was deemed necessary due to the novel methods developed in this experiment, as well as the potential for false positives with such a large dataset and so many dependent variables. We hoped that it would reveal any consistent effects across degrees of freedom, and

direct our choice of final analysis parameters based on this. Each crossed condition (task difficulty \times feedback condition) yielded 42 dependent variables for each participant (84 in total): 7 motion trackers \times 6 degrees of freedom (x, y, z, azimuth, elevation, roll).

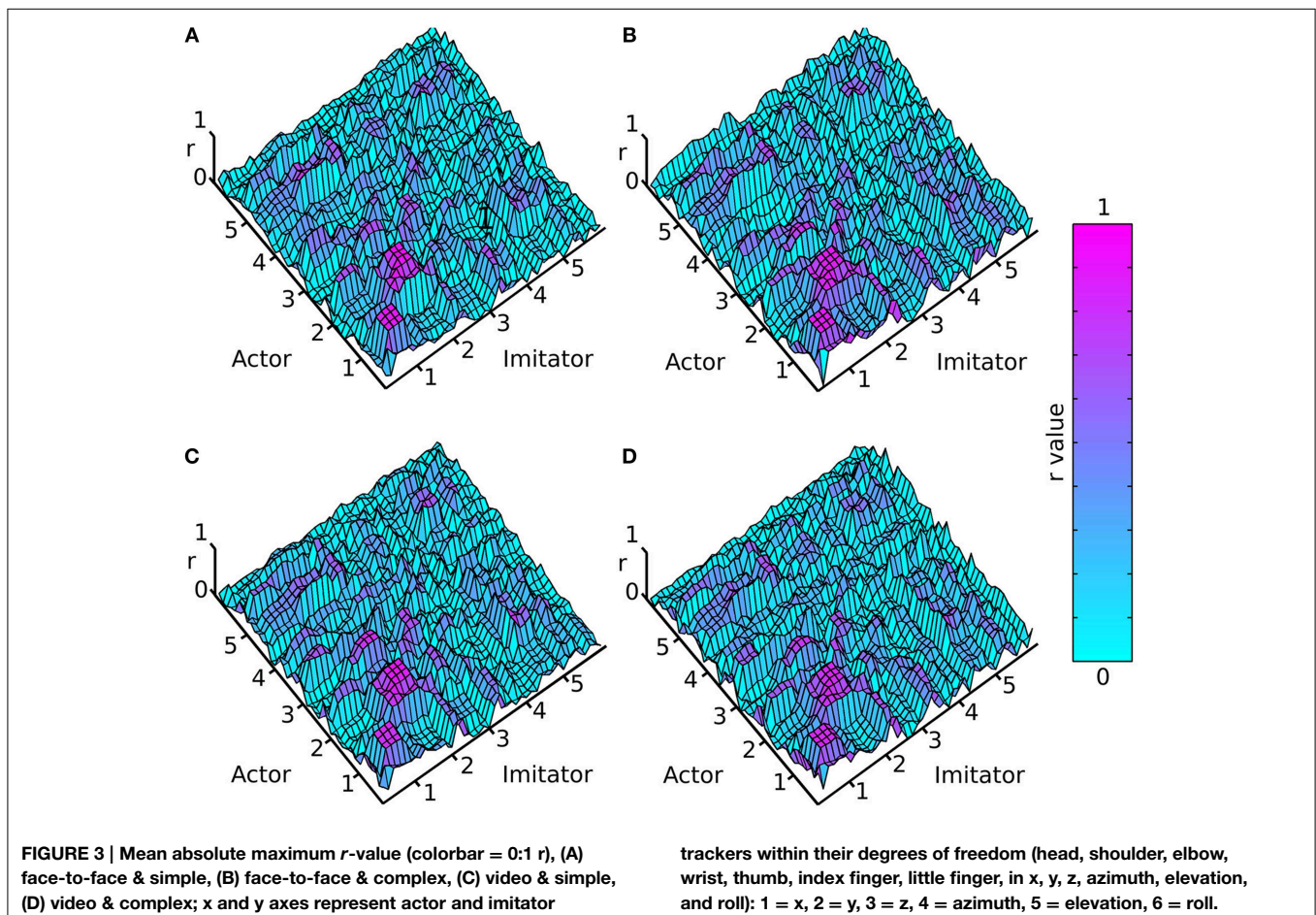
A cross-correlation was performed on each of the crossed conditions over each of the 10 (variable length—see data pre-processing) trials. This was done by shifting the imitator's data relative to the actor's sample by sample over lags of -5 to $+5$ s, and correlating the two time-series for each lag (-1200 to $+1200$ samples). For each of these 10 trials, an absolute maximum r -value between each actor dependent variable and each imitator dependent variable was generated, along with the lag associated with that maximal r -value (as a measure of the best-fitting overall lag between actor and imitator). The lag at maximum r represents the difference (in time) between the actor and imitator datasets at the point at which the maximum r -value was found. These results were averaged across the 10 action trials per participant and then across the 6 participants to generate the surface plots in **Figures 3, 4**.

The surface plots suggested that absolute maximum r -value and the lag associated with it varied widely across dependent variables. The most consistently highly correlated values were the corresponding trackers in their corresponding degrees of freedom. This was emphasized by the highly correlated diagonal contours in the surface map of r -values in **Figure 3** (particularly

in x, y, and z). The greater density of pink coloring in the face-to-face condition r -value plots seemed to suggest that it may be better correlated than the video condition; however it was hard to gauge any large differences between correlations in the difficulty conditions. The surface plots in **Figure 4** suggested that the lag associated with the maximal r -value was, surprisingly, lower in the complex vs. the simple conditions. It also appeared that the video conditions may have had slightly lower lags than the face-to-face conditions, though this was less clear.

Final Analysis Parameters

Based on the exploratory analysis it was decided that an analysis of the entire dataset (12 participants) would benefit from parameters that capture the greatest movement information in the fewest dependent variables. As such, we decided to focus on three elements of the task: joint angles in the arm, grip aperture, and grip position, each of which were calculated for actor and imitator. This analysis was performed on all 12 participants' data. Joint angles of the arm were selected because the angles of all the joints in any given effector across time provide a general representation of the whole movement. Thus, by examining the joint angles between the trunk, shoulder, elbow and wrist, it was possible to develop a reasonably accurate measure of the entire arm movement. This would enable us to compare kinematic, rather than goal outcome accuracy of the imitator.



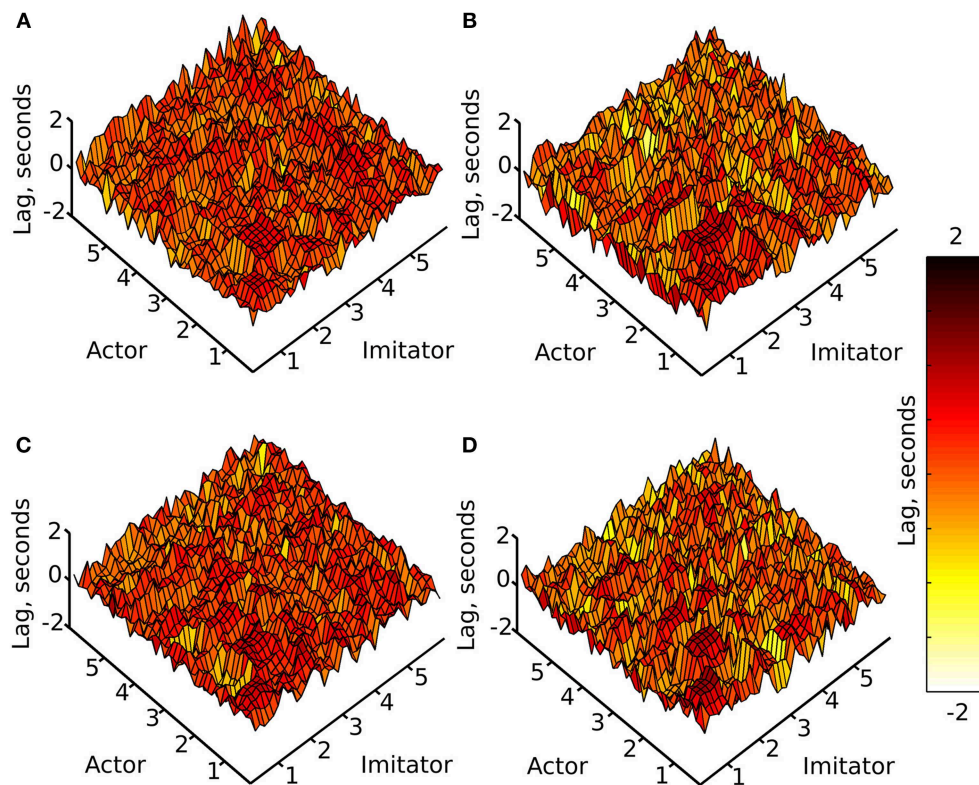


FIGURE 4 | Mean lag at absolute maximum r -value (colorbar = $-2 : +2s$), (A) face-to-face & simple, (B) face-to-face & complex, (C) video & simple, (D) video & complex; x and y axes represent actor and imitator trackers

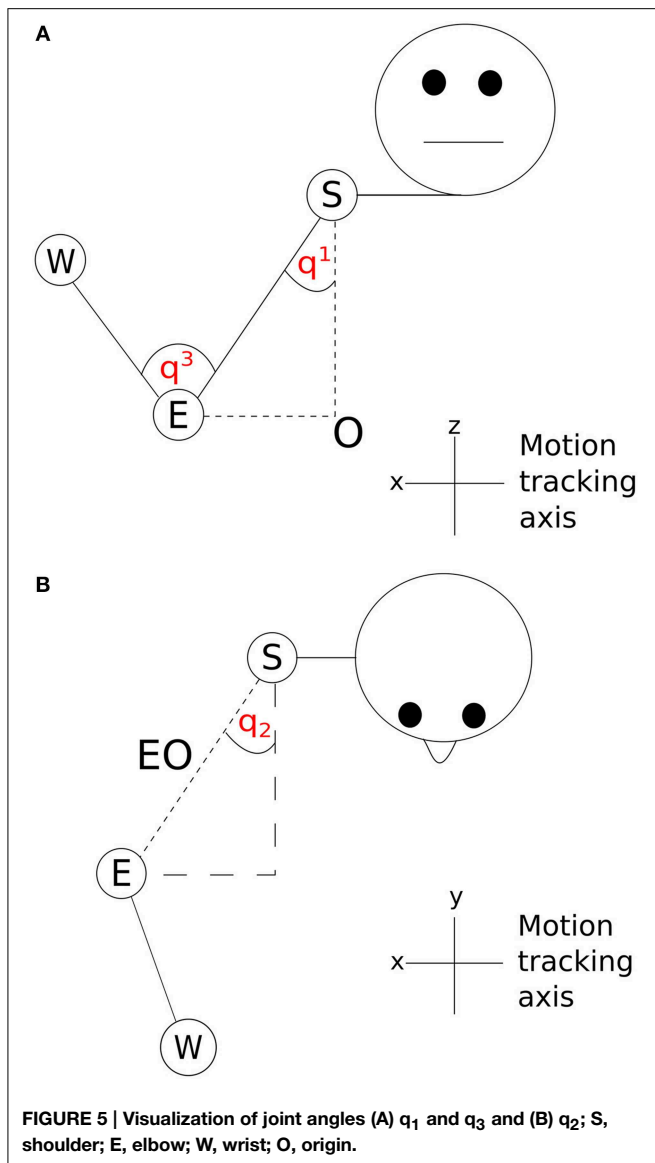
within their degrees of freedom (head, shoulder, elbow, wrist, thumb, index finger, little finger, in x, y, z, azimuth, elevation, and roll): 1 = x, 2 = y, 3 = z, 4 = azimuth, 5 = elevation, 6 = roll.

The two angles between the shoulder and the body in the x and y dimensions (q_1 and q_2) are shown in **Figure 5**. A vector **SO** starting at the shoulder, S and ending at the origin, O was determined by subtracting the z dimension position value of the elbow from the z dimension position value of the shoulder. By using this vector along with the elbow-shoulder vector **ES**, a right angle triangle was formed. Angle q_1 was calculated as the angle between vectors **ES** and **SO** ($q_1 = \cos^{-1} \left(\frac{z_{\text{shoulder}} - z_{\text{elbow}}}{ES} \right)$). A projection of the vector **EO** between the elbow and origin was created in the x and y dimensions. In the x and y dimension a second right angle triangle was created using the vector **EO** and a second vector calculated by subtracting the y dimension position of the elbow from the y dimension position of the shoulder. q_2 was calculated as the angle between **EO** and this second vector ($q_2 = \cos^{-1} \left(\frac{y_{\text{shoulder}} - y_{\text{elbow}}}{EO} \right)$). The inner elbow angle q_3 (**Figure 5**) was calculated through the cosine rule, taking the elbow-to-wrist **EW** and elbow-to-shoulder **ES** as two intersecting vectors ($q_3 = \cos^{-1} \left(\frac{SW^2 - ES^2 - EW^2}{2(ES \times EW)} \right)$). Using joint angles in this way reduced the number of position parameters to examine from nine (3 tracking points \times 3 axes) to three (3 angles, q_1 – q_3).

We also used the grip aperture of the index finger and thumb. Grip aperture is a commonly recorded parameter in kinematics

(Castiello and Ansuini, 2009), and provides a measure of the primary movement required for this task. The grip aperture variable was created by calculating the 3D distance between the index finger and the thumb. Finally, the grip position was recorded. This was done by taking the mean location of the index finger and thumb in x, y, and z. We hoped that this would provide a general measure of task imitation accuracy, rather than movement imitation accuracy, since some authors have claimed that it is the goals of an action that are imitated, rather than the means (Wohlschläger et al., 2003).

These three new DVs were cross-correlated in an identical manner to the exploratory analysis, resulting in absolute maximum r -values and their associated lags for each of the trials across each of the crossed conditions. For participants 11 and 12, the final trial of the complex face-to-face condition was excluded due to the actor's (participant 11) failure to return their hand to the starting point. The means of the r -values and lags across trials was calculated to provide 7 DVs (q_1 , q_2 , q_3 , grip aperture, grip position in x, y, and z) for each participant across the two experimental conditions. For each of these new DVs mean r -values between participants across the 10 trials per crossed condition were converted to Z-values using the Fisher transformation ($Z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$), where \ln is the natural



logarithm of a number. This allowed parametric statistics to be used on the r -values.

Results

Repeated measures MANOVAs were run on the Z -values and lags at absolute maximum r -value, for joint angles (q_1 – q_3) and grip position (x , y , z). A Two-Way repeated measures ANOVA was run on the Z -values and lags at absolute maximum r -value for the grip aperture values. The MANOVAs and ANOVA compared the mean Z -value and mean lag of the 10 trials between the feedback and difficulty conditions across all 12 experiments (24 sessions). The results of the MANOVAs are given in **Tables 1, 2**, and mean values are shown in **Figures 6, 7**.

The MANOVA on Z -values (**Table 1**), measuring the strength of correlation between actor and imitator, revealed 5 significant

effects. Both the x [$F_{(1, 11)} = 9.41$, $p = 0.011$, partial $\eta^2 = 0.461$] and y [$F_{(1, 11)} = 6.77$, $p = 0.025$, partial $\eta^2 = 0.381$] grip positions showed a significant effect of feedback, with the face-to-face condition more highly correlated than the video condition (mean \pm SE difference in Z -values = 0.179 ± 0.058 for x , and 0.145 ± 0.056 for y data), providing some support in favor of our hypothesis. The mean Z -values for x were equivalent to r -values of 0.889 for face-to-face feedback and 0.845 for video feedback. For y the equivalent r -values were 0.907 for face-to-face feedback and 0.878 for video feedback. Both the x [$F_{(1, 11)} = 6.27$, $p = 0.029$, partial $\eta^2 = 0.363$] and y [$F_{(1, 11)} = 13.8$, $p = 0.003$, partial $\eta^2 = 0.557$] grip positions showed significant effects of task difficulty, with the simple condition more highly correlated than the complex (mean \pm SE difference in Z -values = 0.158 ± 0.063 for x , and 0.215 ± 0.058 for y). The mean Z -values for x were equivalent to r -values of 0.887 for simple task difficulty and 0.848 for complex task difficulty. For y the equivalent r -values were 0.913 for simple task difficulty and 0.870 for complex task difficulty. These two significant univariate effects also resulted in a significant multivariate effect in multivariate grip position for task difficulty, $F_{(3, 9)} = 7.32$, $p = 0.009$, partial $\eta^2 = 0.709$. The mean Z -values for this multivariate variable were equivalent to r -values of 0.856 for simple task difficulty and 0.811 for complex task difficulty.

The lag MANOVA (**Table 2**) revealed 4 significant effects. There was a significant effect of feedback in joint angle q_2 , $F_{(1, 11)} = 5.57$, $p = 0.038$, partial $\eta^2 = 0.336$, with the video condition showing a longer delay than the face-to-face (mean \pm SE difference = 0.302 ± 0.128 s). The multivariate grip position was significant for task difficulty, $F_{(3, 9)} = 3.95$, $p = 0.047$, partial $\eta^2 = 0.586$, with the complex condition significantly more delayed than the simple (mean \pm SE difference = 0.155 ± 0.053 s). The y grip position also showed a significant effect of task difficulty, $F_{(1, 11)} = 10.7$, $p = 0.007$, partial $\eta^2 = 0.494$, with the complex condition significantly slower than the simple (mean \pm SE difference = 0.178 ± 0.054 s). Finally, there was a significant interaction between task difficulty and feedback in the x grip position, $F_{(1, 11)} = 5.93$, $p = 0.033$, partial $\eta^2 = 0.350$, where simple conditions showed longer imitation lags than complex when observed face-to-face (mean \pm SE difference = 0.031 ± 0.087 s), but imitation in the complex conditions was later than the simple when observed via video (mean \pm SE difference = 0.268 ± 0.087 s).

Discussion

We examined the effects of face-to-face vs. video feedback on imitation in a transitive imitation task, hypothesizing that video feedback would result in less accurate imitation and that a simpler task would result in more accurate imitation than a complex one. After running an exploratory analysis, we chose to perform a more focused statistical analysis on grip position, joint angles in the arm, and grip aperture.

In the correlation (Z -value) analysis, only the grip position variables revealed significant effects of feedback and task complexity. Grip position can be taken as a general measure of accuracy in our imitation task, since it measures the position of

TABLE 1 | MANOVA and ANOVA main effects and interactions for mean Z-value.

Mean(±SE) Z-value					Main effect				Feedback*difficulty interaction					
Feedback			Task difficulty											
Face-to-face		Video	Simple	Complex	Feedback		Difficulty		F*		Partial η ²			
					F*	p	Partial η ²	F*	p	F*	p	Partial η ²		
Joint angle	Multivariate	0.791 (0.062)	0.770 (0.045)	0.797 (0.052)	0.764 (0.055)	1.27	0.344	0.297	1.29	0.337	0.300	0.916	0.471	0.234
	q1	0.934 (0.111)	0.980 (0.067)	0.947 (0.083)	0.967 (0.092)	0.327	0.579	0.029	0.105	0.751	0.009	0.245	0.631	0.022
	q2	0.995 (0.067)	0.894 (0.061)	0.984 (0.074)	0.905 (0.051)	2.16	0.170	0.164	1.39	0.264	0.112	1.90	0.195	0.147
	q3	0.443 (0.085)	0.436 (0.058)	0.458 (0.061)	0.421 (0.065)	0.008	0.930	0.001	0.866	0.372	0.073	0.014	0.909	0.001
		1.26 (0.041)	1.14 (0.029)	1.28 (0.045)	1.13 (0.035)	3.32	0.071	0.529	7.32	0.009	0.709	0.145	0.930	0.046
Grip position	x	1.42 (0.047)	1.24 (0.061)	1.41 (0.057)	1.25 (0.054)	9.41	0.011	0.461	6.27	0.029	0.363	0.126	0.729	0.011
	y	1.51 (0.043)	1.37 (0.029)	1.55 (0.040)	1.33 (0.035)	6.77	0.025	0.381	13.8	0.003	0.557	0.216	0.651	0.019
	z	0.850 (0.079)	0.819 (0.038)	0.874 (0.076)	0.795 (0.057)	0.198	0.665	0.018	0.842	0.379	0.071	0.025	0.887	0.002
Grip aperture		0.186 (0.063)	0.142 (0.072)	0.180 (0.043)	0.149 (0.074)	0.310	0.589	0.027	0.389	0.545	0.034	0.165	0.693	0.015

* Degrees of freedom for multivariate analyses are (3, 9), for univariate (1, 1); significant p-values (<0.05) are in bold.

TABLE 2 | MANOVA and ANOVA main effects and interactions for mean lag.

Mean lag(±SE) in seconds				Main effect				Feedback*difficulty interaction	
Feedback		Task difficulty		Feedback		Difficulty			
Face-to-face		Video	Simple	Complex	F*	p	Partial η^2	F*	p
									Partial η^2
Joint angle	Multivariate	0.485 (0.096)	0.523 (0.085)	0.572 (0.108)	1.87	0.206	0.383	2.67	0.110
								0.508	0.687
									0.145
	q ₁	0.722 (0.106)	0.711 (0.080)	0.831 (0.096)	0.590	0.458	0.051	1.43	0.257
								0.578	0.463
									0.050
	q ₂	0.466 (0.120)	0.447 (0.112)	0.789 (0.138)	5.57	0.038	0.336	4.48	0.058
								1.56	0.237
									0.124
	q ₃	0.267 (0.198)	0.412 (0.141)	0.095 (0.203)	0.016	0.902	0.001	2.36	0.152
								0.823	0.384
									0.070
Grip position	Multivariate	0.666 (0.071)	0.652 (0.051)	0.807 (0.069)	0.673	0.590	0.183	3.95	0.047
								2.24	0.153
									0.427
	x	0.734 (0.062)	0.742 (0.060)	0.861 (0.076)	1.69	0.22	0.133	3.70	0.081
								5.93	0.033
									0.350
	y	0.712 (0.077)	0.674 (0.040)	0.852 (0.078)	1.70	0.219	0.134	10.7	0.007
								2.80	0.122
									0.203
	z	0.553 (0.138)	0.540 (0.098)	0.707 (0.114)	0.51	0.49	0.044	1.38	0.265
								0.122	0.733
									0.011
Grip aperture		0.043 (0.201)	0.275 (0.171)	−0.178 (0.262)	0.003	0.955	<0.001	3.55	0.086
								0.163	0.694
									0.015

* Degrees of freedom for multivariate analyses are (3, 9), for univariate (1, 11); significant p-values (<0.05) are in bold.

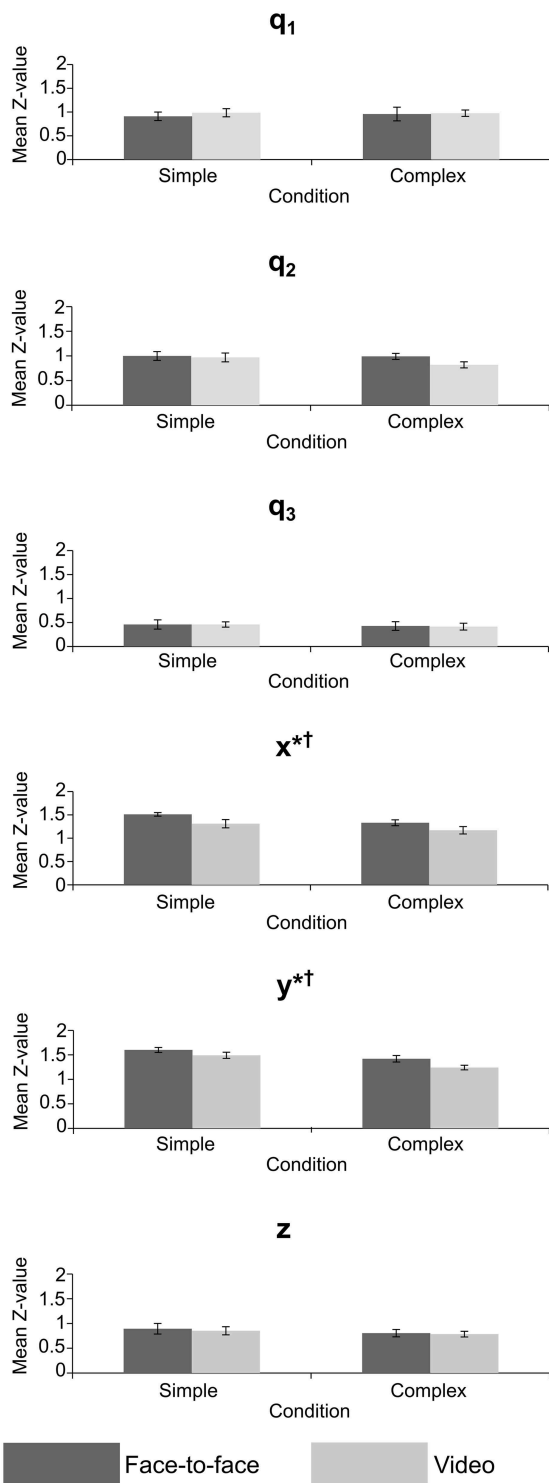


FIGURE 6 | Mean Z-value for crossed conditions in the joint angle and grip position variables, error bars = standard error; *significant effect of feedback, †significant effect of task difficulty. $p < 0.05$ for all significant effects—see Table 1 for exact values.

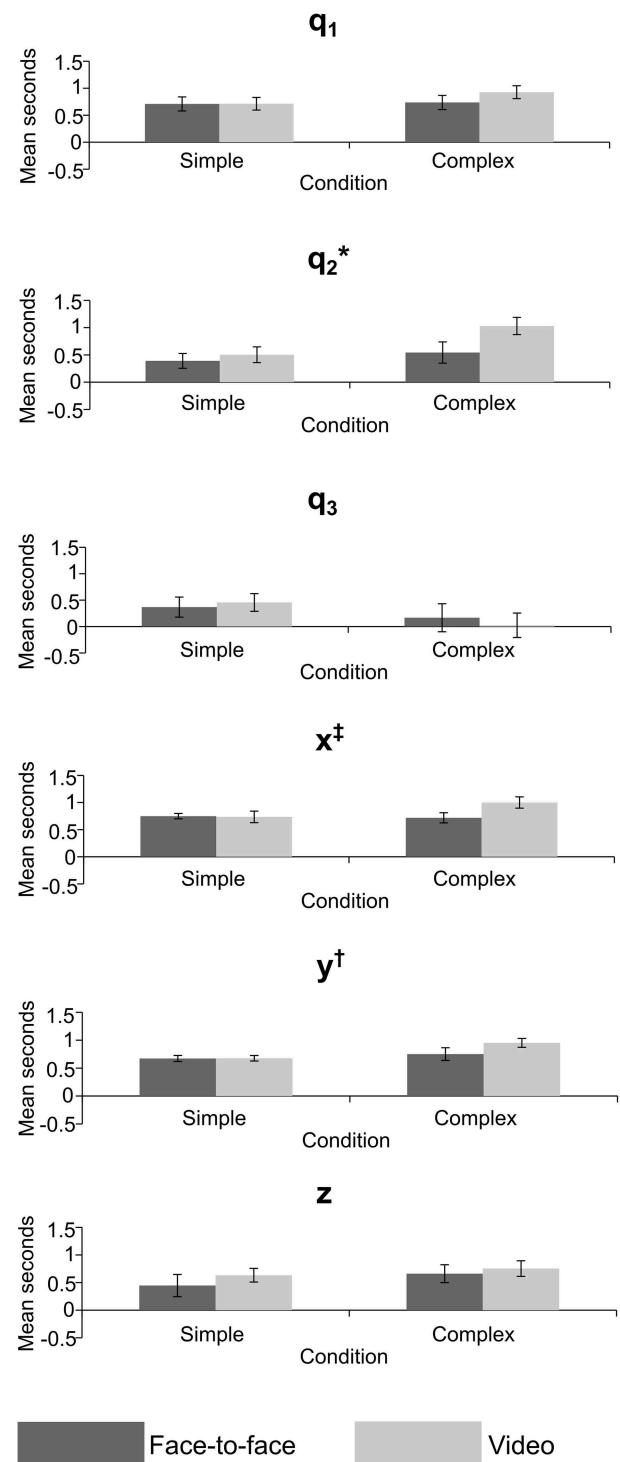


FIGURE 7 | Mean lag in seconds for crossed conditions in the joint angle and grip position variables, error bars = standard error; *significant effect of feedback, †significant effect of task difficulty, ‡significant interaction. $p < 0.05$ for all significant effects—see Table 2 for exact values.

the object effectors (index finger and thumb) from the starting point, across the movement of the balls, and then the return of the hand to the starting point. The significant differences suggested that video feedback reduced the accuracy of transitive imitated actions for left-right (x) and forward-back (y) dimensions of motion, but not for up-down (z). This supports our hypothesis that video feedback would be less highly correlated than face-to-face observation. Imitators were worse at completing the imitation task when required to view the actor through a live video feed. The source of this effect is most likely the difference in visual information provided by the video and face-to-face feedback conditions, but it is also possible that increased motivation driven by the ecological validity of the face-to-face condition is responsible (Järveläinen et al., 2001). However, the continued presence of the actor in the room during both feedback conditions suggests either that this explanation is lacking, or that such an effect may be strong enough to compensate for the imitator's knowledge about the actor's location. These are important findings when considering previous imitation research that has used video stimuli, particularly for studies using object-directed actions. At the very least these studies have not accounted for the effect of visual feedback and may be lacking in ecological validity. It is likely that imitation was altered in these studies, with accuracy being reduced by video feedback.

Comparing simple and difficult tasks, the forward-back and left-right dimensions of grip position also showed significant effects, with the simple task more highly correlated than the complex one, suggesting our manipulation of task difficulty was effective. The lack of significant interactions between feedback and difficulty in the correlation analyses suggests that the effects of face-to-face vs. video feedback were not affected by task complexity.

Despite the significant results in the grip position analysis, grip aperture and joint angles showed no such effects. This may be the result of imitators copying the motion of the ball (the goal), but failing to imitate the broader motion of the actor's arm. This is likely due to our use of a transitive task, and may lend credence to claims that transitive imitation is primarily goal-directed, and that it is the object of the goal that is imitated, rather than the associated body movements (Wohlschläger et al., 2003; but see Leighton et al., 2010). However, a number of other factors may have influenced this outcome. It may also be due to our use of anatomical, rather than mirror imitation, or the fact that imitators had to shift their attention between the actor's game board and their own, thus limiting the resources available to imitate movements outside of the task constraints. In addition, grip aperture showed no effects of feedback or difficulty. This may be because the proportion of time that grip aperture was changing was too low to detect significant effects. When both actor and imitator were holding a ball, there was no longer a time-varying correlation between their (constant) grip apertures.

What remains to be explained from the correlation analysis is why the grip position in the forward-back and left-right directions were significant, whilst up-down was not. One explanation is that up-down movements were not influenced by the effects of the video condition. Certainly up-down movements of the balls were clearer to observe in the video condition than

forward-back. Movements forward-back were hard to distinguish in the video condition without depth information (i.e., pegs that were lined up in front of each other were less distinguishable compared to those going left to right). However, the up-down effects were in the same direction as other dimensions (**Figure 6**), suggesting that the effect was too weak to be detected. The absence of significant effects for joint angles and grip aperture may indicate that some aspects of object-directed imitation are not strongly affected by video feedback. Eye-tracking could have been useful in this respect. Measurement of imitator eye movements could have shown whether they were concentrating on the actor's movements in general, rather than the end point of the ball (the goal).

The results of the lag analysis were less consistent than the correlation analysis. The most interesting result was for joint angle q_2 —the rotation of the upper arm about the shoulder—where imitation was significantly later in the video than face-to-face condition. This may be related to the reach-to-grasp action, and the difference in lag between face-to-face and video conditions may reflect a delayed approach toward the balls by the imitator. This could again be related to the ecological validity or motivation in the video condition. The significant multivariate effect for grip position suggests that overall, imitators acted later to accurately imitate the ball movements in the complex condition. The same effect was also shown in univariate analysis for the forward-back movements, meaning that they were imitated more slowly in complex tasks, potentially reflecting a greater use of this dimension in complex tasks (i.e., for the actor to move their hand to other balls). Movement of grip position left-right showed a significant interaction. Whilst the effect of the video condition was in the predicted direction, the difference in the face-to-face condition for left-right movement may be due to a better level of prediction by the imitators for complex rather than simple conditions in this direction, though it is unclear why this would be the case.

The differences between face-to-face and video observation may partly be due to the ± 25 ms SD in the video projection latency. This temporal jitter surprised us, and was not controlled for in our experiment or analysis. This variable is also likely not controlled in previous research using pre-recorded video stimuli, such that researchers cannot be sure of a constant level of visual quality in their stimuli. Varying visual quality at any one time in a video could alter participant responses in a way that is not consistent with the variable being measured. We believe that researchers would benefit from providing this measure of standard deviation, or some other measure of temporal precision of video stimuli.

Some aspects of our experimental approach may have limited the reliability and validity of our results. Allowing actors to move in any way they chose, rather than in 10 consecutive movements, may have resulted in data more indicative of real life transitive motor activity. However, we felt it was important to maintain some element of control over the way in which participants moved for a number of reasons. By providing a relatively fixed way in which the actor was required to move, it ensured that their actions had a specific aim. As mentioned above, intention is potentially important in action observation (Becchio et al.,

2012), and allowing the actor to move completely freely may have resulted in changes in their aims across conditions. Secondly, we believed that having a set aim across the trials better reflected imitation in real life tasks that have a definite goal and action profile (for example, serving in a game of tennis). This paradigm also ensured that trials could be compared to each other across participants and conditions with reasonable accuracy.

Additionally, a confound in the order of the difficulty conditions may have affected the results with regards to practice effects, but if practice effects were strong, the effects should be in the opposite direction to those found. Using the same participants as both actor and imitator may also have affected the results, with participants playing the role of actor first potentially displaying greater skill at the imitation task. However, an even number of participants ensured that condition order was counterbalanced. Two out of the 12 participants tested were male, and differences in gender may have in some way influenced the results, since there is evidence for differences in simulation strategies between males and females (Kessler and Wang, 2012).

Lastly, our treatment of joint angles, though novel in the research of imitation kinematics, was not entirely optimal. First and foremost, q_1 – q_3 were not “true” joint angles in that they did not pass through the center of the joints. This was impossible to avoid with motion trackers on the surface of the skin, and has been commented on before by previous (non-imitative) research using joint angle kinematics (e.g., Murphy et al., 2006). We do not believe that this undermines the analysis, since the joint angle calculations can be seen as a best estimate, and are likely to closely resemble the true joint angle motion of the actor and the imitator. In addition to this, q_3 did not take into account the rotation of the wrist. However, since we used joint angles as a general measure of arm movement, and not as a way to define the position of the hand, this was also of little concern to our analysis.

Future research may choose to focus on neural differences between face-to-face and video feedback in transitive imitation. This is especially timely considering it is 14 years since Järveläinen et al. (2001) found measurable differences in motor cortex activity between observation of motor actions in face-to-face and video stimuli. Changes in the activity of the motor cortex are likely accompanied by changes in regions including the inferior frontal gyrus, inferior parietal lobule, and posterior superior temporal sulcus (Molenberghs et al., 2009). Translating our design to neuroimaging or neurostimulation may further develop our understanding of the neural effects of video feedback.

Another avenue for research could aim to discover where the difference between face-to-face and video feedback lies. Is it due to the lack of real two-person interaction, or rather due to visual differences between video and real life observation? The findings of Järveläinen et al. (2001) suggest that it could be the latter, but there is a growing consensus regarding the importance of two-person interactions in social psychological research (Schippers et al., 2010; Yun et al., 2012; Liu and Pelowski, 2014). In this experiment the difference could also be due to the reduced social context available to the actor. Perhaps a more reliable way of using pre-recorded video stimuli in the future would involve videoing an actor in an actual imitation task, rather than just performing actions of their own accord (though this could create new problems). As mentioned in the introduction, it is still unclear how an observer can constrain their own motor system in order to imitate an action (the correspondence problem). Our experiment suggests that this process may be influenced in some way by variables beyond simple motor observation, such as the visual quality of the observed movement or the extent to which it is likely to result in a real, two-person interaction. This is worth considering when testing different aspects of imitation. Social aspects of imitation may be more influenced by the lack of real face-to-face interaction, whilst motor aspects may be more influenced by the visual fidelity of video stimuli.

To conclude, it is evident that there are detrimental effects of video stimuli on the accuracy of imitation which may have been overlooked in previous research. This is evident in positional information regarding task-specific, object-directed movement. However, other aspects of transitive imitation (joint angles, grip aperture), may not be affected by the use of video stimuli. Future research should aim to develop new methods of examining imitation that are less reliant on video stimuli, and more closely adhere to the idea of imitation as a method of social communication. This would ensure the development of a more complete understanding of human imitation.

Acknowledgments

This research was supported by the Economic and Social Research Council (grant number ES/J500148/1 to ATR) and the Medical Research Council (grant number MR/K014250/1 to NPH). The authors would like to thank Dr. Yoshikatsu Hayashi and Nicolas Thorne Terry for their expertise and ideas regarding analysis, and Siobhán Ludden for her suggested reading.

References

- Becchio, C., Manera, V., Sartori, L., Cavallo, A., and Castiello, U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Front. Hum. Neurosci.* 6:117. doi: 10.3389/fnhum.2012.00117
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition* 106, 894–912. doi: 10.1016/j.cognition.2007.05.004
- Braadbaart, L., Walter, G. D., and Williams, J. H. G. (2012). Neural correlates of individual differences in manual imitation fidelity. *Front. Integr. Neurosci.* 6:91. doi: 10.3389/fnint.2012.00091
- Brass, M., and Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends Cogn. Sci.* 9, 489–495. doi: 10.1016/j.tics.2005.08.007
- Campione, G. C., and Gentilucci, M. (2011). Is the observation of the human kinematics sufficient to activate automatic imitation of transitive actions? *Behav. Brain Res.* 225, 201–208. doi: 10.1016/j.bbr.2011.07.025
- Castiello, U., and Ansuini, C. (2009). “Kinematic assessment of grasping,” in *Sensorimotor Control of Grasping: Physiology and Pathophysiology*, eds D. A. Nowak and J. Hermsdörfer (Cambridge: Cambridge University Press), 20–32.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi: 10.1093/brain/119.2.593

- Hardwick, R. M., and Edwards, M. G. (2012). Motor interference and facilitation arising from observed movement kinematics. *Q. J. Exp. Psychol.* 65, 840–847. doi: 10.1080/17470218.2012.672995
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15061–15065. doi: 10.1073/pnas.95.25.15061
- Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn. Sci.* 16, 114–121. doi: 10.1016/j.tics.2011.12.007
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazzotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526–2528. doi: 10.1126/science.286.5449.2526
- Järveläinen, J., Schürmann, M., Avikainen, S., and Hari, R. (2001). Stronger reactivity of the human primary motor cortex during observation of live rather than video motor acts. *Neuroreport* 12, 3493–3495. doi: 10.1097/00001756-200111160-00024
- Kessler, K., and Wang, H. F. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: differences predicted by sex and social skills score. *Spat. Cogn. Comput.* 12, 133–158. doi: 10.1080/13875868.2011.634533
- Knudson, D. (2007). *Fundamentals of Biomechanics*. New York, NY: Springer.
- Leighton, J., Bird, G., and Heyes, C. (2010). ‘Goals’ are not an integral part of imitation. *Cognition* 114, 423–435. doi: 10.1016/j.cognition.2009.11.001
- Liu, T., and Pelowski, M. (2014). Clarifying the interaction types in two-person neuroscience research. *Front. Hum. Neurosci.* 8:276. doi: 10.3389/fnhum.2014.00276
- Molenberghs, P., Cunnington, R., and Mattingley, J. B. (2009). Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neurosci. Biobehav. Rev.* 33, 975–980. doi: 10.1016/j.neubiorev.2009.03.010
- Murphy, M. A., Sunnerhagen, K. S., Johnels, B., and Willén, C. (2006). Three-dimensional kinematic motion analysis of a daily activity drinking from a glass: a pilot study. *J. Neuroeng. Rehabil.* 3:18. doi: 10.1186/1743-0003-3-18
- Naish, K. R., Houston-Price, C., Bremner, A. J., and Holmes, N. P. (2014). Effects of action observation on corticospinal excitability: muscle specificity, direction, and timing of the mirror response. *Neuropsychologia* 64, 331–348. doi: 10.1016/j.neuropsychologia.2014.09.034
- Naish, K. R., Reader, A. T., Houston-Price, C., Bremner, A. J., and Holmes, N. P. (2013). To eat or not to eat? Kinematics and muscle activity of reach-to-grasp movements are influenced by the action goal, but observers do not detect these differences. *Exp. Brain Res.* 225, 261–275. doi: 10.1007/s00221-012-3367-2
- Oztop, E., Kawato, M., and Arbib, M. A. (2013). Mirror neurons: functions, mechanisms and models. *Neurosci. Lett.* 540, 43–55. doi: 10.1016/j.neulet.2012.10.005
- Patterson, R. (2009). Human factors of stereo displays: an update. *J. Soc. Inf. Disp.* 17, 987–996. doi: 10.1889/JSID17.12.987
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180. doi: 10.1007/BF00230027
- Pierpaoli, C., Ferrante, L., Manzoni, T., and Fabri, M. (2014). Anatomical or mirror mode imitation? A behavioral approach. *Arch. Ital. Biol.* 152, 20–31.
- Ruysschaert, L., Warreyn, P., Wiersema, J. R., Metin, B., and Roeyers, H. (2013). Neural mirroring during the observation of live and video actions in infants. *Clin. Neurophysiol.* 124, 1765–1770. doi: 10.1016/j.clinph.2013.04.007
- Sartori, L., Becchio, C., and Castiello, U. (2011). Cues to intention: the role of movement information. *Cognition* 119, 242–252. doi: 10.1016/j.cognition.2011.01.014
- Schippers, M. B., Roebroeck, A., Renken, R., Nanetti, L., and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9388–9393. doi: 10.1073/pnas.1001791107
- Wild, K. S., Poliakoff, E., Jerrison, A., and Gowen, E. (2010). The influence of goals on movement kinematics during imitation. *Exp. Brain Res.* 204, 353–360. doi: 10.1007/s00221-009-2034-8
- Wohlschläger, A., Gattis, M., and Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 358, 501–515. doi: 10.1098/rstb.2002.1257
- Yun, K., Watanabe, K., and Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. *Sci. Rep.* 2:959. doi: 10.1038/srep00959

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Reader and Holmes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Social cues to joint actions: the role of shared goals

Lucia M. Sacheli^{1,2,3*}, Salvatore M. Aglioti^{1,2} and Matteo Candidi^{1,2*}

¹ Department of Psychology, University of Rome "Sapienza", Rome, Italy, ² Istituto di Ricovero e Cura a Carattere Scientifico, Fondazione Santa Lucia, Rome, Italy, ³ Department of Psychology, University of Milano-Bicocca, Milan, Italy

OPEN ACCESS

Edited by:

Maurizio Gentilucci,
University of Parma, Italy

Reviewed by:

Elena Daprati,
Università di Roma Tor Vergata, Italy
Ric Dalla Volta,
University Magna Graecia, Italy

*Correspondence:

Lucia M. Sacheli and Matteo Candidi,
Department of Psychology, University
of Rome "Sapienza", Via dei Marsi 78,
Rome I-00185, Italy
lucia.sacheli@uniroma1.it;
matteo.candidi@uniroma1.it

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 01 April 2015

Accepted: 06 July 2015

Published: 30 July 2015

Citation:

Sacheli LM, Aglioti SM and Candidi M
(2015) Social cues to joint actions:
the role of shared goals.
Front. Psychol. 6:1034
doi: 10.3389/fpsyg.2015.01034

In daily life, we do not just move independently from how others move. Rather, the way we move conveys information about our cognitive and affective attitudes toward our conspecifics. However, the implicit social substrate of our movements is not easy to capture and isolate given the complexity of human interactive behaviors. In this perspective article we discuss the crucial conditions for exploring the impact of "interpersonal" cognitive/emotional dimensions on the motor behavior of individuals interacting in realistic contexts. We argue that testing interactions requires one to build up naturalistic and yet controlled scenarios where participants reciprocally adapt their movements in order to achieve an overarching "shared goal." We suggest that a shared goal is what singles out real interactions from situations where two or more individuals contingently but independently act next to each other, and that "interpersonal" socio-emotional dimensions might fail to affect co-agents' behaviors if real interactions are not at place. We report the results of a novel joint-grasping task suitable for exploring how individual sub-goals (i.e., correctly grasping an object) relate to, and depend from, the representation of "shared goals."

Keywords: joint-action, shared goals, socio-emotional context, interpersonal perception, kinematics, grasping

Introduction

"The difference between a helping hand and an outstretched palm is a twist of the wrist"

L. Leamer, King of the Night.

In order to explore the neuro-cognitive bases of interpersonal encounters, social neuroscience needs to shift from "isolation paradigms" (Becchio et al., 2010), which investigate "offline" social cognition from the point of view of a (passive) observer (Pfeiffer et al., 2013), to an active, "second-person" approach (Schilbach et al., 2013), which validates the idea that—in real life—"online" social interaction is much more than just the concurrent recruitment of the essentially isolated social knowledge of individuals (see also Gallotti and Frith, 2013). This implies adopting experimental set-ups that (i) explore the emergence of closed-loop processes (i.e., allowing partners' reciprocal adjustments during the interaction), and (ii) take into account the emotional engagement that characterizes social encounters (Schilbach et al., 2013).

This issue becomes essential when studying "joint actions (JAs)," which we refer to here, defined as activities involving two or more individuals who need to coordinate their actions in time and space with the aim to realize together a desired change in the environment (Sebanz et al., 2006). This scenario requires dynamic experimental paradigms where the agent's individual goal is inherently linked to that of a partner thus depending on mutual adjustments, and where participants perceive themselves as a "couple" which is acting together as a unity *because* they

share an overarching common goal. In the present perspective article we suggest that only such experimental paradigms will allow scholars to study the “socio-emotional nature” of interpersonal coordination and will create a context suitable for exploring whether, and how, socio-emotional variables impact the quality of the interaction. We will focus on on-line interactions that—in our view—best highlight why mutual adjustments are based on shared goals. Indeed, while in turn-taking situations each individual is “passive” at some point during the interaction, on-line interactions require synchronicity in space and time: thus, they require that co-agents actively understand the partner’s behavior and predict his/her action goal while also monitoring their own action execution. This situation requires adapting one’s own action goal to a shared representation of the interaction. This is not a mere difference in complexity, but it is a difference in quality: without the constraint of synchronicity, the interaction may reproduce a condition where one individual is a passive observer.

In what follows, we try and provide an experimentally useful definition of “shared goals,” and we describe why we believe shared goals single out JAs from situations where an agent passively observes or mechanically react to the actions of other individuals. Then, we will operationalize how shared goals can be investigated in a well-controlled interactive task and explain why the analysis of kinematics in general (see, for instance Noy et al., 2011; D’Ausilio et al., 2012; Vesper et al., 2013; Vesper and Richardson, 2014) and grasping kinematics in particular (Georgiou et al., 2007; Becchio et al., 2008a,b; Sartori et al., 2009) might be a powerful instrument to explore the neuro-cognitive instantiations of shared goals. Thus, we will describe a set-up that we specifically designed to investigate the relation between individual and joint goals during an interactive grasping task (Sacheli et al., 2012, 2013, 2015). Our studies provide empirical evidence that motor tasks that include shared goals are suitable for exploring the impact of the socio-emotional context on planned interpersonal coordination.

Defining Shared Goals

Although there is evidence of “proto” forms of cooperative activities in non-human species (Mendres and de Waal, 2000; Seed et al., 2008; Plotnik et al., 2011), studies suggest that the tendency to interact and pursue common goals is typically human and shows up early in development (Tomasello et al., 2005; Warneken et al., 2006). Importantly, the tendency to share goals and intentions with others might support the establishment of social bonds: the efficacy of the interaction itself and the emotional reactions to it may also influence the process of coding others as in-group or out-group members (Hommel et al., 2009; Iani et al., 2011).

In the present perspective article, we focus on (on-line) JAs as a way to realize shared goals. Influential studies suggest that performing successful joint-actions depends on the ability to: (i) share representations, (ii) predict others’ actions, and (iii) integrate predicted effects of one’s own and others actions (Sebanz et al., 2006). Crucially, this definition highlights that interacting individuals cannot directly access a partner’s motor plan and thus need to infer it from his/her overt behavior (aside

from environmental cues). Moreover, since reactive processes do not suffice in supporting the fine-tuned temporal contingency required by on-line interpersonal coordination, co-agents cannot simply react to the partner’s behavior but need to predict it (Knoblich and Jordan, 2003). Predictive coding is (at least partially) based on predictive sensorimotor processes triggered by the observation of others’ actions (Kilner et al., 2007). However, a fundamental question concerns *what* is actually “shared” of motor representations during on-line interpersonal interactions. We suggest that “shared goals” (Butterfill, 2012) create a link between interacting co-agents by integrating, in a unique motor plan, the representation of one’s own and a partner’s action (see also Knoblich et al., 2011).

According to Butterfill (2012), three features of shared goals are crucial: (i) there is a single shared goal, *G*, to which each agent’s actions are (or will be) individually directed; (ii) each agent expects each of the other agents to perform an action directed to the shared goal *G*; (iii) each agent expects this goal *G* to occur as a common effect of all actions directed toward it, i.e., both his or her own and the partner’s ones. Thus, a shared goal is both “in common” between co-agents and “divided up” into individual sub-goals that each actor needs to achieve to fully accomplish the intended JA.

In keeping with computational (Wolpert and Ghahramani, 2000) and neurophysiologic studies (Fogassi et al., 2005; Grafton and Hamilton, 2007) indicating that the motor system represents individual goals according to a hierarchical structure, we suggest that JA and shared goal representations are not independent from this organization: just as individual muscular synergies are coordinated in complex actions by the need to achieve a desired (individual) motor goal, interpersonal motor synergies are shaped by the presence of shared goals which organize co-agents’ behaviors (Chersi, 2011; Candidi et al., 2015). In keeping, framing research on JA as *research on actions involving two or more agents sharing a common goal* implies suggesting JAs are characterized by a “hierarchical structure” where the accomplishment of a (shared) overarching goal depends on the fulfillment of the sub-goals that each interacting partner is required to achieve. For instance, the overarching common goal of moving a table together is achieved only when both partners accomplish their own individual sub-goal (e.g., pulling and pushing the table in the right direction) by dynamically adapting to each other in space and time.

Importantly, the definition of “shared goals” provided above does not overlap with the one of “shared representations” as they have been defined by studies on joint attention (Sebanz et al., 2003, 2005, 2007). Studies on joint attention typically investigate conditions where one binary choice task with two competitive target stimuli is split between two participants, with each participant responding to only one of the targets in turn-taking because he/she has “his/her own target” to respond to (e.g., paradigms leading to joint Simon effect, Tsai et al., 2008; Flanker effect, Atmaca et al., 2011, and SNARC effect, Atmaca et al., 2008). In these tasks, participants have to attend to one target and to ignore the other. Thus, in principle their performance in the joint condition should resemble the one in individual go/no-go tasks. However, participants involuntarily take the co-actor’s task into account albeit they are explicitly instructed to

ignore it. This suggests that humans have a tendency to form “task co-representations” which specify not only one’s own but also a co-actor’s task, even if the co-actor’s task is irrelevant to (or even interfering with) one’s own task fulfillment. Although the ability to co-represent a task is obviously crucial in JA, the above studies resemble interference effects reported in studies on action-perception coupling (Brass et al., 2000, 2001; Kilner et al., 2003). Namely, they may tap incidental and automatic (i.e., “passive”) processes recruited when agents act independently but contingently. Accordingly, it has been suggested that what participants co-represent in these “task sharing” scenarios is *that* another agent is present and *when* it is his/her turn, but not *what* the other agent needs to do (Wenke et al., 2011). On the contrary, shared goals imply that partners have clear in mind both what they need to do (i.e., their own sub-goal), what the partner needs to do (i.e., his/her sub-goal) and their common effects. Hence, shared goals are “active” ingredients of our motor planning: they enable co-agents to dynamically integrate predicted effects of the partner’s action within the agent’s motor plan. In the following section we will explain why a joint-grasping set-up provides an excellent opportunity to investigate the hierarchical structure (where co-agents’ sub-goals depend on overarching shared goals) that—in our view—characterize motor planning during JAs.

Grasping Kinematics: From Individual Transitive Behavior to Interpersonal Goal Sharing

Prehension, i.e., the capacity to reach and grasp, is the key behavior that allows humans to change their environment, and it has been largely described both in terms of its kinematic features (Jeannerod, 1981, 1984) and in terms of its neural bases (see Castiello, 2005, for a review), thus becoming an “experimental test-case” for the study of transitive, goal directed actions (Grafton, 2010). Indeed, prehension is a somewhat stereotyped movement in which maximum grip aperture (i.e., the thumb-finger maximum distance) is a landmark always occurring at 60–70% of the reach trajectory and highly correlated with object size (Jeannerod, 1981, 1984; see also Bootsma et al., 1994; Smeets and Brenner, 1999). Thus, grasping kinematics follows stereotypical patterns if other factors do not intervene. Importantly, however, grasping kinematics also depends on its desired end-goal. In fact, not only objects features (e.g., texture, weight and fragility; Johansson and Westling, 1988; Weir et al., 1991; Savelsbergh et al., 1996) but also the intentions of an agent (e.g., grasping an object to lift it, to place it in a precise location or to use it, Ansuini et al., 2006, 2008) modify grasping pre-shaping, i.e., the relative position of fingers during the reaching phase, and the contact points of the fingers on the object (Sartori et al., 2011). Finally and most importantly, prehension kinematics is also modulated by social factors as the co-agent’s communicative (Sartori et al., 2009; Ferri et al., 2011; Innocenti et al., 2012) or cooperative/competitive intention (Georgiou et al., 2007; Becchio et al., 2008a,b; see Becchio et al., 2012, for a review). Thus, grasping kinematics becomes an ideal candidate to explore how the socio-emotional context modulates agents’ overt behavior

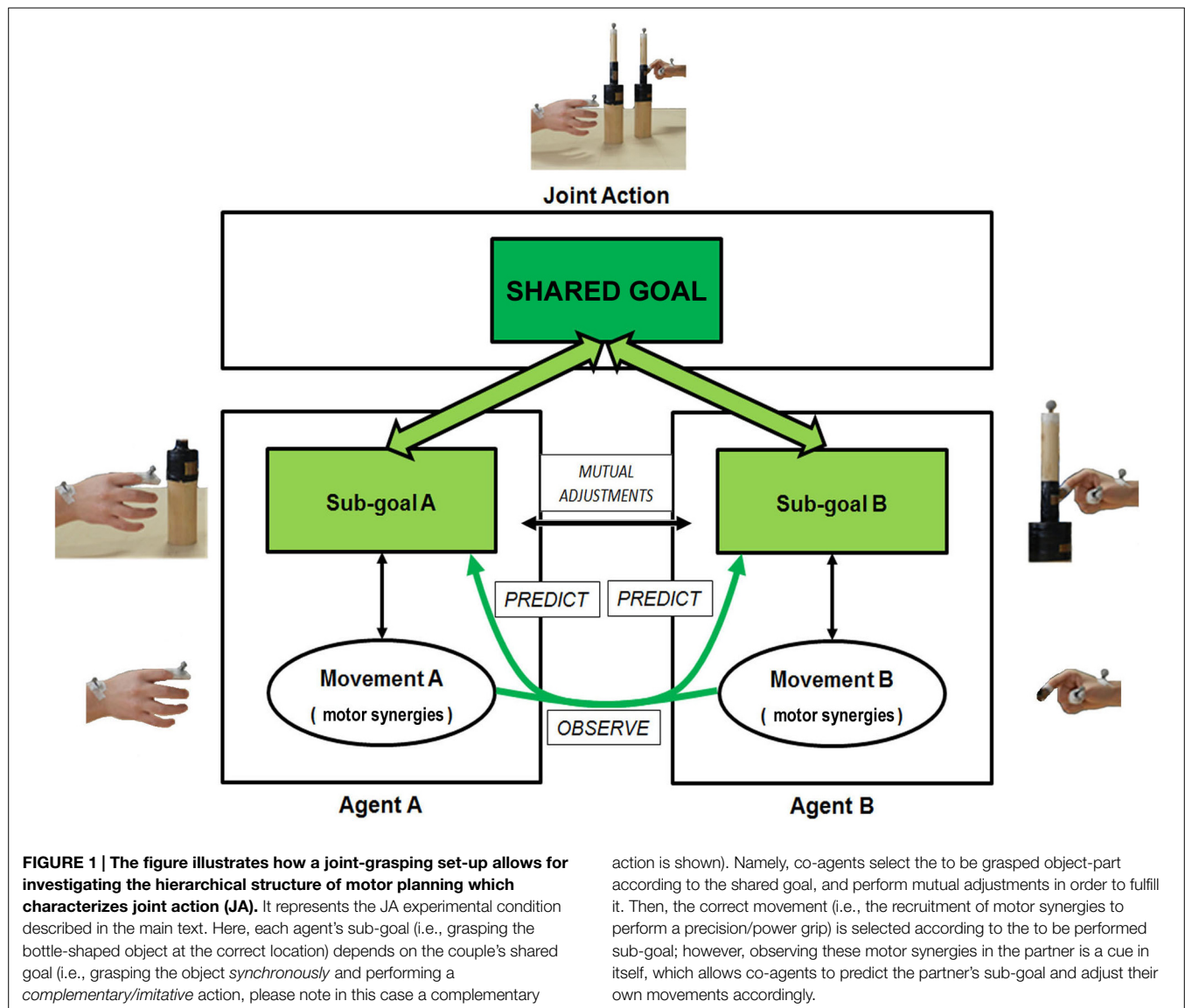
during realistic face-to-face interactions, by using a set-up where object properties (i.e., the physical “sub-goal” of each agent) is kept constant (and cannot thus modulate kinematics) but the co-agents’ “shared” goal and the socio-emotional context are modulated instead. Suggestions have been made that the observation of movement kinematics is what allows an observer to infer the agent’s intention by simply noting details of his/her overt behavior (Ansuini et al., 2014; see also D’Ausilio et al., 2015). For instance, we can distinguish when a given action (say, making a pass) is used for its pragmatic goal (e.g., passing the ball to the teammate) or for a communicative one (e.g., signaling to a co-actor the direction of the pass) from minimal motor cues.

The Implementation of Shared Goals in a Joint-Grasping Task

Taking advantage from early attempts to apply grasping kinematic analysis to research on JA, in recent years we developed a joint-grasping task where each of two individuals sitting one in front of the other is required to reach and grasp a bottle-shaped object. The objects provided to each individual are identical and designed to prompt a precision grip (when grasping a small cylinder in the higher part of the bottle) or a power grip (when grasping a large cylinder in the lower part of the bottle; see **Figure 1**).

Participants are required to reach-and-grasp the bottle in the correct part following different instructions. Crucially, however, each participant needs to perform the task as synchronously as possible with the partner. The more participants are synchronous, the higher their common payoff. As synchronicity with the partner is essential to fulfill the instructions—in this scenario as well as in many daily life situations—Grasping Asynchrony is the critical dependent variable indexing the success of interpersonal coordination (Sacheli et al., 2012, 2013).

Four features of this paradigm are crucial. Two participants are instructed to perform a face-to-face motor task (i) implying a *shared goal* (i.e., be synchronous) which is dependent on participants’ ability to achieve their own motor sub-goals (i.e., grasping the bottle-shaped object at the correct location), and which also implies that (ii) each participant’s motor sub-goal is dependent on the partner’s action (i.e., the task requires *mutual adjustments*); moreover, in different experimental conditions, participants have (iii) to perform either *imitative* or *complementary* movements with respect to their partner’s one, and (iv) to adjust to the partner’s movements either in *time* only [“synchronization” (Synchr) condition, requiring to be synchronous only] or in time and space [“joint action” (JA) condition, requiring to be synchronous and to adapt to the partner’s sub-goal]. Importantly, in the JA condition participants do not know where to grasp the object in advance: both partners only receive an auditory cue that specifies whether they have to perform an imitative action (precision–precision or power–power grip) or a complementary action (precision–power grip or *vice versa*) as a couple. As a consequence, they have to reciprocally adapt their movements on-line in order to select which action (e.g., precision–precision or power–power grip in case of imitative actions) they are going to perform based on the movement of their partner. Thus, although in principle both Synchr and JA



imply a “shared goal” (i.e., be synchronous) according to the Butterfill’s (2012) definition, only the JA condition would capture a situation where participants need to predict the partner’s action and sub-goal in order to select their own action and sub-goal to achieve the shared goal (i.e., not only be synchronous with but also complementary/imitative to your partner): namely, JA requires participants to predict (and represent) *what* the partner is doing (see **Figure 1**). On the contrary, the partner’s sub-goal might be totally disregarded in the Synchr condition, at least in its spatial features. In this regard, the Synchr condition implies “task sharing” and not necessarily “shared goals” (see the distinction outlined above).

Thanks to such task structure (which includes shared goals) and the peculiar feature of grasping movements, we have been able to explore how co-agents’ (individual) behavior is modulated by socio-emotional variables (Sacheli et al., 2012, 2013, 2015).

In one study, by applying this set-up we showed that a negative interpersonal perception (induced by the feeling that the partner

has mined one’s own self-esteem, Caprara et al., 1987) strongly modulates the ability to reciprocally coordinate in JA (Sacheli et al., 2012, see **Figure 2**). Specifically, when participants interact within a negative interpersonal scenario (i.e., negatively biased group), their performance in the JA condition is significantly lower than in the Synchr one (**Figure 2A**), suggesting they act “each one on their own”: they do not represent the shared goal and hence disregard the partner’s sub-goal, and this impairs the performance when mutual adjustments are required (i.e., in JA). As a matter of fact, the analyses on participants’ movement kinematics demonstrate that participants’ maximum grip aperture in JA is less variable, indicating they perform less movement corrections. This evidence supports the idea that participants are less prone to represent and adapt to the partner’s action and sub-goal (**Figure 2B**). On the contrary, we showed that in a neutral interpersonal situation pairs of participants achieve the same level of performance in Synchr and JA (**Figure 2A**): this suggests they represent the task as what Vesper et al. (2010) define a “me + X

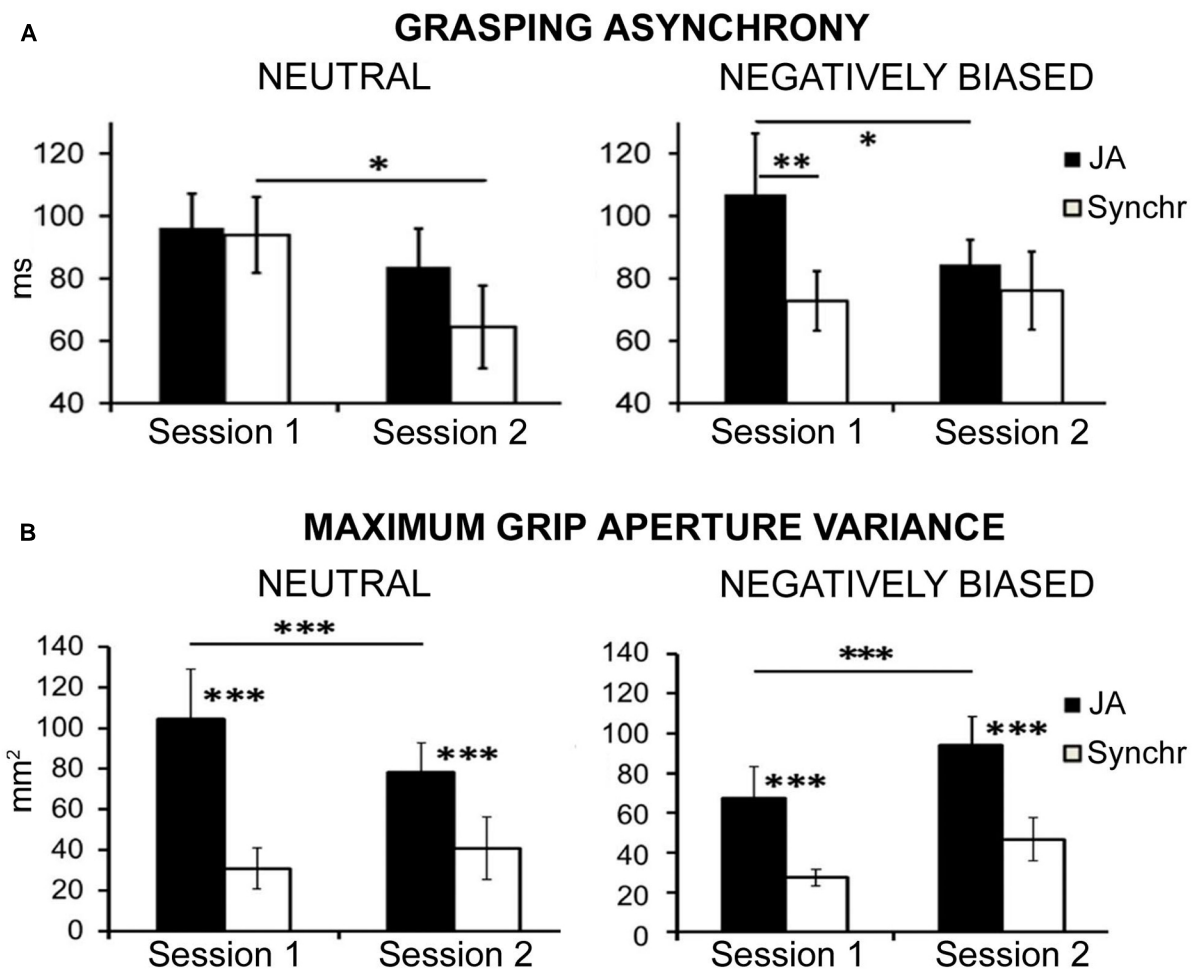


FIGURE 2 | Modified from Sacheli et al. (2012). The graphs illustrate results comparing behavioral performance in terms of grasping asynchrony (i.e., time-delay between partners' grasp times on the bottle, **A**) and grasping kinematics (in terms of maximum grip aperture variance, **B**) of two groups of participants. In one group the interpersonal perception is left neutral (Neutral group), while in the other it is negatively biased (Negatively biased group). **(A)** In the Neutral group, participants initially (session 1) achieve the same level of performance in the synchronization (Synchr) and in the joint action (JA) condition. On the contrary, in the Negatively biased group, participants' performance in JA is significantly lower (i.e., grasping asynchrony is higher, indicating poorer performance) than in Synchr: this suggests they are acting

"each one on his own," and are thus not able to reciprocally adapt in order to achieve the JA shared goal. **(B)** Coherently, analysis of kinematics reveals that maximum grip aperture is much less variable (indicating less movement corrections and thus less reciprocal adjustments) in the Negatively biased group, supporting the hypothesis that they are less prone to adapt to the partner's action. However, in the second session maximum grip aperture variance in negatively biased participants increases in JA. This effect is paralleled by an improvement in interpersonal coordination as measured by grasping asynchrony. See the main text for an interpretation of these results. JA, joint action condition; Synchr, synchronization condition. Error bars indicate SEM. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

mode," i.e., including the partner's movement in their own motor plan, independently from the experimental condition, namely even when they are not necessarily required to do so (i.e., in the Synchr condition).

Hence, a negative interpersonal bond reduces the tendency to map others' behavior onto one's own sensorimotor system for the purpose of representing the shared goal of the joint-grasping task: participant act independently from each other and do not reciprocally adapt, as if they did not automatically resolve back to a sensorimotor representation of the partner's movement for the sake of shared goal fulfillment. Conversely, this "shared goal representation" is established in neutral interpersonal situations.

Importantly, in a second session of the experiment, negatively biased participants improve their performance in the JA condition, and this is paralleled by an increase in movement corrections as shown by grasping kinematics. This suggests that acting together might itself facilitate the creation of a social bond between interacting co-agents and change the way partners represent the task: from representing it as an "individual" grasping task where two agents act in synchrony but independently ("task sharing" mode) to representing the task as a joint grasping task having an overarching, cooperative shared goal ("shared goal" mode). Accordingly, one might hypothesize that JA tasks like the one described here might be exploited to investigate whether acting together reduces biases toward

other individuals. For instance, we showed that the JA condition is also modulated by racial prejudices (Sacheli et al., 2015), as sensorimotor simulation recruited during JA (indexed by visuo-motor interference measured by the comparison between complementary and imitative actions) negatively correlates with the individual ethnic bias (i.e., it is reduced when interacting with the out-group partner in biased participants only). Studies indicate that unconscious mimicry of others' postures and mannerisms during interactions may have the social scope of promoting affiliation (Lakin and Chartrand, 2003; van Baaren et al., 2004, 2009), and that the voluntary mimicry of out-group members reduces racial stereotypes (Inzlicht et al., 2012). In a similar vein, the reinforcement of social bonds that arises during prolonged motor interactions (Sacheli et al., 2012) may exert the same powerful modulation.

Conclusion

The present perspective focuses on the idea that the presence of a shared goal is what qualifies an on-line interaction as being

penetrable to interpersonal cues. *Vice versa*, the presence of shared goals during mutually adaptive interactions may promote affiliation between interacting individuals by reinforcing their emotional bond, and this may be reflected in subtle changes in co-agents' interactional behaviors that can be captured analysing their movement kinematics. The present perspective article intends to suggest that future research on the socio-emotional components of motor interactions do not necessarily require "complex" interactional set-up: indeed, even extremely instrumental and overlearned movements (such as grasping movements) can be shaped by the emotional context in which interaction takes place, provided that the interactive task implies a shared goal.

Acknowledgments

Funded by the EU Information and Communication Technologies Grant (VERE project, FP7-ICT-2009-5, Prot. Num. 257695) and the Italian Ministry of Health (grant RC11.G and RF-2010-2312912).

References

- Ansuini, C., Cavallo, A., Bertone, C., and Becchio, C. (2014). The visible face of intention: why kinematics matters. *Front. Psychol.* 5:815. doi: 10.3389/fpsyg.2014.00815
- Ansuini, C., Giosa, L., Turella, L., Altoè, G., and Castiello, U. (2008). An object for an action, the same object for other actions: effects on hand shaping. *Exp. Brain Res.* 185, 111–119. doi: 10.1007/s00221-007-1136-4
- Ansuini, C., Santello, M., Massaccesi, S., and Castiello, U. (2006). Effects of end-goal on hand shaping. *J. Neurophysiol.* 95, 2456–2465. doi: 10.1152/jn.01107.2005
- Atmaca, S., Sebanz, N., and Knoblich, G. (2011). The joint flanker effect: sharing tasks with real and imagined co-actors. *Exp. Brain Res.* 211, 371–385. doi: 10.1007/s00221-011-2709-9
- Atmaca, S., Sebanz, N., Prinz, W., and Knoblich, G. (2008). Action co-representation: the joint SNARC effect. *Soc. Neurosci.* 3, 410–420. doi: 10.1080/17470910801900908
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., and Castiello, U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Front. Hum. Neurosci.* 6:117. doi: 10.3389/fnhum.2012.00117
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008a). The case of Dr. Jekyll and Mr. Hyde: a kinematic study on social intention. *Conscious. Cogn.* 17, 557–564. doi: 10.1016/j.concog.2007.03.003
- Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008b). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition* 106, 894–912. doi: 10.1016/j.cognition.2007.05.004
- Becchio, C., Sartori, L., and Castiello, U. (2010). Toward you: the social side of actions. *Curr. Dir. Psychol. Sci.* 19, 183–188. doi: 10.1177/0963721410370131
- Bootsma, R. J., Marteniuk, R. G., MacKenzie, C. L., and Zaal, F. T. J. M. (1994). The speed-accuracy trade off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Exp. Brain Res.* 98, 535–541. doi: 10.1007/BF00233990
- Brass, M., Bekkering, H., and Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychol. (Amst.)* 106, 3–22. doi: 10.1016/S0001-6918(00)00024-X
- Brass, M., Bekkering, H., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain Cogn.* 44, 124–143. doi: 10.1006/brcg.2000.1225
- Butterfill, S. (2012). Joint action and development. *Philos. Q.* 62, 1467–9213. doi: 10.1111/j.1467-9213.2011.00005.x
- Candidi, M., Sacheli, L. M., and Aglioti, S. M. (2015). From muscles synergies and individual goals to interpersonal synergies and shared goals: mirror neurons and interpersonal action hierarchies: comment on "Grasping synergies: a motor-control approach to the mirror neuron mechanism" by D'Ausilio et al. *Phys. Life Rev.* 12, 126–128. doi: 10.1016/j.plrev.2015.01.023
- Caprara, G. V., Gargaro, T., Pastorelli, C., Prezza, M., Renzi, P., and Zelli, A. (1987). Individual differences and measures of aggression in laboratory studies. *Pers. Indiv. Dif.* 8, 885–893. doi: 10.1016/0191-8869(87)90140-1
- Castiello, U. (2005). The neuroscience of grasping. *Nat. Rev. Neurosci.* 6, 726–736. doi: 10.1038/nrn1744
- Chersi, F. (2011). Neural mechanisms and models underlying joint action. *Exp. Brain Res.* 211, 643–653. doi: 10.1007/s00221-011-2690-3
- D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., et al. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS ONE* 7:e35757. doi: 10.1371/journal.pone.0035757
- D'Ausilio, A., Bartoli, E., and Maffongelli, L. (2015). Grasping synergies: a motor-control approach to the mirror neuron mechanism. *Phys. Life Rev.* 12, 91–103. doi: 10.1016/j.plrev.2014.11.002
- Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., and Gentilucci, M. (2011). Social requests and social affordances: how they affect the kinematics of motor sequences during interactions between conspecifics. *PLoS ONE* 6:e15855. doi: 10.1371/journal.pone.0015855
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. doi: 10.1126/science.1106138
- Gallotti, M., and Frith, C. D. (2013). Social cognition in the we-mode. *Trends Cogn. Sci.* 17, 160–165. doi: 10.1016/j.tics.2013.02.002
- Georgiou, I., Becchio, C., Glover, S., and Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition* 102, 415–433. doi: 10.1016/j.cognition.2006.01.008
- Grafton, S. T. (2010). The cognitive neuroscience of prehension: recent developments. *Exp. Brain Res.* 204, 475–491. doi: 10.1007/s00221-010-2315-2
- Grafton, S. T., and Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Hum. Mov. Sci.* 26, 590–616. doi: 10.1016/j.humov.2007.05.009
- Hommel, B., Colzato, L. S., and van den Wildenberg, W. P. M. (2009). How social are task representations? *Psychol. Sci.* 20, 794–798. doi: 10.1111/j.1467-9280.2009.02367.x
- Iani, C., Anelli, F., Nicoletti, R., Arcuri, L., and Rubichi, S. (2011). The role of group membership on the modulation of joint action. *Exp. Brain Res.* 211, 439–445. doi: 10.1007/s00221-011-2651-x
- Innocenti, A., De Stefani, E., Bernardi, N. F., Campione, G. C., and Gentilucci, M. (2012). Gaze direction and request gesture in social interactions. *PLoS ONE* 7:e36390. doi: 10.1371/journal.pone.0036390
- Inzlicht, M., Gutsell, J. N., and Legault, L. (2012). Mimicry reduces racial prejudice. *J. Exp. Soc. Psychol.* 48, 361–365. doi: 10.1016/j.jesp.2011.06.007

- Jeannerod, M. (1981). "Intersegmental coordination during reaching at natural visual objects," in *Attention and Performance IX*, eds J. Long, and A. Baddeley (Hillsdale: Erlbaum), 153–168.
- Jeannerod, M. (1984). The timing of natural prehension movements. *J. Mot. Behav.* 16, 235–254. doi: 10.1080/00222895.1984.10735319
- Johansson, R. S., and Westling, G. (1988). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Exp. Brain Res.* 71, 59–71. doi: 10.1007/BF00247522
- Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). The mirror-neuron system: a Bayesian perspective. *Neuroreport* 18, 619–623. doi: 10.1097/WNR.0b013e3281139ed0
- Kilner, J. M., Paulignan, Y., and Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Curr. Biol.* 13, 522–525. doi: 10.1016/S0960-9822(03)00165-9
- Knoblich, G., Butterfill, S., and Sebanz, N. (2011). "Psychological research on joint action: theory and data," in *The Psychology of Learning and Motivation*, ed. B. Ross (Burlington: Academic Press), 59–101.
- Knoblich, G., and Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1006–1016. doi: 10.1037/0278-7393.29.5.1006
- Lakin, J., and Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychol. Sci.* 14, 334–339. doi: 10.1111/1467-9280.14481
- Mendres, K. A., and de Waal, F. B. (2000). Capuchins do cooperate: the advantage of an intuitive task. *Anim. Behav.* 60, 523–529. doi: 10.1006/anbe.2000.1512
- Noy, L., Dekel, E., and Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20947–20952. doi: 10.1073/pnas.1108155108
- Pfeiffer, U. J., Timmermans, B., Vogeley, K., Frith, C. D., and Schilbach, L. (2013). Towards a neuroscience of social interaction. *Front. Hum. Neurosci.* 7:22. doi: 10.3389/fnhum.2013.00022
- Plotnik, J. M., Lair, R., Suphachoksahakun, W., and de Waal, F. B. (2011). Elephants know when they need a helping trunk in a cooperative task. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5116–5121. doi: 10.1073/pnas.1101765108
- Sacheli, L. M., Candidi, M., Pavone, E. F., Tidoni, E., and Aglioti, S. M. (2012). And yet they act together: interpersonal perception modulates predictive simulation and mutual adjustments during a joint-grasping task. *PLoS ONE* 7:e50223. doi: 10.1371/journal.pone.0050223
- Sacheli, L. M., Christensen, A., Giese, M. A., Taubert, N., Pavone, E. F., Aglioti, S. M., et al. (2015). Prejudiced interactions: implicit racial bias reduces predictive simulation during joint action with an out-group avatar. *Sci. Rep.* 5, 8507. doi: 10.1038/srep08507
- Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., and Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Exp. Brain Res.* 226, 473–486. doi: 10.1007/s00221-013-3459-7
- Sartori, L., Becchio, C., Bara, B. G., and Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Conscious. Cogn.* 18, 766–772. doi: 10.1016/j.concog.2009.06.004
- Sartori, L., Straulino, E., and Castiello, U. (2011). How objects are grasped: the interplay between affordances and end-goals. *PLoS ONE* 6:e25203. doi: 10.1371/journal.pone.0025203
- Savelsbergh, G. J. P., Steenbergen, B., and van der Kamp, J. (1996). The role of fragility information in the guidance of the precision grip. *Hum. Mov. Sci.* 15, 115–127. doi: 10.1016/0167-9457(95)00039-9
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/S0140525X12000660
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76. doi: 10.1016/j.tics.2005.12.009
- Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others actions: just like ones own? *Cognition* 88, B11–B21. doi: 10.1016/S0010-0277(03)00043-X
- Sebanz, N., Knoblich, G., and Prinz, W. (2005). How two share a task: corepresenting stimulus-response mappings. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 1234–1246. doi: 10.1037/0096-1523.31.6.1234
- Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., and Frith, C. D. (2007). Is it really my turn? An event-related fMRI study of task sharing. *Soc. Neurosci.* 2, 81–95. doi: 10.1080/17470910701237989
- Seed, A. M., Clayton, N. S., and Emery, N. J. (2008). Cooperative problem solving in rooks (*Corvus frugilegus*). *Proc. Biol. Sci.* 275, 1421–1429. doi: 10.1098/rspb.2008.0111
- Smeets, J. B. J., and Brenner, E. (1999). A new view on grasping. *Motor Control* 3, 237–271.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/S0140525X05000129
- Tsai, C. C., Kuo, W. J., Hung, D. L., and Tzeng, O. J. (2008). Action co-representation is tuned to other humans. *J. Cogn. Neurosci.* 20, 2015–2024. doi: 10.1162/jocn.2008.20144
- van Baaren, R., Janssen, L., Chartrand, T., and Dijksterhuis, A. (2009). Where is the love? The social consequences and moderators of mimicry in humans. *Proc. R. Soc. B* 364, 2381–2389. doi: 10.1098/rstb.2009.0057
- van Baaren, R. B., Holland, R. W., Kawakami, K., and van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychol. Sci.* 15, 71–74. doi: 10.1111/j.0963-7214.2004.01501012.x
- Vesper, C., Butterfill, S., Knoblich, G., and Sebanz, N. (2010). A minimal architecture for joint action. *Neural Netw.* 23, 998–1003. doi: 10.1016/j.neunet.2010.06.002
- Vesper, C., and Richardson, M. J. (2014). Strategic communication and behavioral coupling in asymmetric joint action. *Exp. Brain Res.* 232, 2945–2956. doi: 10.1007/s00221-014-3982-1
- Vesper, C., van der Wel, R. P., Knoblich, G., and Sebanz, N. (2013). Are you ready to jump? Predictive mechanisms in interpersonal coordination. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 48–61. doi: 10.1037/a0028066
- Warneken, F., Chen, F., and Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Dev.* 77, 640–663. doi: 10.1111/j.1467-8624.2006.00895.x
- Weir, P. L., MacKenzie, C. L., Marteniuk, R. G., and Cargoe, S. L. (1991). Is object texture a constraint on human prehension? kinematic evidence. *J. Mot. Behav.* 23, 205–210. doi: 10.1080/00222895.1991.10118363
- Wenke, D., Atmaca, S., Holländer, A., Liepelt, R., Baess, P., and Prinz, W. (2011). What is shared in joint action? Issues of co-representation, response conflict, and agent identification. *Rev. Phil. Psychol.* 2, 147–172. doi: 10.1007/s13164-011-0057-0
- Wolpert, D. M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3(Suppl.), 1212–1217. doi: 10.1038/81497

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Sacheli, Aglioti and Candidi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Complementary actions

Luisa Sartori^{1,2*} and Sonia Betti¹

¹ Dipartimento di Psicologia Generale, Università di Padova, Padova, Italy, ² Cognitive Neuroscience Center, Università di Padova, Padova, Italy

Complementary colors are color pairs which, when combined in the right proportions, produce white or black. Complementary actions refer here to forms of social interaction wherein individuals adapt their joint actions according to a common aim. Notably, complementary actions are incongruent actions. But being incongruent is not sufficient to be complementary (i.e., to complete the action of another person). Successful complementary interactions are founded on the abilities: (i) to simulate another person's movements, (ii) to predict another person's future action/s, (iii) to produce an appropriate incongruent response which differ, while interacting, with observed ones, and (iv) to complete the social interaction by integrating the predicted effects of one's own action with those of another person. This definition clearly alludes to the functional importance of complementary actions in the perception–action cycle and prompts us to scrutinize what is taking place behind the scenes. Preliminary data on this topic have been provided by recent cutting-edge studies utilizing different research methods. This mini-review aims to provide an up-to-date overview of the processes and the specific activations underlying complementary actions.

Keywords: action observation, perception–action coupling, social interactions, motor resonance, transcranial magnetic stimulation

Introduction

Motor resonance is defined as the subliminal activation of the motor system—and of the *imitative* response—while observing actions performed by others (reviewed in Heyes, 2011). Gallese (2001) explained that: “when we observe actions performed by other individuals our motor system ‘resonates’ along with that of the observed agent” (pp. 38–39). Numerous neurophysiological studies have in fact demonstrated that a motor resonant mechanism is at work in the motor, premotor, and the posterior parietal cortices when individuals are instructed to observe goal-directed actions being executed by another or others (for review, see Fadiga et al., 2005; Heyes, 2011; Rizzolatti et al., 2014). The discovery of mirror neurons in monkeys provided the physiological model for this perception–action coupling mechanism (Rizzolatti and Craighero, 2004). Located in the ventral premotor cortex (area F5) and the posterior parietal cortex, mirror neurons were found to fire *both* when a monkey carried out a goal-directed action as well as when it observed that same action being performed by another subject (di Pellegrino et al., 1992). Motor resonance appears then to pre-activate the motor system of an observer in order to represent and interpret the movements of another person even before the “go” signal has been given and activation remains for the most part on an unconscious level (Costantini et al., 2011).

While actions that are observed and those that are being planned appear functionally equivalent (Knoblich and Flach, 2001), it is unclear if the visual representation of an observed action inevitably leads to its motor representation. This is particularly true with regard to *complementary* (from Latin *complementum*; i.e., that fills up) actions, a specific class of movements which differ from –although

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Lucia M. Sacheli,
Sapienza University of Rome, Italy
Lincoln J. Colling,
Australian Catholic University,
Australia

*Correspondence:

Luisa Sartori,
Dipartimento di Psicologia Generale,
Università di Padova, Via Venezia 8,
35131 Padova, Italy
luisa.sartori@unipd.it

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 27 January 2015

Accepted: 17 April 2015

Published: 01 May 2015

Citation:

Sartori L and Betti S (2015)
Complementary actions.
Front. Psychol. 6:557.
doi: 10.3389/fpsyg.2015.00557

interacting with- an observed action (Sebanz et al., 2006; Knoblich et al., 2011). In the case, for example, that someone hands us a mug by its handle, we will automatically, without giving it a second thought, grab the mug using a whole-hand-grasp (the most appropriate grasping posture in this particular situation). The types of grasps adopted by the two interacting agents are incongruent, but they are nevertheless appropriate and complementary.

As a working definition, complementary actions refer here to any form of social interaction wherein two (or more) individuals coordinate and mutually complete their incongruent actions, rather than performing imitative behaviors. In this respect, we can define as *complementary affordances* all the action possibilities in which suitable motor programs aiming to bring a joint goal to completion are activated (such as grasping and offering a coin when seeing an open hand in sign of request). Depending on its posture and context, therefore, an extended open hand could lead to a donation, to a handshake or to an infinite number of other actions (Sartori et al., 2009). Activation of a complementary affordance is an important social tool, and it suggests that the automatic, rapid decoding of social cues influences intentional behavior in our everyday interactions, maximizing the efficiency of our responses. These examples illustrate the functional importance of complementary actions in the action-perception domain (Graf et al., 2009), and they prompt us to examine the mechanisms involved in producing those responses.

Behavioral Studies of Complementary Actions

Since the direct matching between observed and performed actions is thought to occur automatically, when we observe an action which differs from our intended action we have to inhibit the tendency to imitate (Brass et al., 2005). While the mechanism leading to automatic imitation is relatively well-studied (Heyes, 2011), it is less clear how this automatic tendency is brought under control.

Evidence that task representation plays a pivotal role in shaping our actions has been provided by a series of studies (Newman-Norlund et al., 2007a,b; van Schie et al., 2008b; Poljac et al., 2009) in which participants were explicitly instructed to prepare imitative or complementary actions after viewing a virtual actor grasp a manipulandum using either a precision grip (PG; i.e., opposition between the index finger and thumb) or a whole-hand grasp (WHG; i.e., opposition of the thumb with the other fingers). As expected, participants were faster at preparing their response in imitative contexts if the action to be carried out was congruent with what they had observed. When, instead, they were expected to carry out complementary actions, they responded faster when their action was dissimilar to the one they had just observed. The task representation (imitative vs. complementary) seems then to overrule long-term stimulus-response associations, influencing the way that action-perception coupling takes place. Further evidence concerning this flexible perception-action coupling was produced by a 3D motion capture study (Ocampo and Kritikos, 2010) in which reaching and grasping parameters of congruent responses were found to improve in imitative contexts, and incongruent responses were facilitated in complementary

contexts. Consistent with these findings, Longo et al. (2008) demonstrated that also the level of action coding can be modified (e.g., toward coding in terms of movements) depending on task requirements. Taken together, these data challenge the idea that action observation automatically leads to imitation in the observer and suggest that, depending on the context, observed actions can prime incongruent responses.

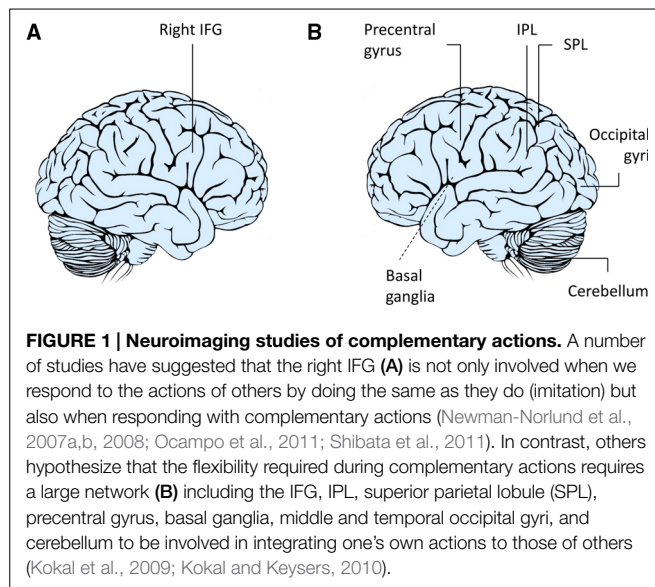
Recently, Sacheli et al. (2012, 2013) showed that participants involved in face-to-face interactions can mutually adjust their movements in time and space even in the absence of instructions to either imitate or perform a complementary response. This demonstrates that priming does not strictly depend on task-constraints, and that humans might indeed be able to actively shift from imitative to complementary actions, thanks to neuro-cognitive processes that still needs to be clarified.

Neuroimaging Studies of Complementary Actions

Few studies have examined the neural circuitry behind joint actions, and in particular the human mirror neuron system's (hMNS) involvement in complementary forms of social interaction. Might the hMNS provide a substrate for complementary actions? And if not, what role do other brain systems play?

In a pioneering experiment, the response of the hMNS was specifically investigated in imitative and complementary action contexts using functional magnetic resonance imaging (fMRI; Newman-Norlund et al., 2007a,b). Signals were recorded while the participants prepared to grasp a manipulandum in one of two ways—with a WHG or a PG—after they viewed an actor carrying out that action. It was found that preparation for complementary actions resulted in an increased blood-oxygen-level-dependent (BOLD) signal in the right inferior frontal gyrus (IFG) and in the bilateral inferior parietal lobule (IPL), two core components of the mirror system (**Figure 1**). This finding can be explained in terms of *different kinds* of mirror neurons: strictly congruent mirror neurons, which respond to identical actions, both observed and performed ones, and broadly congruent mirror neurons, which respond to non-identical observed and performed actions and objects linked to them (Fogassi and Gallese, 2002). It is also possible that in the complementary condition, when participants observe an action drawing attention to an object eliciting a different action, an interplay takes place between mirror and canonical neurons with the latter responding both during the time the action is being executed and also while the objects linked to those behaviors are perceived (Rizzolatti and Craighero, 2004). The need to carry out a complementary action involving a different object might then imply a combination of mirror and canonical neurons coding for different types of actions at different times of the sequence. The hypothesis that different classes of mirror neurons serve to integrate observed and executed actions during complementary kinds of social interaction is certainly an appealing one.

Newman-Norlund et al. (2007a,b, 2008) also hypothesized that a joint action could preferentially recruit right lateralized components of the mirror system since right inferior frontal activations are linked to inhibition processes (Brass et al., 2005). Planning and executing complementary actions in this



framework would mean, first of all, actively inhibiting the natural tendency to imitate observed actions. In the light of recent debates revolving around mirror mechanisms (Gallese and Sinigaglia, 2011; de Bruin and Gallagher, 2012), some have theorized that mirror neurons transform perceptual information regarding an intentional action in terms of the observer's own action possibilities (Gazzola et al., 2007). The idea that the hMNS could link perceived actions with appropriate motor plans was confirmed by an fMRI study designed by Ocampo et al. (2011) who studied the neural activations underlying execution of actions that were unlike the ones observed. As expected, activity within the right IPL and right IFG—core regions of the hMNS—was greatest in the imitative context when the participants responded with actions that were similar to the hand actions observed. Interestingly, activity within these regions also increased when dissimilar actions were performed, indicating that there are increased demands linked to remapping stimulus-response associations (Figure 1A). Shibata et al. (2011) likewise found that the right IFG was involved in mediating higher-order action understanding linked to a complementary action request. Overall, these findings seem to suggest that there are two separate processes and that both are supported by fronto-parietal brain regions. The first process operates at a simple motor level within contexts that require similar responses. The second allows the observer to inhibit those responses and to prepare an action that is compatible with the task demands at hand.

A more integrated description of neural circuits underlying complementary actions was recently outlined by Kokal et al. (2009; Kokal and Keysers, 2010; Figure 1B). Participants in an interactive fMRI study were instructed to carry out complementary and imitative actions in real-time cooperation with an experimenter ("Joint Action"), by performing the same actions individually ("Execution"), or by simply observing the experimenter's actions ("Observation"). This experiment raised our understanding of social interactions to an entirely new level by specifically mapping the contribution of the hMNS (i.e., common

voxels for both execution and observation) as well as the areas specifically involved in the joint actions (i.e., voxels exceeding the sum of execution and observation). The areas responsible for this integration process were located bilaterally in the IFG, IPL, precentral gyrus, superior parietal lobule, middle and temporal occipital gyri, and cerebellum.

Two anatomically separate networks have thus been delineated: one that would decipher observed and executed actions into a single common code (Etzel et al., 2008) and another that would integrate this information to successfully achieve common goals. These findings show that although the hMNS plays a critical role in translating all actions into a common code, their flexible remapping seems to be performed elsewhere. It would seem then that any potential discrepancy between an observed action and a complementary response can be resolved flexibly in a two-step manner. During the first step, the observed action is processed in order to predict its goal. During the second, associations are made between the action observed and the appropriate response needed to accomplish a complementary goal. Crucially, Erlhagen et al. (2006) proposed an anatomical model based on animal studies differentiating direct (automatic) and flexible routes for action-perception coupling. The model involves four interconnected brain areas, namely the superior temporal sulcus (STS), area PF (Brodmann area 7b), area F5, and the prefrontal cortex (PFC). The STS-F5 connection, allowing for the matching between a visual description of an action and its motor representation, would represent the neural basis of the direct route for the automatic imitation of an observed action. More importantly, when required, the flexible action-perception coupling is realized in the model by the connection between the PF area and the PFC through which goal representations from the PFC can modulate and set the coupling between visual (STS) and motor (F5) representations (Erlhagen et al., 2006).

Notably, the temporal course of the low- and high-level systems interaction has long been debated.

If output from control systems guide and modulates the mirror system, this would represent a top-down process. The STORM model (social top-down response modulation) suggests that the decision to imitate or to inhibit imitation initially draws on social signals and is most likely supported by a brain network including medial Prefrontal Cortex (mPFC) and temporoparietal junction (TPJ), two core areas of the so-called Mentalizing system, engaged when participants judge other people's mental state (Wang and Hamilton, 2012; Hamilton, 2015). Recently, Cross et al. (2013) have proposed a model for conflict processing in case of incongruence between observed and executed actions. When preparation to avoid imitation is not possible, medial prefrontal regions (mPFC and anterior cingulate cortex) would first detect imitative conflict and send information to anterior insula which would process conflict resolution, suppressing the unwanted motor activation. The hMNS would be therefore the target of top-down mechanisms of conflict resolution. In contrast, if the hMNS leads to an automatic tendency to imitate an observed action and this information feeds up to a monitoring system, this represents a bottom-up process (Brass et al., 2009). According to Ubaldi et al. (2015; see also Barchiesi and Cattaneo, 2013), early mirror responses (150 ms from onset of visual stimuli) would

be followed by later rule-based non-mirror responses (300 ms). These data seem to indicate that a fast bottom-up process mediated by the dorsal visual stream produces automatic imitative responses. Whereas rule-based visuomotor associations would be mediated by a slower top-down system, relying on the PFC.

Neurophysiologic Studies of Complementary Actions

Action observation automatically activates corresponding motor representations in an observer, and the stronger support for this process comes from single-pulse transcranial magnetic stimulation (spTMS) over the primary motor cortex (M1) and concomitant electromyography (EMG; e.g., Fadiga et al., 1995). This technique allows to measure modulations in an observer's corticospinal (CS) excitability while he/she watches an agent performing an action. A statistically significant increase in TMS-induced motor evoked potentials (MEP) amplitudes in the corresponding muscles indicates that observers are specifically attuned to the observed action and at what time it does occur. The facilitation of CS excitability provided the first physiological evidence for a direct matching in humans between action perception and action execution (for review, see Fadiga et al., 2005), and made it possible to explore motor system reactions in interactive contexts. A series of recent neurophysiologic studies were designed to assess the facilitation of CS excitability while participants observed video-clips evoking imitative and complementary gestures (Sartori et al., 2011b, 2012, 2013a,b,c). In one of these studies (Sartori et al., 2012), TMS-induced MEPs were recorded from the participants' hand muscles while they observed an actor grasping an object and then unsuccessfully attempting to complete a task (e.g., pouring coffee in a cup which was strategically placed out of her reach but in the video foreground, close to the observer's right hand). An almost imperceptible movement of the actor's hand was interpreted as a request to move the out-of-reach cup closer to the actor so that she could complete the action (Figure 2). Notably, the type of grasp the participant observed and the one that was needed to complete the actor's task were mismatched in all of the videos (i.e., a WHG performed by the actor vs. a PG required of the observer, and *vice versa*). As the participants were instructed to remain motionless throughout the task, the degree to which the motor system was activated provided an index of the CS activity elicited by action preparation. Moreover, as no explicit instructions were imparted to the participants, the experiment uncovered spontaneous tendencies to fulfill an implicit request embedded in a social interaction. This experiment was particularly enlightening in view of the fact that most studies typically ask participants to perform actions that are not associated with any meaningful behavior in real-world settings or utilize paradigms aiming to uncover dispositions formed during the execution of the experimental task itself (e.g., in imitation vs. complementary blocks) rather than spontaneous tendencies. Study results showed that a matching mechanism at the beginning of an action sequence turned into a complementary one as soon as the request for a reciprocal action became evident (*functional shift*). The muscle-specificity of MEP recordings highlighted the interplay between the initial tendency to resonate with what was observed and

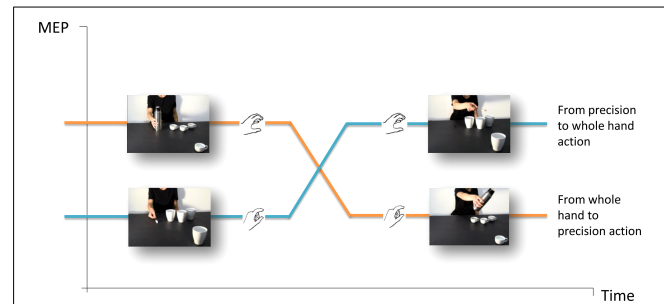


FIGURE 2 | The functional shift. A fundamental requirement for successful complementary actions is the capacity to smoothly and efficiently switch from observing another person's gestures to planning a corresponding reciprocal action. TMS-induced MEPs were recorded from participants' hand muscles in response to observing an actor grasping an object and then trying vainly to fulfill a task (e.g., pouring coffee) in a cup which was strategically placed out of her reach but in the video foreground, close to the observer's right hand (Sartori et al., 2012, 2013b,c). The type of grasp observed and the one that was required were reciprocally mismatched in all the videos (i.e., a WHG performed by the actor vs. a PG requested of the observer, and *vice versa*).

the subsequent inclination to implicitly prepare for a dissimilar complementary action (Figure 2).

At this point a new important question arose: at what point does this functional switch occur? A new experiment was designed in which TMS was delivered at *five* different timepoints corresponding to five kinematic landmarks characterizing the observed action (Sartori et al., 2013b,c). The most critical was the fourth (T_4) timepoint when the actor's hand trajectory began to significantly move toward the out-of-reach object. A TMS pulse was delivered precisely at that moment to investigate whether participants were able to predict the actor's trajectory even before the action became explicit. The control condition that was designed consisted in the actor bringing her hand back to its initial position—with the out-of-reach object still visible in the foreground. The results showed that the participants quickly discriminated between an action driven by a social goal and one that was not, simply by observing the kinematic cues signaling the direction of the actor's hand. These findings have direct implications with regard to action representation theories because they suggest that intention attribution (i.e., social vs. individual) is sensitive to kinematic constraints. As different types of intentional actions have distinctive motion signatures, observers appear to take note of precocious differences in kinematics during action observation in order to be able to predict the actor's intentions (Kilner et al., 2007; Sartori et al., 2009, 2011a; Becchio et al., 2010, 2012a,b; Manera et al., 2011).

A Working Memory Hypothesis

A dual process seems then to underlie joint actions: a low-level motor resonance analyzes and stores information on observed actions, while a high-level system would flexibly integrate our and others' motor intentions and select the most appropriate response and time course to achieve joint goals (van Schie et al., 2008a). It can be hypothesized then that the hMNS' function is similar to that of a working memory, although specifically tailored for

action. Mirroring the responses of others might be useful to constantly track and monitor the interacting partners, and to support temporal coordination and action planning (Colling et al., 2013) while cognitive control systems come into play to distinguish self- and other-related representation, to inhibit unwanted imitative responses and to enforce self-generated actions (Brass et al., 2009; Cross et al., 2013). As in the case of the working memory, distinct elements would be kept on-line while others are being processed (Gibson, 2000). We therefore suggest an extension of the previous models of imitation control involving a cross-talk and a *simultaneous* activation of low- and high-level systems.

Complementary actions are the ideal way to test this topic. During complex social interaction/s, the individual needs to keep information relative to the observed action available while contemporaneously attempting to process a response. In this type of context, the mirror system may be involved in keeping action-related information on hold to enable other brain areas to extract the meaning of the action observed to achieve a joint goal. Notably, observing another person's actions priming for an incongruent reaction can lead to a motor resonant response in the observer's corresponding muscles as well as a simultaneous preparation in different effectors necessary for achieving a complementary response (Sartori et al., 2015). The relation between observed and executed actions seems then to be coordinated by a social associative memory which apparently matches some actions to their natural social responses regardless of who is actually performing the action (Chinellato et al., 2013). Under this model, there would be no difference between congruent, incongruent and complementary responses, as long as they have been associatively linked. In this vein, Catmur et al., (2007; 2008; 2009; see also Heyes, 2001, 2011; Cooper et al., 2013) have proposed that flexibility in action perception

coupling may be gained thanks to associative sequence learning (i.e., the ASL theory) developed during social interactions. They strongly suggests that overlearned responses are able to modulate the motor priming effect: when a specific behavior is contingent on a non-matching behavior (e.g., extending the right hand when observing a right hand), an incongruent association is formed.

Conclusion

The research outlined here shows that motor resonance elicited by action observation is modulated depending on its context: when an observed gesture is socially relevant (i.e., there is an implicit or explicit request) anticipatory complementary activations follow. The assumption that observing an action automatically triggers the inclination to imitate probably arose because most studies did not explicitly challenge the automaticity or flexibility of the visuomotor transformation process. The data outlined here have contributed to shedding light on the functioning of the human motor system in social contexts and on the types of social behavior frequently occurring in real-world settings. From a wider perspective, we can theorize that defining the conditions and the modalities by which motor resonant responses to action observation can be modulated may prove to have specific translational implications leading to the development of novel neuro-rehabilitation protocols for patients with localized lesions to cortical motor areas (e.g., ischemic stroke) and for pathologies such as autism (Hamilton, 2015). More distant horizons may include developing models of brain mechanisms underlying social interactions in the effort to endow artificial agents such as robots with the ability to perform meaningful complementary actions in response to observed actions.

References

- Barchiesi, G., and Cattaneo, L. (2013). Early and late motor responses to action observation. *Soc. Cogn. Affect. Neurosci.* 8, 711–719. doi: 10.1093/scan/nss049
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., and Castiello, U. (2012a). Grasping intentions: from thought experiments to empirical evidence. *Front. Hum. Neurosci.* 6:117. doi: 10.3389/fnhum.2012.00117
- Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., and Castiello, U. (2012b). Social grasping: from mirroring to mentalizing. *Neuroimage* 61, 240–248. doi: 10.1016/j.neuroimage.2012.03.013
- Becchio, C., Sartori, L., and Castiello, U. (2010). Toward you: the social side of actions. *Curr. Dir. Psychol. Sci.* 19, 183–188. doi: 10.1177/0963721410370131
- Brass, M., Derrfuss, J., and Von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: a functional double dissociation. *Neuropsychologia* 43, 89–98. doi: 10.1016/j.neuropsychologia.2004.06.018
- Brass, M., Ruby, P., and Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2359–2367. doi: 10.1098/rstb.2009.0066
- Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., and Heyes, C. M. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. *Eur. J. Neurosci.* 28, 1208–1215. doi: 10.1111/j.1460-9568.2008.06419.x
- Catmur, C., Walsh, V., and Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2369–2380. doi: 10.1098/rstb.2009.0048
- Catmur, C., Walsh, V., and Heyes, C. M. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531. doi: 10.1016/j.cub.2007.08.006
- Chinellato, E., Ognibene, D., Sartori, L., and Demiris, Y. (2013). “Time to change: deciding when to switch action plans during a social interaction,” in *Biomimetic and Biohybrid Systems*, eds N. F. Lepora, A. Mura, H. G. Krapp, P. F. M. J. Verschure, and T. J. Prescott (Berlin: Springer), 47–58.
- Colling, L. J., Knoblich, G., and Sebanz, N. (2013). How does “mirroring” support joint action. *Cortex* 49, 2964–2965. doi: 10.1016/j.cortex.2013.06.006
- Cooper, R. P., Cook, R., Dickinson, A., and Heyes, C. M. (2013). Associative (not Hebbian) learning and the mirror neuron system. *Neurosci. Lett.* 540, 28–36. doi: 10.1016/j.neulet.2012.10.002
- Costantini, M., Committeri, G., and Sinigaglia, C. (2011). Ready both to your and to my hands: mapping the action space of others. *PLoS ONE* 6:e17923. doi: 10.1371/journal.pone.0017923
- Cross, K. A., Torrisi, S., Losin, E. A. R., and Iacoboni, M. (2013). Controlling automatic imitative tendencies: interactions between mirror neuron and cognitive control systems. *Neuroimage* 83, 493–504. doi: 10.1016/j.neuroimage.2013.06.060
- de Bruin, L., and Gallagher, S. (2012). Embodied simulation, an unproductive explanation: comment on Gallese and Sinigaglia. *Trends Cogn. Sci.* 16, 98–99. doi: 10.1016/j.tics.2011.12.003
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain. Res.* 91, 176–180. doi: 10.1007/BF00230027
- Erlhagen, W., Mukovskiy, A., and Bicho, E. (2006). A dynamic model for action understanding and goal-directed imitation. *Brain Res.* 1083, 174–188. doi: 10.1016/j.brainres.2006.01.114
- Etzet, J. A., Gazzola, V., and Keysers, C. (2008). Testing simulation theory with cross-modal multivariate classification of fMRI data. *PLoS ONE* 3:e3690. doi: 10.1371/journal.pone.0003690

- Fadiga, L., Craighero, L., and Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. *Curr. Opin. Neurobiol.* 15, 213–218. doi: 10.1016/j.conb.2005.03.013
- Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* 73, 2608–2611.
- Fogassi, L., and Gallese, V. (2002). The neural correlates of action understanding in non-human primates. *Adv. Conscious. Res.* 42, 13–36. doi: 10.1075/aicr.42.03fog
- Gallese, V. (2001). The 'shared manifold' hypothesis. From mirror neurons to empathy. *J. Conscious. Stud.* 8, 33–50.
- Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends. Cogn. Sci.* 15, 512–519. doi: 10.1016/j.tics.2011.09.003
- Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., and Keysers, C. (2007). Apathics born without hands mirror the goal of hand actions with their feet. *Curr. Biol.* 17, 1235–1240. doi: 10.1016/j.cub.2007.06.045
- Gibson, E. (2000). "The dependency locality theory: a distance-based theory of linguistic complexity" in *Image, Language, Brain*, eds Y. Miyashita, A. Marantz, and W. O'Neil (Cambridge, MA: MIT Press), 95–126.
- Graf, M., Schütz-Bosbach, S., and Prinz, W. (2009). *Motor Involvement in Action and Object Perception: Similarity and Complementarity. Grounding Sociality: Neurons, Minds, and Culture*. New York, NY: Psychology Press.
- Hamilton, A. F. de C. (2015). The neurocognitive mechanisms of imitation. *Curr. Opin. Behav. Sci.* 3, 63–67. doi: 10.1016/j.cobeha.2015.01.011
- Heyes, C. (2001). Causes and consequences of imitation. *Trends Cogn. Sci.* 5, 253–261. doi: 10.1016/S1364-6613(00)01661-2
- Heyes, C. (2011). Automatic imitation. *Psychol. Bull.* 137, 463–83. doi: 10.1037/a0022288
- Kilner, J. M., Friston, K. J., and Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cogn. Process.* 8, 159–166. doi: 10.1007/s10339-007-0170-2
- Knoblich, G., Butterfill, S., and Sebanz, N. (2011). "Psychological research on joint action: theory and data," in *The Psychology of Learning and Motivation—Advances in Research and Theory*, ed. B. Ross (London: Academic Press), 59–101.
- Knoblich, G., and Flach, R. (2001). Predicting the effects of actions: interactions of perception and action. *Psychol. Sci.* 12, 467–472. doi: 10.1111/1467-9280.00387
- Kokal, I., Gazzola, V., and Keysers, C. (2009). Acting together in and beyond the mirror neuron system. *Neuroimage* 47, 2046–2056. doi: 10.1016/j.neuroimage.2009.06.010
- Kokal, I., and Keysers, C. (2010). Granger causality mapping during joint actions reveals evidence for forward models that could overcome sensory-motor delays. *PLoS ONE* 5:e13507. doi: 10.1371/journal.pone.0013507
- Longo, M. R., Kosobud, A., and Bertenthal, B. I. (2008). Automatic imitation of biomechanically possible and impossible actions: effects of priming movements versus goals. *J. Exp. Psychol. Hum.* 34, 489–501. doi: 10.1037/0096-1523.34.2.489
- Manera, V., Becchio, C., Cavallo, A., Sartori, L., and Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Exp. Brain Res.* 211, 547–556. doi: 10.1007/s00221-011-2649-4
- Newman-Norlund, R. D., Bosga, J., Meulenbroek, R. G., and Bekkering, H. (2008). Anatomical substrates of cooperative joint-action in a continuous motor task: virtual lifting and balancing. *Neuroimage* 41, 169–177. doi: 10.1016/j.neuroimage.2008.02.026
- Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. G., and Bekkering, H. (2007a). Exploring the brain basis of joint action: co-ordination of actions, goals and intentions. *Soc. Neurosci.* 2, 48–65. doi: 10.1080/17470910701224623
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., and Bekkering, H. (2007b). The mirror neuron system is more activated during complementary compared with imitative action. *Nat. Neurosci.* 10, 817–818. doi: 10.1038/nn1911
- Ocampo, B., and Kritikos, A. (2010). Placing actions in context: motor facilitation following observation of identical and non-identical manual acts. *Exp. Brain Res.* 201, 743–751. doi: 10.1007/s00221-009-2089-6
- Ocampo, B., Kritikos, A., and Cunningham, R. (2011). How frontoparietal brain regions mediate imitative and complementary actions: an fMRI study. *PLoS ONE* 6:e26945. doi: 10.1371/journal.pone.0026945
- Poljac, E., van Schie, H. T., and Bekkering, H. (2009). Understanding the flexibility of action-perception coupling. *Psychol. Res.* 73, 578–586. doi: 10.1007/s00426-009-0238-y
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., and Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706. doi: 10.1152/physrev.00009.2013
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230
- Sacheli, L. M., Candidi, M., Pavone, E. F., Tidoni, E., and Aglioti, S. M. (2012). And yet they act together: interpersonal perception modulates visuo-motor interference and mutual adjustments during a joint-grasping task. *PLoS ONE* 7:e50223. doi: 10.1371/journal.pone.0050223
- Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., and Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Exp. Brain Res.* 226, 473–486. doi: 10.1007/s00221-013-3459-7
- Sartori, L., Becchio, C., Bulgheroni, M., and Castiello, U. (2009). Modulation of the action control system by social intention: unexpected social requests override preplanned action. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1490–1500. doi: 10.1037/a0015777
- Sartori, L., Becchio, C., and Castiello, U. (2011a). Cues to intention: the role of movement information. *Cognition* 119, 242–252. doi: 10.1016/j.cognition.2011.01.014
- Sartori, L., Cavallo, A., Buccioni, G., and Castiello, U. (2011b). Corticospinal excitability is specifically modulated by the social dimension of observed actions. *Exp. Brain Res.* 211, 557–568. doi: 10.1007/s00221-011-2650-y
- Sartori, L., Betti, S., and Castiello, U. (2013a). When mirroring is not enough: that is, when only a complementary action will do (the trick). *Neuroimage* 24, 601–604. doi: 10.1097/WNR.0b013e3283630a66
- Sartori, L., Betti, S., and Castiello, U. (2013b). Corticospinal excitability modulation during action observation. *J. Vis. Exp.* 82, e51001. doi: 10.3791/51001
- Sartori, L., Buccioni, G., and Castiello, U. (2013c). When emulation becomes reciprocity. *Soc. Cogn. Affect. Neurosci.* 8, 662–669. doi: 10.1093/scan/nss044
- Sartori, L., Betti, S., Chinellato, E., and Castiello, U. (2015). The multiform motor cortical output: kinematic, predictive and response coding. *Cortex* doi: 10.1016/j.cortex.2015.01.019. [Epub ahead of print].
- Sartori, L., Cavallo, A., Buccioni, G., and Castiello, U. (2012). From simulation to reciprocity: the case of complementary actions. *Soc. Neurosci.* 7, 146–158. doi: 10.1080/17470919.2011.586579
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76. doi: 10.1016/j.tics.2005.12.009
- Shibata, H., Inui, T., and Ogawa, K. (2011). Understanding interpersonal action coordination: an fMRI study. *Exp. Brain Res.* 211, 569–579. doi: 10.1007/s00221-011-2648-5
- Ubaldo, S., Barchiesi, G., and Cattaneo, L. (2015). Bottom-up and top-down visuomotor responses to action observation. *Cereb. Cortex* 25, 1032–1041. doi: 10.1093/cercor/bht295
- van Schie, H. T., Koelewijn, T., Jensen, O., Oostenveld, R., Maris, E., and Bekkering, H. (2008a). Evidence for fast, low-level motor resonance to action observation: an MEG study. *Soc. Neurosci.* 3, 213–228. doi: 10.1080/17470910701414364
- van Schie, H. T., Van Waterschoot, B. M., and Bekkering, H. (2008b). Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1493–1500. doi: 10.1037/a0011750
- Wang, Y., and Hamilton, A. F. D. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Front. Hum. Neurosci.* 6:153. doi: 10.3389/fnhum.2012.00153

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Sartori and Betti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Grasping actions and social interaction: neural bases and anatomical circuitry in the monkey

Stefano Rozzi* and Gino Coudé

Department of Neuroscience, University of Parma, Parma, Italy

OPEN ACCESS

Edited by:

Claudia Gianelli,
University of Potsdam, Germany

Reviewed by:

Alex Pitti,
University of Cergy-Pontoise, France
Umberto Castiello,
Università di Padova, Italy

*Correspondence:

Stefano Rozzi,
Department of Neuroscience,
University of Parma, Via Volturmo
39-E, 43125 Parma, Italy
stefano.rozzi@unipr.it

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 01 April 2015

Accepted: 29 June 2015

Published: 14 July 2015

Citation:

Rozzi S and Coudé G (2015)
Grasping actions and social
interaction: neural bases and
anatomical circuitry in the monkey.
Front. Psychol. 6:973.
doi: 10.3389/fpsyg.2015.00973

The study of the neural mechanisms underlying grasping actions showed that cognitive functions are deeply embedded in motor organization. In the first part of this review, we describe the anatomical structure of the motor cortex in the monkey and the cortical and sub-cortical connections of the different motor areas. In the second part, we review the neurophysiological literature showing that motor neurons are not only involved in movement execution, but also in the transformation of object physical features into motor programs appropriate to grasp them (through visuo-motor transformations). We also discuss evidence indicating that motor neurons can encode the goal of motor acts and the intention behind action execution. Then, we describe one of the mechanisms—the mirror mechanism—considered to be at the basis of action understanding and intention reading, and describe the anatomo-functional pathways through which information about the social context can reach the areas containing mirror neurons. Finally, we briefly show that a clear similarity exists between monkey and human in the organization of the motor and mirror systems. Based on monkey and human literature, we conclude that the mirror mechanism relies on a more extended network than previously thought, and possibly subserves basic social functions. We propose that this mechanism is also involved in preparing appropriate complementary response to observed actions, allowing two individuals to become attuned and cooperate in joint actions.

Keywords: motor, mirror neurons, intention, motor goal, grasping, parietal

Introduction

Over the last 50 years, sensorimotor neuroscience has produced an extensive body of work dedicated to the study of grasping. The motor act of grasping is multifaceted and lies at the crossroad between action and perception. Here, a distinction should be drawn between the grasping motor act and the action of grasping. A grasping motor act can be defined as a series of joint movements, like clasping the fingers on an object, aimed at achieving the motor goal of seizing. An action of grasping consists in a sequence of fluently linked motor acts that altogether are aiming at the achievement of an overarching behavioral goal. For instance, a grasping action would consist in reaching, grasping a fruit and bringing it to the mouth for eating. Under normal circumstances, a grasping motor act is executed within a sequence of other motor acts together forming a grasping action. Such an action can be driven by a wide gamut of needs and aimed at a variety of overarching goals such as feeding, exploring the environment, or interacting with other individuals. Interestingly, socially appropriate behaviors require a continuous monitoring of the social environment. Accordingly, numerous studies both on monkey and human focused on analyzing the motor behavior, especially

grasping actions, to investigate basic social interactions. Altogether, these studies demonstrate that grasping is modulated by the social context in which it occurs. This, in turn, implies that the motor system, that actually produces the behavior itself, is involved in a larger network encoding social aspects of real life. However, the neural mechanisms at the basis of this coupling between social cognition and motor behavior have not yet been fully unveiled. In this review, we describe the basic neural mechanisms underpinning grasping and show how these same mechanisms are also at the bases of cognitive abilities that are basic aspects of social cognition such as action understanding and intention reading. This paper mainly focus on the anatomical and functional literature based on the macaque monkey model. Indeed, monkeys have been used for brain studies since the beginning of the twentieth century (Brodmann, 1909; Bucy, 1933, 1935; Fulton, 1935), and we owe to monkey studies a huge part of our knowledge on the neuroanatomy and neurophysiology of the motor system. This is especially true for the neural bases and the anatomical circuitry involved in grasping. Monkeys are capable of using their hands for grasping in a way that is very similar to humans. Evolution is opportunistic and conservative: working mechanisms tend to be retained through generations and species and novelty tends to be built by adapting extant material and processes to the new demands. Without denying the obvious gap existing between the monkey brain and the human brain (see Iriki and Sakura, 2008; Passingham, 2009), we think that the macaque model remains invaluable for the anatomophysiological study of grasping. Thus, in the first part of the paper we describe the anatomical structure of the motor cortex in the monkey and the cortical and sub-cortical connections of the areas forming it. In the second section, we review the monkey functional literature showing two important aspects: (1) that motor neurons are not only involved in movement execution, but also in the sensory-motor transformation for grasping, and (2) that a population of these neurons encodes the goal of grasping motor acts and the motor intention behind action execution. In the third part, we describe one of the mechanisms—the mirror mechanism—considered to be at the basis of action understanding and intention reading. In particular we discuss how important aspects of the social environment such as the spatial representation of self, objects and others, modulate the motor and mirror neurons activity, influencing monkeys behavioral responses. In the last part, we briefly show that mechanisms similar to those described in the monkey are also present in the homo species.

Anatomy of the Motor System

The Motor Cortex: General Organization

At the end of the nineteenth century the general view of the organization of the motor system was that the movements were controlled by subcortical centers, while the cerebral cortex was involved cognitive functions. This view was challenged by pioneering studies demonstrating that the electrical stimulation of a specific part of the frontal cortex (motor cortex) evoked body movements in different species of animals (Fritsch and Hitzig, 1870; Ferrier, 1873; see Porter and Lemon, 1993). The idea that

the motor cortex contains a simple map of the muscles was in line with Jackson's observations on the epileptic seizures in human patients. At the beginning of the twentieth century, Campbell (1905) identified a possible anatomical substrate accounting for Jackson's observations in his architectonic map of the human cerebral cortex. Campbell's (1905) view was that the precentral cortex was implicated in motor control, while the intermediate sector was involved in what will be later called "higher order motor functions." A similar view emerged from Brodmann's (1909) work, where he confirmed the existence of two motor areas, area 4 and area 6, and provided a more detailed map of the frontal lobe both in monkeys and humans. The idea that architectonic differences reflects functional specificity was later supported by Fulton (1935), who showed that the ablation of area 6 produces specific deficits in the execution of skilled movements. This observation led him to refer to this region as *premotor cortex*. However, a few years later, Woolsey's electrophysiological studies (Woolsey and Settlage, 1952) casted doubts about the existence of a high order motor area rostral to area 4, and led him to conclude that area 4 and posterior area 6 form together a functional entity, while the not electrically excitable rostral area 6 does not belong of the motor cortex.

Brodmann's definition of area 6 as a single architectonic entity was also challenged by subsequent anatomical studies in which this sector was divided in different areas (e.g., Vogt and Vogt, 1919; Von Bonin and Bailey, 1947; Barbas and Pandya, 1987). Recently, a more objective assessment of areal borders was provided by combining cytoarchitectonic and neurochemical techniques (see Geyer et al., 2000; Belmalih et al., 2007). This multiarchitectonic approach yielded a more refined map of the motor cortex of the macaque monkey (Figure 1; Matelli et al., 1985, 1991; Belmalih et al., 2009). In this parcellation, area F1 roughly corresponds to Brodmann's area 4 (primary motor cortex), whereas the mesial, dorsal, and ventral sectors of Brodmann's area 6 are each divided into caudal and rostral areas. This parcellation has been further validated by converging functional evidence showing that each of these architectonic subdivisions are connectionally and functionally distinct. The resulting map indicates that the macaque motor cortex is a mosaic of distinct areas and contains a multiplicity of body movement representations, each emphasizing different categories of behavior and playing a specific role in motor control (see Rizzolatti et al., 1998). Thus, the "mapping from cortex to muscles is not fixed, as was once thought, but instead is fluid, changing continuously on the basis of feedback in a manner that could support the control of higher-order movement parameters" (Graziano, 2006).

Connections of the Motor Areas of the Monkey

Connectional studies are warranted to gather clues about their functional role and complete the picture etched through architectonic studies. By means of tract tracing studies, it has been shown that each motor area is characterized by a specific pattern of connections. Based on these general connectivity patterns, the premotor areas have been grouped into two major classes (Rizzolatti and Luppino, 2001): the caudal (F2, F3, F4, F5p, and F5c) and the rostral (F5a, F6, and F7) premotor areas. In the following paragraphs, we describe the descending and cortical

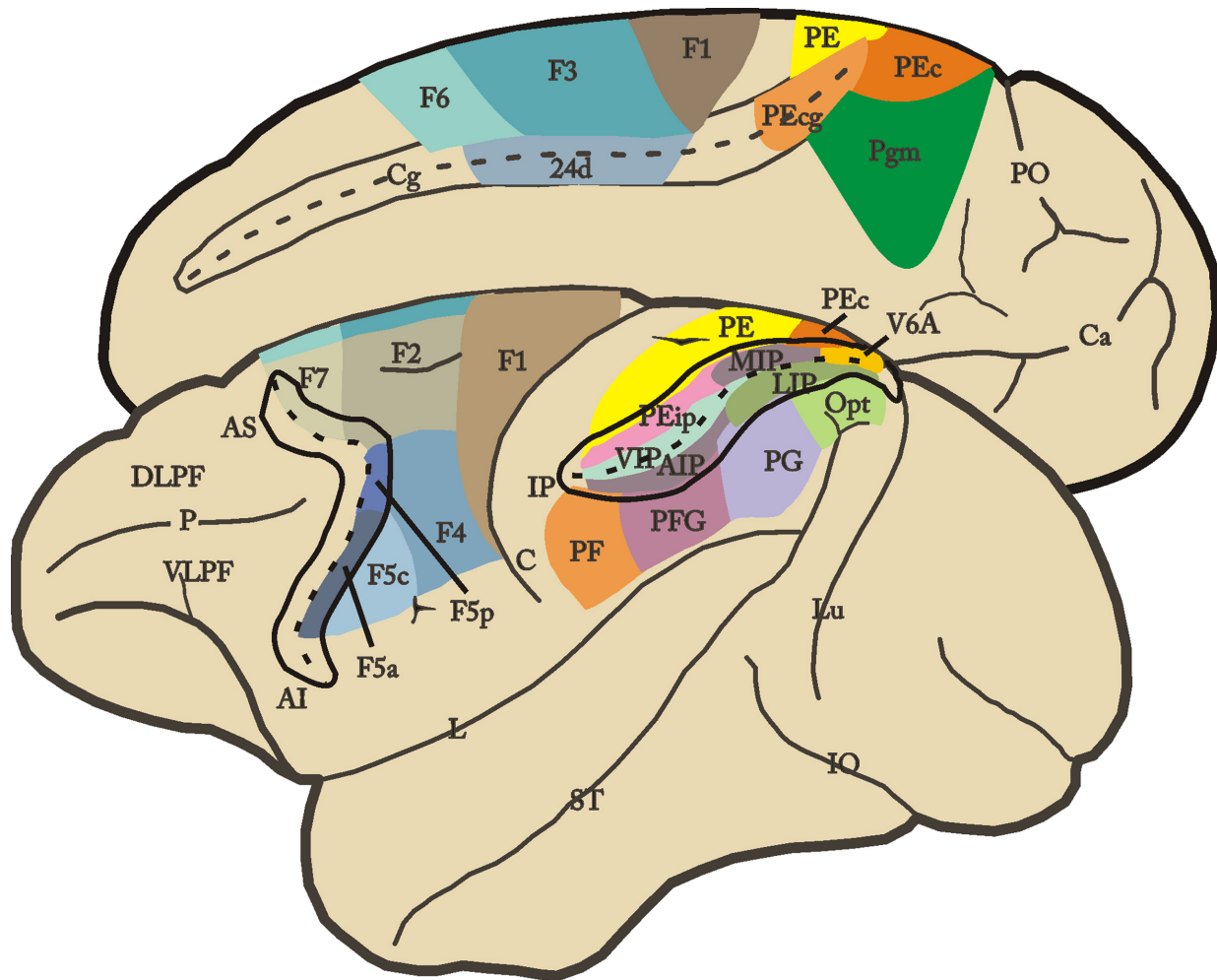


FIGURE 1 | Lateral and mesial views of the monkey brain showing the parcellation of the agranular frontal and posterior parietal cortex. Intraparietal, arcuate and cingulated sulci are shown unfolded. For the nomenclature and definition of the agranular frontal and posterior parietal areas, see text. AI, inferior arcuate sulcus; AS,

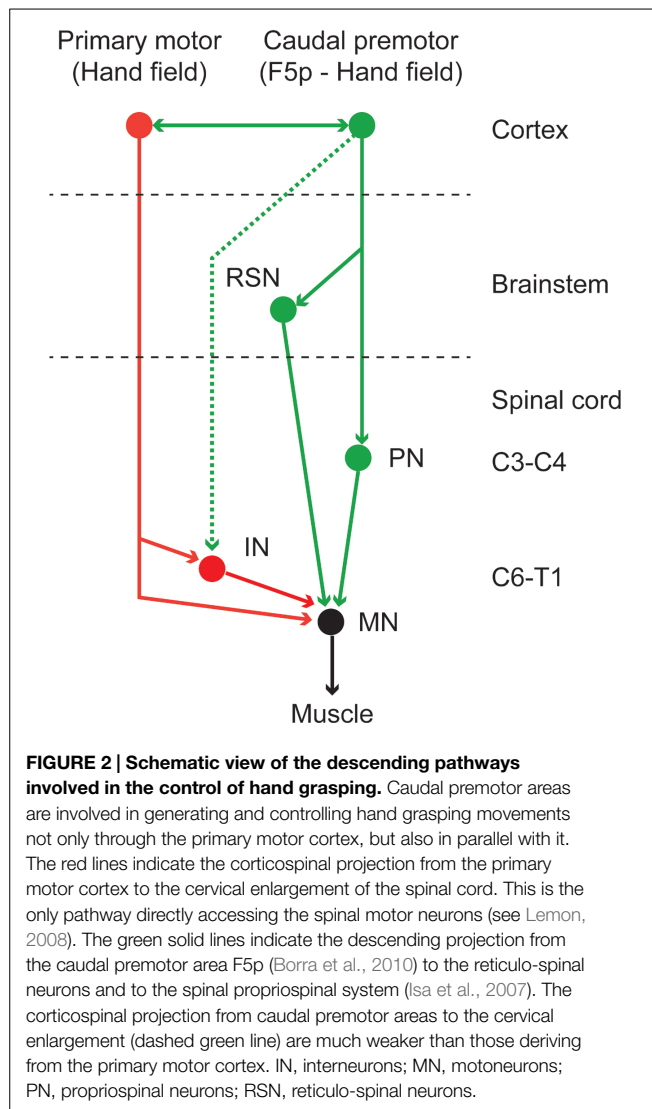
connections of these motor areas, and draw hypotheses on their possible functional role.

Descending Motor Pathways and Intrinsic Motor Connections

As a whole, the motor cortex is source of different descending motor pathways, each providing it with an access to the brainstem and spinal motor centers. Strick and coworkers (Dum and Strick, 1991; He et al., 1993, 1995) showed that the corticospinal projections are somatotopically organized and that originate both from the primary motor area and from all the caudal premotor areas. Similarly, the face and mouth cortical motor representations are sources of corticobulbar projections (Morecraft et al., 2001). The corticospinal projections mostly terminate in the intermediate zone of the spinal cord, and only F1 is source of monosynaptic projections to spinal motor neurons (Porter and Lemon, 1993). This means that F1 is the final common

pathway, at the cortical level, for controlling skilled movements. However, the presence of corticospinal projections from all the caudal premotor areas clearly indicates that these areas are also involved in generating and controlling movements, not only through F1, but also in parallel with it, as also confirmed by the evidence that each of them is also somatotopically connected with F1. For example, a descending indirect pathway connecting the caudal premotor area F5p with the cervical propriospinal system was recently described and is deemed to be involved in the control of dexterous fingers movements (Sasaki et al., 2004; Isa et al., 2007; Borra et al., 2010; see also Lemon, 2008; Alstermark and Isa, 2012). **Figure 2** depicts a schematic view of the descending pathways enabling hand motor control.

In contrast, none of the rostral premotor areas project directly to the spinal cord. Their descending projections reach different parts of the brainstem (Keizer and Kuypers, 1989). Furthermore, they are not directly connected with F1, and generally have a



widespread pattern of connections with other motor areas. The radically different pattern of descending projection characterizing rostral and caudal areas hints to the fact that they probably are subserving different functions. The rostral areas are thought to be only indirectly involved in the generation of motor behavior through their sub-cortical projections and through their cortical connections with the caudal premotor areas.

Cortico-Cortical Connections

The cortical connections of the frontal motor areas involve mainly two brain regions: parietal cortex and prefrontal cortex (see Rizzolatti et al., 1998, 2014; Rizzolatti and Luppino, 2001). The connections between frontal motor and posterior parietal areas are very strong and reciprocal. Anatomical and functional evidence show that the posterior parietal cortex consists in a mosaic of areas similar to the motor cortex (Figure 1), each area is involved in processing specific aspects of sensory information and controlling different effectors (e.g., mouth, hand, arm, and eyes). In general, most IPL areas and the posterior areas of the

SPL process either strictly visual or visual and somatosensory information, while the rostral areas of the SPL mainly deal with somatosensory information (Hyvärinen, 1982; Caminiti et al., 1996; Rizzolatti et al., 1997; Wise et al., 1997; Colby, 1998; Rizzolatti and Matelli, 2003; Rozzi et al., 2008). A series of largely segregated anatomical circuits linking parietal and motor areas can be identified according to the pattern of predominant connections. These circuits integrate specific motor and sensory signals and participate to particular aspects of sensory-motor transformations, and should be thus considered the functional units of the cortical motor system. The processing undertaken by these functional units results in the generation of *potential motor acts*. In the following section, we describe the anatomy and function of one of these circuits (AIP-F5), and discuss its role in transforming visual information about an object into potential motor acts appropriate to grasp it.

The second strongest source of cortical connections of the motor areas is the prefrontal cortex. Prefrontal connections primarily involve the rostral premotor areas (Barbas, 1988; Preuss and Goldman-Rakic, 1989; Luppino et al., 1993; Lu et al., 1994; Rizzolatti and Luppino, 2001; Saleem and Kondo, 2008; Gerbella et al., 2010, 2013; Borra et al., 2011). Specifically, the dorsal part of the lateral prefrontal cortex (DLPF) projects to F7, its ventral part (VLPF) projects to F5a, whereas both DLPF and VLPF project to F6. Our knowledge of the anatomo-functional organization of the prefrontal cortex is much less detailed than that of the parietal cortex. It is generally accepted that these regions are involved in “higher order” functions such as working memory, planning of actions, and motivation (see Miller and Cohen, 2001; Tanji and Hoshi, 2008). Thus, these projections could play a role in selecting the potential motor acts generated as the result of sensorimotor transformations, weighting their suitability according to context, abstract rules, memorized information, and behavioral goals. The interplay between prefrontal cortex and frontal motor areas could be at the basis of the transformation of *potential* actions into *actual* actions.

Functional Properties of Motor Neurons: From Grasping to Intention

Visuo-Motor Transformations for Grasping

Grasping requires the adjustment of hand conformation to the size and shape of an object. A very efficient way of to accomplish this duty has evolved in the motor system. It consists in a direct linkage between the representations of object physical features and of potential motor acts, allowing the capacity of coding objects in term of actions to execute upon them. The process of transforming object properties into corresponding potential grasping actions relies on a specific circuit linking parietal area AIP and premotor area F5. The neural properties of these areas have been widely studied. We know that area F5 contains purely motor and sensory-motor neurons, some of which responsive to the presentation of visual stimuli (Rizzolatti et al., 1988). These F5 visuo-motor neurons fall into two main classes: canonical and mirror neurons, although, recently, the additional hybrid class of “canonical-mirror” neurons has been identified (Bonini et al.,

2014). In this section, we will describe the properties and the functional role of the canonical neurons.

Canonical neurons are mostly located in area F5p and discharge during the presentation of 3D objects (Murata et al., 1997; Raos et al., 2006). They have been systematically studied by means of a paradigm that allows one to separate activity related to object presentation, action preparation and action execution. For the major part, canonical neurons selectively respond to objects of a certain size, shape and orientation. Typically, their visual and motor specificity are congruent, and it was demonstrated that their activity does not depend on attention to stimuli, intention to act, or motor preparation (Murata et al., 1997). The most likely explanation for the canonical neurons discharge proposes that object presentation activates a representation of the observed object in motor format. In other words, when an object appears in the visual scene, the discharge of a specific set of canonical neurons code a *potential grasping act* congruent with the physical properties of the presented object. Note that this occurs independently of whether the act will be actually executed or not. In support of this explanation is the observation that a canonical neuron can show a visual response of the same intensity to the presentation of objects of different shape that are grasped in the same way (Murata et al., 1997; Raos et al., 2006).

As mentioned above, F5 has strong anatomical connections with the AIP area (Luppino et al., 1999; Borra et al., 2008; Gerbella et al., 2011). The functional properties of the neurons located in this area have been studied using the same paradigm adopted for the study of F5 canonical neurons (Taira et al., 1990; Sakata et al., 1995; Murata et al., 2000). By using this paradigm, AIP neurons have been divided into three classes: *motor-dominant*, *visual and motor*, and *visual-dominant* neurons. *Motor-dominant* neurons discharge during grasping either if the action is performed in light or in darkness, but do not fire during simple object fixation. *Visual-dominant* neurons discharge during grasping in light and during object fixation, but not when grasping is performed in darkness. Finally, *visual and motor neurons* discharge stronger during grasping in light than in darkness, and also discharge during object fixation.

The evidence that AIP and F5 are nodes of a circuit involved in visuo-motor transformations for grasping was strongly supported by inactivation studies. In particular, the inactivation of either AIP (Gallese et al., 1994) or F5 (Fogassi et al., 2001) has been shown to cause important deficits in shaping the hand according to the stimulus physical characteristics during hand transport before landing on the object. Note that, once touched, the object is correctly grasped, thus showing the lack of pure motor deficits.

Several models have been proposed to explain the role of AIP and F5 in visuo-motor transformation for grasping (Taira et al., 1990; Jeannerod et al., 1995; Fagg and Arbib, 1998; Rizzolatti and Luppino, 2001; Fluet et al., 2010). Despite the fact that there is no complete agreement among these models, they share the common idea that when an object is presented, AIP neurons extract specific aspects of its intrinsic features and provide F5 with a multiple description of the possible ways to grasp it. This corresponds to what Gibson defined as *affordances* (Gibson, 1979). The lateral prefrontal cortex would activate a set of AIP

and F5 neurons according to the behavioral goal, object nature, and context. Indeed, an object can be grasped with various types of grip depending not only on its physical features, but also on the different behavioral contexts. For instance, in recent studies, monkeys were trained to associate two different grip-types with corresponding color cues. The results showed that in both AIP and F5, a set of neurons were active after cue presentation, showing context-dependent grasp planning activity (Baumann et al., 2009; Fluet et al., 2010). The information about the chosen grip, according to the models, would be then sent from F5 to F1, where the movements are coded, and the final command for the execution is generated. Indeed, recent physiological experiments demonstrated that the activation of F5 is able to generate object-oriented actions through the modulation of F1 motor output (Cerri et al., 2003; Shimazu et al., 2004; Prabhu et al., 2009). Note, however, that the existence of corticospinal projections from F5p (see above; Borra et al., 2010) indicates that this area could be involved in the generation and control of movements not only through F1, but also in parallel with it (**Figure 2**). In particular, the F5 connections with the cervical propriospinal neurons appears to be involved in the control of dexterous fingers movements (Isa et al., 2007; Kinoshita et al., 2012). The exact functional role of these projections is still only partially understood, but evidence suggest that they may play a role in the functional recovery observed after lesions of the motor cortex. In particular, in New World monkeys, the ventral premotor hand field expands and develops new cortical connections after lesions of the primary motor cortex (Frost et al., 2003; Dancause et al., 2005, 2006; see Nudo, 2007). In macaque, it has been shown that after intensive post-lesion motor training, the ventral premotor hand field (including F5p) undergoes plastic changes and shows recovery-related increases in activity (Nishimura et al., 2007; see Nishimura and Isa, 2012).

Coding Grasping Goal: The “Vocabulary of Motor Acts”

Planning and executing an *action*, such as grasping and eating an apple, implies to have an *overarching goal* (to eat the apple), to select the appropriate sequence of *motor acts*—each with its specific motor goal (reaching, grasping, bringing to the mouth, biting)—and to execute the sequence of *movements* forming each motor act (see Rizzolatti et al., 2014). Attaining action goals relies on the precise integration of the processes carried out at each of these hierarchical levels and on their accurate timing. It is well known that area F1 and F5 are mainly involved in movement implementation, and in coding the goal of motor acts, respectively. Area F5 neurons typically encode motor acts performed with the hand or the mouth (Kurata and Tanji, 1986; Gentilucci et al., 1988; Rizzolatti et al., 1988; Hepp-Reymond et al., 1994; Ferrari et al., 2003). Electrophysiological studies revealed that a large proportion of F5 neurons encode specific motor acts such as grasping or tearing, rather than simple movements (Rizzolatti et al., 1988). Typically, an F5 hand motor neuron discharge during finger movements aimed at taking possession of an object (grasping) but not during similar movements aimed at different goals (e.g., scratching). In addition, F5 neurons activate when the same goal is achieved by using different

effectors/movements (e.g., taking possession either with the right hand, the left hand, or the mouth, **Figure 3A**). Interestingly, many neurons code specific grip types such as precision grip, finger prehension or whole hand prehension. Concerning the timing of grasping, some neurons discharge during the whole unfolding of the motor act, and others during a specific part of it (e.g., shaping of the hand).

Altogether, these data led to the proposal that F5 contains a “vocabulary” of motor acts (Rizzolatti et al., 1988). The “words” of this motor vocabulary are represented by different populations of neurons, some coding the general goal of a motor act, others coding how a specific motor act has to be executed or specifying the temporal aspects of the motor act to be executed (see Jeannerod et al., 1995). Neuroanatomical data show that F5 is densely connected with the parietal areas AIP, PF, PFG, and SII (Petrides and Pandya, 1984; Matelli et al., 1986; Cavada and Goldman-Rakic, 1989; Luppino et al., 1999; Rozzi et al., 2006; Borra et al., 2008; Gerbella et al., 2011). Areas F5 and PFG also share numerous functional properties (Leinonen et al., 1979; Hyvärinen, 1981; Rozzi et al., 2008; Bonini et al., 2010), and both contain motor neurons coding goal directed motor acts. A definitive demonstration that motor neurons indeed code motor acts has been provided by a study in which the same motor goal was achieved by employing opposite movements (Umiltá et al., 2008). Monkeys were trained to grasp objects using “normal” pliers, that is pliers that require hand closure in order to take possession of the object, and “reverse” pliers that require hand opening to achieve the same goal. The correlation between the neuron discharge and the hand movements revealed that a population of F5 neurons code goal achievement (i.e., taking possession of the target object) independently of the type of fingers movement employed (flexion or extension, **Figure 3B**).

Coding Motor Intention

Based on the data described in the previous section, a dissociation seems to exist between goal and movement in the motor system. One can therefore hypothesize that some population of neurons would code an even higher level of goal representation, possibly reflecting the overarching goal of the action, and expect to find neurons discharging differently during the execution of a motor act (e.g., grasping) according to the overarching goal of the whole action (e.g., eating). Recently, a series of experiments were carried out to test this hypothesis (Fogassi et al., 2005; Bonini et al., 2010, 2012). Grasping neurons were recorded from areas PFG and F5 in two conditions: in the first condition, the monkey grasped a piece of food and brought it to the mouth for eating; in the second, the monkey grasped an object or a piece of food in order to place it into a container. Some neurons discharged stronger during grasping to eat, and weaker or did not discharge at all during grasping to place in the container. Others had an opposite behavior (**Figure 3C**). Note that the differential discharge occurred during the actual grasping execution, and that the grasping act itself—consisting in closing the hand on the object—was exactly the same in the two conditions. The kinematics of reaching movements, the grip force exerted, the type of object involved—metallic cube or food—or the amount of underlying motivation could not account

for the differential activation of the neurons in the two conditions (Fogassi et al., 2005; Bonini et al., 2010). The discharge of these motor neurons, besides coding goals at the motor acts level, also reflects the overarching goal of the actions. Such neurons could play an important role in linking the specific motor acts belonging to an action in an appropriate motor chain, allowing the correct and fluid execution of the corresponding movement sequence. Beside this role in kinematic fluidity, their activation could have significant implications at a cognitive level. The firing of these neurons, together with that of the other neurons involved in the same action, represent the neural correlate of the overarching goal underlying the action, that is, the *motor intention* of the acting individual. Having a motor system wired as such could have been important in the phylogenetic development of the ability to read other's intentions. One of the mechanisms possibly underlying this capacity relies on the mirror system and is discussed in the following sections.

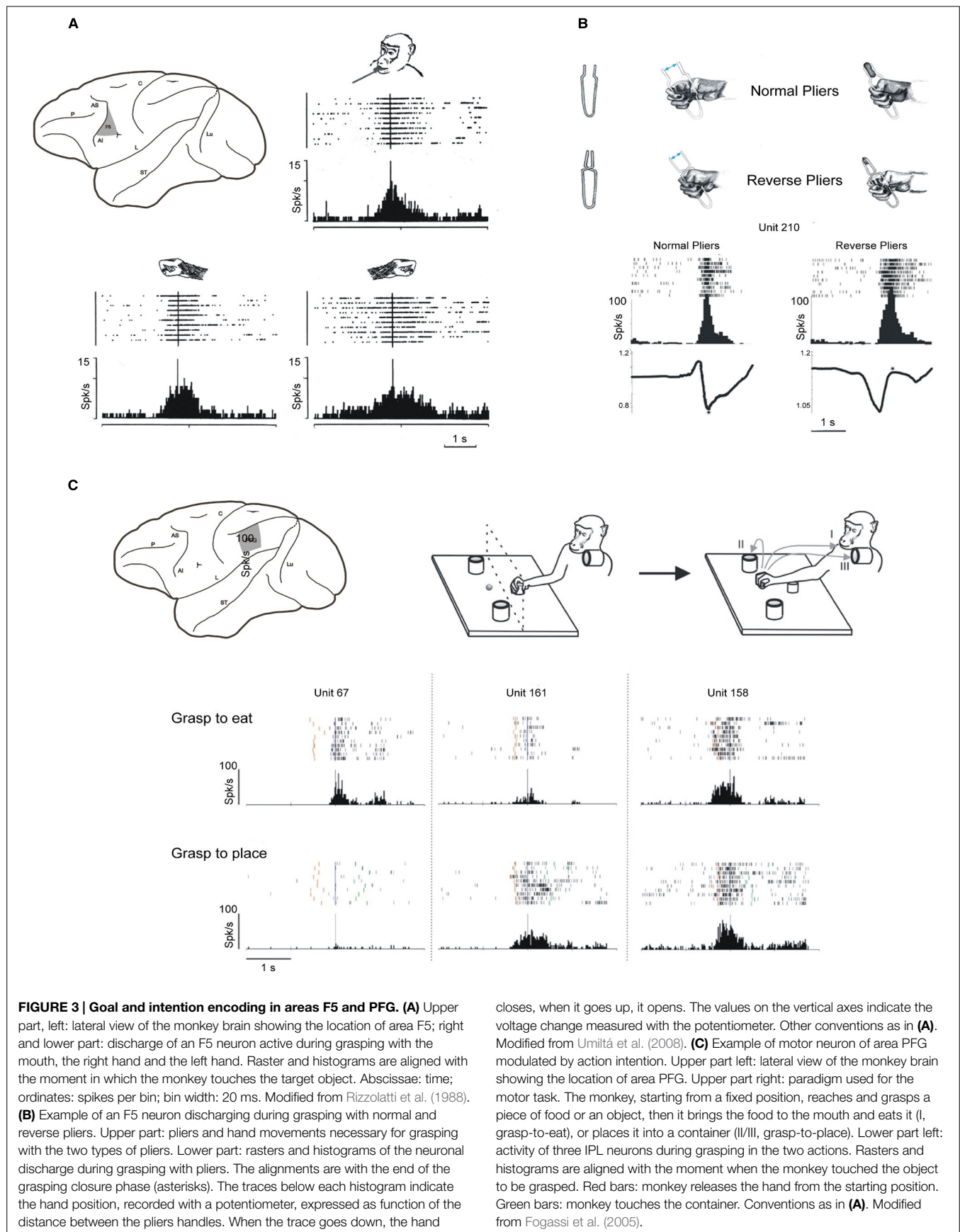
The Mirror Mechanism

The discovery of the *mirror mechanism* radically changed our views on the functional role of the motor system. It is now largely accepted that the same neurons involved in motor coding can also underpin social abilities such as understanding actions, reading others' intentions and programming contextually appropriate motor responses (see Rizzolatti et al., 2014). The existence of mirror neurons also shed light on how some basic processes involved in social cognition can be mediated by the motor system. Altogether, these functions represent fundamental aspects for social relations in primate and human societies (Sebanz and Knoblich, 2009; Bach et al., 2011; see Sebanz and Knoblich, 2008).

Simply put, mirror neurons discharge when a subject either actually performs a motor act or simply observes the same act being performed by someone else (**Figures 4A,B**). In other words, the observation of an action triggers in the observer's brain a representation of that action in a motor format (Di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996a; see Rizzolatti and Sinigaglia, 2010; Rizzolatti et al., 2014). The fact that an action representation in motor format can be activated by mere observation raise an important question: why don't we automatically move when observing an action? Recent single neuron experiment showed that the activity of a significant portion of pyramidal tract neurons of area F5 is modulated by action observation (Kraskov et al., 2009), and either increase or decrease their discharge. This finding indicates that mirror neurons activity can be transmitted to the spinal cord. Considering that more than one-fourth of pyramidal tract neurons show suppression of discharge during observation, while increase firing rate during active movement, the authors suggested that this inhibitory effect might play a role in preventing movement generation during action observation.

Mirror Neurons and Action Understanding

It has been proposed that the mirror mechanism, by matching the visual description of a motor act with its motor representation



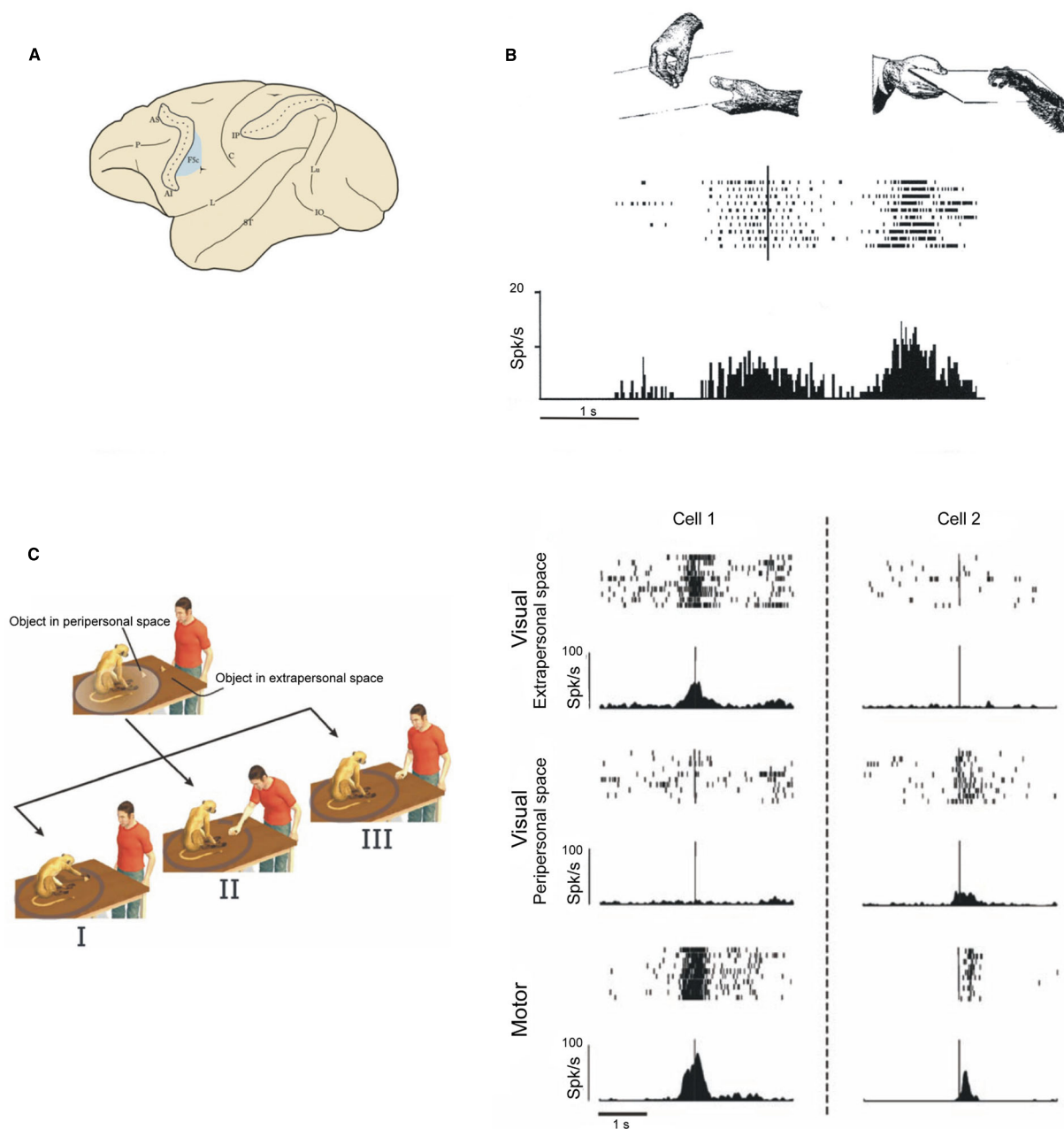


FIGURE 4 | Examples of F5 mirror neurons. (A) Lateral view of the monkey brain showing the location of area F5c. **(B)** Mirror neuron responding during observation and execution of a hand grasping motor act. Conventions as in **Figure 3A**. Modified from Rizzolatti et al. (1996a). **(C)** Examples of mirror neurons whose visual response is modulated by the distance at which the observed act is performed. Left: experimental paradigm. I the monkey grasps a piece of food. II

and III the experimenter grasps a piece of food located in the monkey peripersonal or extrapersonal space, respectively. Right: Cell 1: the visual response is higher during observation of grasping performed in the extrapersonal space; Cell 2: the visual response is higher during observation of grasping performed in the peripersonal space. Each panel shows a raster plot and the spike density function of the neuron response. Modified from Caggiano et al. (2009).

in terms of goal, allows the observer to understand what another individual is doing. Such a process would be possible because the observation of an act automatically retrieve its motor representation by tapping into the observer's motor vocabulary (described in the previous section, see Rizzolatti and Sinigaglia, 2010; Rizzolatti and Fogassi, 2014; Rizzolatti et al.,

2014). This implies that a representation of the motor goal of an act can be triggered by sensory information. The nature of the sensory information capable of activating mirror neurons has been investigated in two neurophysiological studies. In the first, mirror neurons have been demonstrated to discharge both when the monkey can fully observe an experimenter grasping

an object, and when he can only see part of the action, due to its crucial part (the hand-object interaction) being hidden by a screen (Umiltà et al., 2001). Interestingly, there was no neuronal discharge if the monkey knew that there was no object to grasp behind the screen, suggesting that, in the absence of a full visual description, mirror neurons use mnemonic-contextual information to retrieve the motor representation of the observed motor act. In the second study, sensory information about the motor act was presented to the monkey in an acoustic and/or visual format. It revealed that some mirror neurons (audio-visual mirror neurons), discharge not only during the execution and the observation of a motor act producing a sound (e.g., the crackling sound of breaking a peanut), but were also activated when the monkey simply heard the sound made by the action (Kohler et al., 2002; Keysers et al., 2003). Altogether, these data indicate that mirror neurons respond to the goal of others' motor acts also in partial or total absence of visual cues.

Mirror Neurons and Intention Coding

Some mirror neurons share an interesting property with purely motor neurons and encode the motor intention behind the actions performed by other individuals (Fogassi et al., 2005; Bonini et al., 2010). A series of experiments was carried out to assess a possible relation between motor intention and mirror neurons activity. One way of testing this possibility was to verify whether neurons discharging during the execution and observation of grasping acts are influenced by the type of action in which the grasping acts are embedded (Fogassi et al., 2005; Bonini et al., 2010). In this purpose, grasping-related mirror neurons were recorded from parietal area PFG and premotor area F5 while the monkey executed a motor task (motor condition) and observed the same task, performed by an experimenter (visual condition). The experimental paradigm was the same as previously described in the section "coding motor" intention: an identical grasping act was embedded into two different actions, aimed at eating or placing the target in a container, respectively. The results show that in both the motor and visual condition a large proportion of mirror neurons discharged differently during the observation of the grasping act, when it was part of the two different actions (action-goal-related mirror neurons). The neuronal selectivity for the overarching goal expressed during grasping observation has been interpreted as a prediction of the action outcome. Making such a prediction is possible since the monkey knows that a given context like the type of target object or the presence of a container is followed by *grasping-to-place* action. Note, however that these neurons are not activated by contextual cues such as the observation of the target object or of the scene, but by action observation. Accordingly, it was hypothesized that when one of these neurons is activated by the observation of a grasping motor act that is part of a specific motor action (grasp-to-eat or grasp-to-place), it triggers the motor circuitry that constitutes the internal representation of the overarching goal of the sequential action. Thus, mirror neurons, besides the capacity of coding motor acts, provide individuals with a mechanism for understanding others' intentions (Fogassi et al., 2005; Rizzolatti and Sinigaglia, 2010; Rizzolatti et al., 2014).

Mirror Neurons and the Social Context: Space and Agency

Understanding the behavior of others' is one of the building block of social cognition. However, in social animals in which object-oriented behaviors usually occur in the presence of other individuals, understanding the action goal and the intention behind it is not sufficient to frame this action in its social context. In this purpose, it is also very relevant to evaluate actions with respect to the position in space where they occur, and especially with respect to the observer's position. For example, if an individual is grasping an object close to an observer, an interaction is possible. The observer has the actual possibility to interfere with the grasping action and prevent it to happen, or cooperate to it. Cooperative behaviors are common in humans, but are also documented in monkeys (Mendres and de Waal, 2000; Visalberghi et al., 2000; Visco-Comandini et al., 2015). If the mirror mechanism was only involved in action and intention understanding, the spatial location of the observed action and the vantage point of the observer would be irrelevant. However, if space also plays a role in tuning motor responses appropriate to others' actions, as first proposed by Jeannerod (2006), these spatial factors could possibly modulate the neural discharge of mirror neurons. This hypothesis has been empirically tested in different experiments. The results showed that, although in most cases the visual response of mirror neurons is invariant with respect to spatial features, the discharge of some of them is modulated by the direction of the hand movement, the space sector (right or left) in which the motor act occurs or the hand (right or left) used by the observed agent (Gallese et al., 1996; Rozzi et al., 2008).

The effect of the distance at which an action occurs on the discharge of mirror neurons was systematically tested in a recent experiment (Caggiano et al., 2009). In this study the same motor act was executed within the monkey reaching space (peripersonal space) or outside it (extrapersonal space). About half of the studied mirror neurons discharged differently in the two conditions. Of them, 50% discharged stronger when the monkey observed the experimenter grasping a piece of food in its peripersonal space and 50% in the extrapersonal space (**Figure 4C**). Crucially, the authors tested whether, in these mirror neurons, space was represented in terms of a metric representation—the geometric distance between the action and the monkey—or in terms of operational representation—the pragmatic space where the monkey can actually act. To this end, a transparent barrier was introduced between the monkey and the site where the experimenter executed the action. In this condition the monkey could see the action, but was prevented from interacting with the object located within its peripersonal space. If a metric representation is at play in the mirror neurons code, peripersonal and extrapersonal space would remain unchanged, while if an operational representation occurs, the introduction of the barrier would lead to a remapping of the peripersonal into extrapersonal space. The results show that when the barrier was introduced, *extrapersonal* mirror neurons started discharging also when the observed action was performed within the peripersonal space, as if this latter were displaced far away. Taken together, these data suggest that a subpopulation of mirror neurons can code differently others' actions depending on the space sector in which they occur. It is very likely that space

location and distance are coded within the mirror neuron system in relation to the often vital possibility to interact or not with others. Thus, mirror neurons, besides being involved in action understanding, could also be important for choosing the motor response appropriate to others' actions in their specific behavioral context.

The issue of space coding is very important also because a large number of primates actions is directed toward oneself (e.g., bringing objects to the mouth), while most of the studies about the mirror system focused on actions directed away from one's body (e.g., reach for and grasp an object). It is well known that in the fundus of the intraparietal sulcus (area VIP) there are bimodal neurons, responding to visual and tactile stimuli, whose tactile receptive fields are located predominantly on the face and the visual receptive fields are in spatial register with the tactile ones (Colby et al., 1993; Duhamel et al., 1998). The electrical stimulation of this area evokes face movements and defensive movements of the arm toward the face (Cooke et al., 2003). Ishida et al. (2010) studied the neural properties of a population of these bimodal neurons, delimiting the extension in depth of their peripersonal space in monkeys either alone, or facing an experimenter. Typically, when a visual stimulus was presented outside the peripersonal space, at more than one meter of distance from the tactile receptive field, no visual response was recorded. However, when an experimenter was standing in front of the monkey at the same distance and a stimulus was moved close to his/her body part corresponding to the neuron tactile receptive field, the response appeared. In other words, other's body space was matched to the monkey's one. This result indicates a possible way for encoding others' peripersonal space, and might extend the role of the mirror mechanism in action understanding to others individuals' actions aimed at themselves. However, in this study the motor responses of the neurons have not been recorded, and it is impossible to tell whether these neurons actually were mirror neurons, nonetheless, it is known that area VIP is strictly connected with premotor area F4 (Matelli et al., 1986; Barbas and Pandya, 1987; Cavada and Goldman-Rakic, 1989; Andersen et al., 1990; Lewis and Van Essen, 2000), where peripersonal space is encoded in terms of reaching movements (Gentilucci et al., 1988; Fogassi et al., 1996; see Graziano, 2006). It is plausible, therefore, that the visual responses actually represent potential motor acts directed toward specific body parts (Gentilucci et al., 1988; Fogassi et al., 1996).

A subsequent study investigated a further important aspect of the observed actions, that is the view-dependence of the visual responses of mirror neurons (Caggiano et al., 2011). To this purpose the monkey was required to observe movies showing the same grasping motor act from three different points of view: in the subjective perspective (0°), and in two types of third-person views, a lateral (90°) and a frontal (180°) one. Among the tested mirror neurons, about three-fourth showed a preference for one of the vantage points, encoding in equal percentage the three different perspectives employed. On the base of these results it has been proposed that *view-independent* mirror neurons encode the goal of the observed motor act irrespective of the visual details of the scene, while *view-dependent* mirror neurons provide a link between the goal of the motor act and its pictorial aspects.

Similarly to the mirror neurons modulated by peripersonal or extra-personal space, *view-dependent* mirror neurons could be important for preparing an adequate response to the observed action. These neurons could be part of a neural circuit of the "social brain" coding the spatial relations at the roots of basic social interactions.

In all the reported studies on mirror neurons, actions were unidirectional and non-interactive, while in nature, most often, monkeys interact within complex social environments in which different individuals share the same social space. Here we refer to social interactions as the acts of two or more individual taking into account of other's actual or potential actions or intentions. By combining a motion capture system with chronic multielectrode recording from different cortical areas (multi-dimensional recording), Fujii et al. (2007, 2008) were able to study the neural activity from monkeys' parietal and premotor cortex in a social context. When two monkeys were sitting one close to the other, and could reach for and grasp food without interacting, parietal activity resulted to be strongly tuned to the use of the arm contralateral to the recorded hemisphere. However, when the food was put in a shared space and a social conflict emerged between the monkeys, the neurons developed different combinations of preferences to self and other motion (Fujii et al., 2007). This evidence indicates that parietal neurons can recognize social cues and provide other areas with a neural code modulated by social information. The same authors also described the responses of premotor and parietal neurons during the observation of action in a task in which two monkeys were present, but could not interact (Fujii et al., 2008). During action execution both premotor and parietal neurons showed a strong preference for actions performed with the arm contralateral to the recording hemisphere (right arm). During the observation of the other monkey action, the premotor neurons preferred the other monkey right arm movements, while the parietal neurons typically lost this laterality preference, showing a wider spectrum of combinatorial responses to own/other right/left responses. Indeed, the arm used (right or left) in a specific context (position of the food on the right or left side) is very relevant to understand the intention of an action. Accordingly, the authors propose that the premotor neurons code information on action's agent and effector as primitives of action recognition within the mirror network, while parietal neurons represent the social space and participate in recognizing others' actions with respect to one's own actions (Fujii et al., 2008).

The mirror mechanism can enable to understand others' action, but in this process, the sense of agency appears to be revoked: the neurons are active both when I act and when I see someone else acting, without moving. The lack of synchronicity between the vision of an action and the somato-motor signals related to action execution probably represents the crucial information for attributing the action to self or others (see Wolpert and Ghahramani, 2000; Gazzola and Keysers, 2009; Pitti et al., 2009). Interestingly, different studies described neurons activated by the observation of actions but not discharging during action execution, in the premotor, parietal and temporal cortex (Perrett et al., 1989; Gallese et al., 1996; Fujii et al., 2008). It was proposed that these neurons, by separating visual and somato-motor information, could play a role in the attribution of agency

to others (Fujii et al., 2008). The same authors also propose that this function does not rely on single cortical areas, but on a larger cortical network capable of integrating visual and somato-motor informations. Possible networks involved in action observation and participating to this function are described in the following section.

Anatomo-Functional Mirror Pathways

Among the cortical areas involved in the processing of visual information, those located in the superior temporal sulcus (STS) are generally considered as the fundamental node coding biological information. In fact, STS contains neurons coding visual information about eye/gaze direction, body/limbs orientation and movement, facial expressions, and biological motion (Bruce et al., 1981; Perrett et al., 1989; Puce et al., 1998; Pelphrey et al., 2003; Tsao and Livingstone, 2008). These features are among the most relevant aspects needed by an individual to interpret others' behavior, and, for this reason, STS is generally considered as the initial stage in processing social cues. To date, very few studies have been done to elucidate how the mirror system interplays with STS or with other parts the "social brain." Notwithstanding the paucity of available data on this topic, a few studies are unveiling possible pathways by which social information can reach mirror neurons' computation.

A recent study employed fMRI technique in the monkey to identify the frontal areas active during the observation of motor acts (Nelissen et al., 2005). By using anatomically defined regions of interest, the authors found that viewing videos showing a hand grasping an object activates F5a, F5p, and the prefrontal areas 45A, 45B, and 46. When the video showed the whole individual grasping an object, and not only a hand, the activation also involved F5c. This indicates that there are multiple representations of others' actions in the monkey frontal cortex and that they can be sensitive to different features. F5a, F5p and the prefrontal areas appear to encode the action as such, while F5c action representation is more centered on the agent doing the action. This result has two important implications: first, a context-dependent processing of the act of grasping is taking place in F5c; second, an input coming from areas processing the visual features of the scene (like STS), makes its way to the premotor cortex.

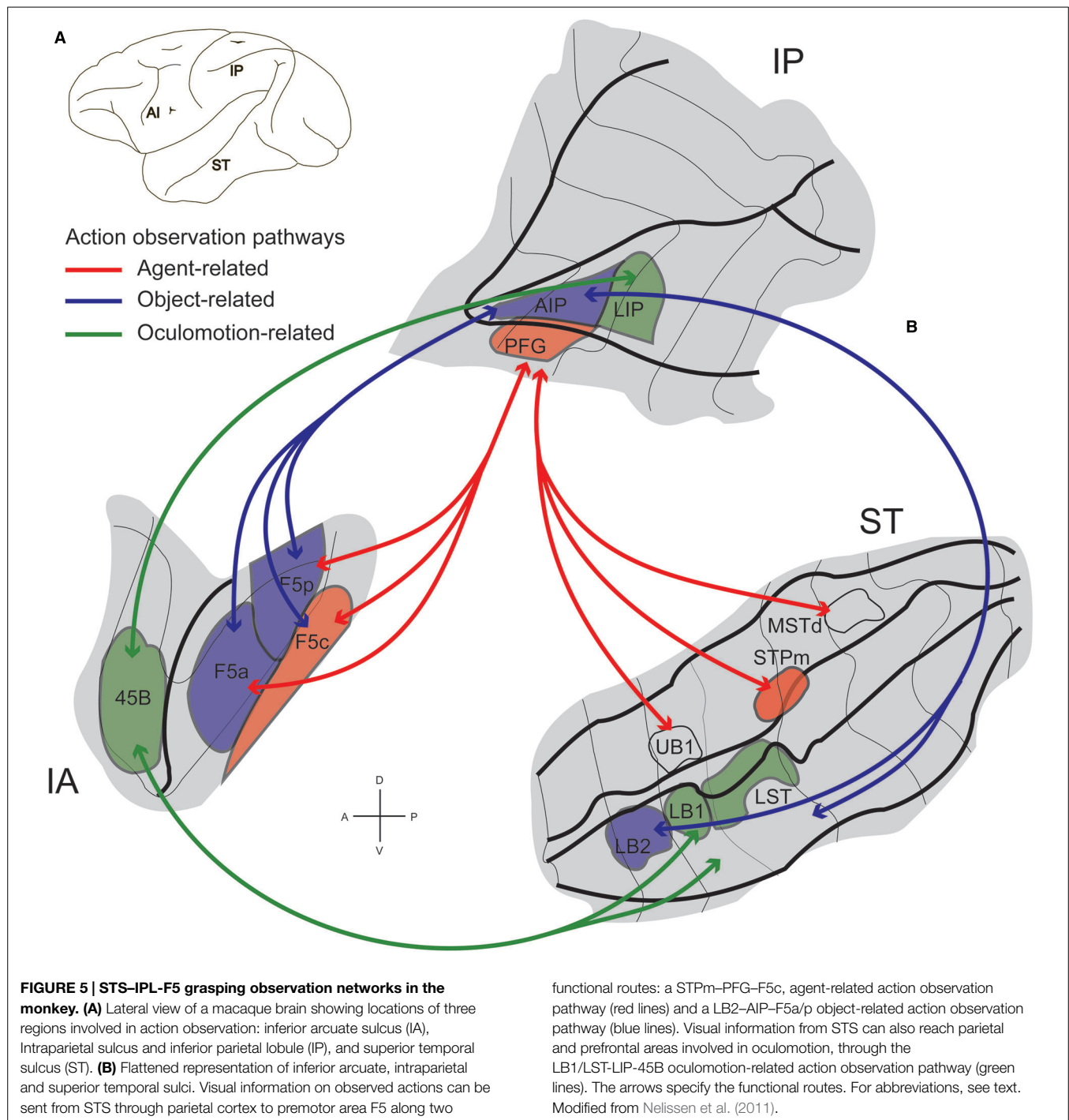
A subsequent anatomo-functional study made by Nelissen et al. (2011) investigated how visual information about action can reach the frontal areas, and concentrated on STS and posterior parietal cortex fMRI activation in the monkeys during the observation of grasping acts. They also correlated functional with connectional data obtained by means of neural tracers injections. The employed videos activated areas in the lower and upper banks of STS, and in the IPL. An analysis based on regions of interest showed that grasping observation activates stronger than control conditions three IPL areas (PFG, on the cortical convexity, AIP and LIP in the lower bank of intraparietal sulcus) and five STS regions (MT/V5, LST, and LB2 in the lower bank, FST in the fundus and STPm in the upper bank). Note that a recent electrophysiological study directly demonstrated the presence of mirror neurons in the AIP area (Pani et al., 2014).

In order to assess which of the STS areas active during action observation are actually connected with the mirror areas, retrograde tracers were injected in the parietal nodes of the mirror system (AIP and PFG). After AIP injections, a widespread STS labeling were found, but the most consistent labeling in all cases was in the lower bank sector LB2 and in the inferotemporal cortex near the lip of STS. Injections in PFG resulted in consistent labeling in STS upper bank sectors MSTd, STPm, and UB1. Note that, of them, only STPm was found to be specifically active during action observation.

This integrated anatomo-functional approach led to the identification of two functional pathways involved in action observation linking STS, IPL, and PMv (**Figure 5**, red and blue). One links STS sector STPm with parietal area PFG that, in turn, is connected with premotor area F5c. The other pathway connects LB2 with AIP that, in turn, is connected with F5a and F5p. Both routes process information necessary for understanding the observed motor act, but each provides a different type of information, possibly playing a specific role in understanding the intention underlying it. In particular, the STPm-PFG-F5c pathway is more concerned with the *agent doing the action*, while the LB2-AIP-F5a/p one with the *details of grip and object identity* (Nelissen et al., 2011).

These pathways show that the parietal regions containing mirror neurons have a direct access to STS information about biological motion, crucial for coding the observed agent's actions and intentions. This direct access implies that PFG/AIP would be a first node where the neural codes for grasping and for social information—like gaze, head, body, or limb orientation or direction—are integrated in a common motor representation that becomes available for mirroring others' actions. This integrated code in which grasping is linked with social information would then be sent to F5 in the premotor cortex.

A further pathway links area 45B in the prearcuate cortex with LST and LB1 in the lower bank of STS, and LIPa in the lower bank of intraparietal sulcus (**Figure 5**, green). Note that monkey area 45B is known to be part of the oculomotor system, probably representing the gateway of highly integrated prefrontal and orbitofrontal information to this system (Moschovakis, 2004; Gerbella et al., 2010). It was proposed that the LST/LB1-LIP-45B pathway could play a role in oculomotor control during action observation (Gerbella et al., 2010; Nelissen et al., 2011). Indeed, gaze behavior mirroring has been found in LIP of macaque monkeys where a sub-population of neurons discharge both when monkeys direct their gaze in a given direction and when they look at a static image of another monkey having the gaze oriented in that same direction (Shepherd et al., 2009). The areas of the LST/LB1-LIP-45B pathway are activated by action observation (Nelissen et al., 2011), but there is no evidence of the presence of grasping mirror neurons in any of them. So the question remains: where does the integration of information related to grasping and gaze direction occurs? This question is even more relevant considering the importance of parsing others' gaze direction for deciphering their intention (see Klein et al., 2009). The existence of this pathway raise the possibility that STS information about biologically or socially relevant gaze targets reaches oculomotor areas LIP and 45B.



However, “oculomotor mirroring,” confirmed in LIP, remains untested in prefrontal cortex. One could reasonably expect to find a population of neurons mirroring gaze behavior in area 45B. Note that in this pathway, gaze information would still be segregated from the one coding grasping. The anatomical pathways through which gaze mirroring would reach the parietal and premotor areas of grasping mirroring still remain to be described, but probably include the prefrontal cortex (see below).

Summing up, the STS information about social cues deriving from biological motion analysis could reach the mirror system directly (STPm-PFG-F5c and LB2-AIP-F5a/p pathways) or indirectly through an “oculomotor” mirroring system (LST/LB1-LIP-45B pathway). These hypotheses are not mutually exclusive, and thus far, there is no data directly confirming any of them. However, there are indirect behavioral and electrophysiological evidence suggesting that information about gaze direction and action observation converge and probably become integrated in

the same neural code. The behavioral data comes, from a well-known human study showing that subjects display the same gaze pattern when performing a grasping action and when observing another individual performing the same action (Flanagan and Johansson, 2003). The electrophysiological evidence comes from preliminary data showing that in monkeys area F5, the activity of some mirror neurons is modulated by the gaze direction of the observed agent (Coudé et al., 2013).

Prefrontal Cortex and Mirror Network

Functional MRI studies demonstrated that prefrontal area 46 is involved in action observation (Nelissen et al., 2005) and motor-related activity in the ventral prefrontal cortex has been described (Tanila et al., 1992; Hoshi et al., 1998; Rozzi et al., 2011). More recently, connectional studies on the ventral prefrontal cortex indicated that a specific sector of ventral area 46 (rostral part of 46VC) and area 12 (intermediate 12r) is connected with different nodes of the mirror pathways (Borra et al., 2011; Gerbella et al., 2013). These nodes include rostral premotor area F5a, IPL areas PFG and AIP and a sector of the ventral bank of rostral STS sector, which overlaps with the fMRI sites activated by action observation. Altogether, these evidence indicate that certain parts of the prefrontal cortex might be considered as actual components of the mirror system, but electrophysiological confirmation of this hypothesis is still lacking. A possible role of the prefrontal cortex within the mirror system, could be to provide the motor representations of the parietal and motor areas with mnemonic and contextual information. The mirror system access to this kind of information could, for instance, allow action understanding when the target object is not actually visible during the unfolding of a grasping action (Umiltá et al., 2001). It could also enable intention understanding by retrieving the meaning of contextual cues previously associated to specific actions (Fogassi et al., 2005). In addition, the ventral prefrontal cortex could provide the parietal and premotor cortex with social contextual cues. Such cues would consist in information about gaze direction or body part orientation, as elaborated in STS. Interestingly another sector of ventral prefrontal cortex (caudal part of 46VC) is strongly connected with frontal and parietal oculomotor areas, as well as with the STS and the other sectors of area 46. This pattern of connections could represent a pathway, though indirect, linking the oculomotor system with the mirror system.

Note that the connections between prefrontal areas and mirror areas are bidirectional. This implies that, from the one side, the VLPF can modulate mirror neuron activity by sending mnemonic and contextual information, from the other, the parieto-premotor areas could provide the prefrontal cortex with motor representations of action goals. Thus, a further role of the prefrontal cortex could consist in recombining the observed motor acts, captured by the parietal and premotor nodes of the mirror system, to produce an action fitting the observed model, allowing imitative learning, as suggested by studies on humans (Buccino et al., 2004, see below). Further studies will have to verify these hypotheses and assess the specific contribution of the prefrontal areas, classically considered to exert a top-down control on sensory and motor areas, to the mirror system.

The Mirror System in Humans, An Anatomo-Functional Perspective

The mirror system is thought to constitute a fundamental part of the vertebrate motor system and has presumably been conserved and adapted through different species, including humans. The previous sections outlined its circuitry and functions in the monkey. Technical and ethical limitations precludes to reach a similar level of details in the description of the human mirror system. However, studies using non-invasive techniques like brain imaging, TMS and EEG/MEG have yield evidence that a mirror system exists in humans (Fadiga et al., 1995; Grafton et al., 1996; Rizzolatti et al., 1996b; Pfurtscheller and Neuper, 1997; Hari et al., 1998; Cochin et al., 1999; Grèzes et al., 1999; Nishitani and Hari, 2000, 2002; Buccino et al., 2001; Gangitano et al., 2001; Perani et al., 2001; Maeda et al., 2002; Muthukumaraswamy and Johnson, 2004; Iacoboni et al., 2005; Oberman et al., 2005; see Pineda, 2005; Rizzolatti et al., 2014).

In EEG studies, Mu waves are detected in the 8–13 Hz frequency range and are thought to be the result of synchronous discharges by resting neurons in the sensorimotor region of the brain (see Kuhlman, 1978; Anderson and Ding, 2011). Mu rhythm suppression occurs during motor preparation, action execution (Neuper et al., 2006), but also during mental imagery and action observation (Cochin et al., 1999; see Pineda, 2005). Brain imaging studies demonstrated a consistent pattern of cortical activity during action observation, involving a network of several brain regions (see Caspers et al., 2010). This action observation network includes Brodmann's areas 44/45, lateral dorsal premotor cortex, supplementary motor area, primary somatosensory cortex, superior parietal lobule, intraparietal cortex, rostral inferior parietal lobule, posterior middle temporal gyrus at the transition to visual area V5, and fusiform face area/fusiform body area.

Interestingly, this human mirror network largely overlaps with the monkey one (IPL, PMv, and caudal part of inferior frontal gyrus). However, various other areas are active in humans during action observation. The only one description of a single neuron mirroring mechanism was provided by Mukamel et al. (2010), recording from areas not belonging to the classical mirror system. The larger number of areas involved in action observation in humans could depend on several factors. First, most of monkey studies have been carried out by means of single neuron recording. This technique is the only one capable of demonstrating the presence of mirror neurons, but lacks the possibility to explore large brain regions at the same time. Thus, it is likely that the monkey mirror system has not yet been fully mapped. This hypothesis is supported by ¹⁴C-deoxyglucose autoradiography experiments in monkeys showing that further regions beyond the classical mirror areas—including superior parietal, somatosensory and primary motor areas—are activated by action observation (Evangelidou et al., 2009; Raos et al., 2014), although the actual presence of mirror neurons in these areas is still to be confirmed. A second hypothesis is that the mirror system in humans have expanded to additional cortical areas, probably acquiring new functions. A third possibility is that the brain activation evidenced by brain imaging studies

during action observation could be related to different aspects of visual processing or to motor preparation, and be independent on the actual presence of mirror neurons. To our knowledge, none of these hypotheses has been ultimately demonstrated. An interesting attempt has been done by Gazzola and Keysers (2009) in a recent fMRI study aimed at identifying the brain regions activated by both action observation and action execution, and thus, likely containing mirror neurons. The single-subject analysis of unsmoothed fMRI data allowed the authors to identify the voxels shared between action observation and action execution in the classical IPL-PMv circuit, but also in the middle cingulate, dorsal premotor, somatosensory, superior parietal, and middle temporal cortex. The activation of areas not belonging to the classically described parieto-premotor mirror circuit could reflect sensory predictions from internal models (Wolpert and Ghahramani, 2000; Gazzola and Keysers, 2009). This process would complete and enrich the information about others' actions encoded by the classical mirror system. Further studies on action observation and execution conducted in human and monkey by means of brain imaging and electrophysiological techniques will be important to demonstrate the presence of mirror neurons in the human cortex and to test and disentangle between the different hypotheses proposed above.

The mirror mechanism, besides being involved in action understanding, could also play a role in learning by imitation. Buccino et al. (2004) specifically investigated this issue by means of fMRI. In this study, naive participants were required to observe images depicting the hand of an expert guitarist playing chords and to imitate them after a delay. Action observation, as expected, activated IPL, PMv and the pars opercularis of the inferior frontal gyrus. Noteworthy, these areas together with the prefrontal cortex (area 46 in the middle frontal gyrus) and the anterior mesial cortex were active during the delay phase preceding movement execution. The authors proposed that area 46 could recombine the observed motor acts, captured by the parieto-premotor mirror system in order to produce an act fitting the observed model.

Others' Actions in Their Social Context

As mentioned above, the mirror mechanism in monkeys is considered to be involved in coding others' actions in their social context (Fujii et al., 2007, 2008; Ishida et al., 2010; Visco-Comandini et al., 2015). The term "social context" encompasses a wide spectrum of settings and can refer to complex interactions, especially in human societies. Whereas some forms of human social interactions appear to be unique in their complexity, other forms are more basic and are probably shared with other primates (see Tomasello and Call, 1997). It is thus likely that the same mirror mechanism is involved in the most basic social interactions in different primate species. Human brain imaging and TMS data seems to support this idea. Indeed, it has been showed that areas pertaining to the mirror system are more strongly activated when subjects performed complementary actions rather than when they performed the same action as the one observed (Newman-Norlund et al., 2007). TMS data by Sartori et al. (2011) also point in the same direction and demonstrate that depending on the context, motor-evoked potentials can reflect the observed movement or an appropriate complementary movement. In this

experiment, when an object was present and the observer was implicitly required to act upon the object in response to the observed action, a shift from symmetrical motor resonance to complementary activations of hand muscles was observed. Thus, action observation does not inevitably lead to symmetrical motor facilitation, that could be useful for imitation, but could also play a role in successfully performing attuned joint actions.

Human data also showed that intentions and social contexts affect kinematics, and conversely kinematics and contexts affect intention understanding. The kinematics of a grasping act differ depending on the final goal of the action (e.g., grasping to move, to throw or to pass, see Becchio et al., 2012). On the other hand, the context provides clues for understanding the intention underlying the observed motor act, and is known to modulate the activity of the caudal sector of IFG, during action observation (Iacoboni et al., 2005). This mean that, also in humans the mirror system is involved in intention coding. In addition, it has been showed that reaching toward an object and grasping it either to move it from one spatial location to another or to place it into the hand of a partner yield different kinematics (Becchio et al., 2010; for similar results see also Mason and Mackenzie, 2005; Meulenbroek et al., 2007). Interestingly, the observation of social movements evokes an activation stronger than non-social ones within mirror areas, including the IFG and the IPL (Becchio et al., 2012). These finding demonstrates that areas within the mirror system are sensitive to very subtle differences in the observed action's kinematics. Most importantly, it suggests that mirror areas in humans are more responsive to social than non-social movements. Similarly, Mu rhythm suppression has been shown to be greater for social interactive actions than for non-social actions (Oberman et al., 2007).

Altogether, these data suggest that during social interaction, human agents decipher the goal of others' ongoing action and integrate it into their own action planning, eliciting different potential complementary responses. Thus, the mirror mechanism, being tuned to social actions, besides its known role in motor cognition, is likely involved in social cognition.

Conclusion

This review was an attempt to outline an updated view of the organization of the neural bases of grasping. Our knowledge of the motor system hinges on a multidisciplinary approach applied to the macaque monkey model. Obvious technical and ethical limitations preclude the application of such method to humans. However, clear homologies have been established between the motor systems of the two species. The basic mechanisms underpinning grasping actions are very likely shared among primates and humans. Among these mechanisms is the neuronal coding of movements in terms of motor goals and the mirror mechanism, allowing to retrieve these goals during action observation. The latter is an in-built motor resonance mechanisms, deemed to be at the core of action understanding. We believe that such neural coding, pertaining to the motor system and originally evolved for guiding behavior, has later become a fundamental component on which social cognition was constructed. However, the possible role of other processes,

for instance involving mentalizing, should not be downplayed and could work in parallel with the mirror system.

The mirror system is not only reflecting what another individual is doing, but also integrates contextual aspects like spatial cue, gaze direction or kinematic parameters. We discuss how this process of internal simulation is at the bases of action and intention understanding in monkeys and humans. Human data also yield a further fundamental function of the mirror mechanism: allowing the preparation of appropriate complementary responses to the observed actions. This latter process could explain how two individuals become attuned to cooperate in a joint action. It also underlines the flexibility of the mirror system.

Complex functions cannot depend on a single brain region but are rather the results of several areas linked together by cortical connections, and forming functionally specialized networks. The grasping execution/observation system is no exception. Clearly, specific sets of temporal, parietal and motor areas contribute to

different aspects of the mirror system functions. This suggests that the mirror neuron network extends probably beyond the motor system to include other cortical sectors. A deeper investigation of the role of these putative nodes of the mirror system, and especially of those located in the prefrontal cortex, will be crucial for defining the relationships between the classical mirror circuit and other centers possibly exerting a top-down control on them. This, in turn, will prompt a better understanding of how information about the social context can influence our comprehension of actions and intentions, and shape our own motor programs.

Acknowledgments

The research was supported by IAP 2011 (contract P7/11); Division of Intramural Research, NICHD and NIH P01HD064653; PRIN (prot. 2010MEFN7), European Commission Grant Cogsystems FP7-250013.

References

- Alstermark, B., and Isa, T. (2012). Circuits for skilled reaching and grasping. *Annu. Rev. Neurosci.* 35, 559–578. doi: 10.1146/annurev-neuro-062111-150527
- Andersen, R. A., Asanuma, C., Essick, G., and Siegel, R. M. (1990). Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J. Comp. Neurol.* 296, 65–113. doi: 10.1002/cne.902960106
- Anderson, K. L., and Ding, M. (2011). Attentional modulation of the somatosensory mu rhythm. *Neuroscience* 180, 165–180. doi: 10.1016/j.neuroscience.2011.02.004
- Bach, P., Bayliss, A. P., and Tipper, S. P. (2011). The predictive mirror: interactions of mirror and affordance processes during action observation. *Psychon. Bull. Rev.* 18, 171–176. doi: 10.3758/s13423-010-0029-x
- Barbas, H., and Pandya, D. N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *J. Comp. Neurol.* 256, 211–228. doi: 10.1002/cne.902560203
- Barbas, H. (1988). Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. *J. Comp. Neurol.* 276, 313–342. doi: 10.1002/cne.902760302
- Baumann, M. A., Fluet, M.-C., and Scherberger, H. (2009). Context-specific grasp movement representation in the macaque anterior intraparietal area. *J. Neurosci.* 29, 6436–6448. doi: 10.1523/JNEUROSCI.5479-08.2009
- Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., and Castiello, U. (2012). Social grasping: from mirroring to mentalizing. *Neuroimage* 61, 240–248. doi: 10.1016/j.neuroimage.2012.03.013
- Becchio, C., Sartori, L., and Castiello, U. (2010). Toward you: the social side of actions. *Curr. Dir. Psychol. Sci.* 19, 183–188. doi: 10.1177/0963721410370131
- Belmalih, A., Borra, E., Contini, M., Gerbella, M., Rozzi, S., and Luppino, G. (2007). A multiarchitectonic approach for the definition of functionally distinct areas and domains in the monkey frontal lobe. *J. Anat.* 211, 199–211. doi: 10.1111/j.1469-7580.2007.00775.x
- Belmalih, A., Borra, E., Contini, M., Gerbella, M., Rozzi, S., and Luppino, G. (2009). Multimodal architectonic subdivision of the rostral part (area F5) of the macaque ventral premotor cortex. *J. Comp. Neurol.* 512, 183–217. doi: 10.1002/cne.21892
- Bonini, L., Maranesi, M., Livi, A., Fogassi, L., and Rizzolatti, G. (2014). Space-dependent representation of objects and other's action in monkey ventral premotor grasping neurons. *J. Neurosci.* 34, 4108–4119. doi: 10.1523/JNEUROSCI.4187-13.2014
- Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F., and Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cereb. Cortex* 20, 1372–1385. doi: 10.1093/cercor/bhp200
- Bonini, L., Ugolotti Serventi, F., Bruni, S., Maranesi, M., Bimbi, M., Simone, L., et al. (2012). Selectivity for grip type and action goal in macaque inferior parietal and ventral premotor grasping neurons. *J. Neurophysiol.* 108, 1607–1619. doi: 10.1152/jn.01158.2011
- Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., et al. (2008). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb. cortex* 18, 1094–1111. doi: 10.1093/cercor/bhm146
- Borra, E., Belmalih, A., Gerbella, M., Rozzi, S., and Luppino, G. (2010). Projections of the hand field of the macaque ventral premotor area F5 to the brainstem and spinal cord. *J. Comp. Neurol.* 518, 2570–2591. doi: 10.1002/cne.22353
- Borra, E., Gerbella, M., Rozzi, S., and Luppino, G. (2011). Anatomical evidence for the involvement of the macaque ventrolateral prefrontal area 12r in controlling goal-directed actions. *J. Neurosci.* 31, 12351–12363. doi: 10.1523/JNEUROSCI.1745-11.2011
- Brodman, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde*. Leipzig: Barth.
- Bruce, C., Desimone, R., and Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* 46, 369–384.
- Buccino, G., Binkofski, F., and Fink, G. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404. doi: 10.1046/j.1460-9568.2001.01385.x
- Buccino, G., Vogt, S., Ritzl, A., Fink, G., and Zilles, K. (2004). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron* 42, 323–334. doi: 10.1016/S0896-6273(04)00181-3
- Bucy, P. C. (1933). Representation of ipsilateral extremities in the cerebral cortex. *Science* 78, 418. doi: 10.1126/science.78.2027.418
- Bucy, P. C. (1935). A comparative cytoarchitectonic study of the motor and premotor areas in the primate cortex. *J. Comp. Neurol.* 62, 293–331. doi: 10.1002/cne.900620203
- Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A., et al. (2011). View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Curr. Biol.* 21, 144–148. doi: 10.1016/j.cub.2010.12.022
- Caggiano, V., Fogassi, L., Rizzolatti, G., and Thier, P. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science* 324, 403–406. doi: 10.1126/science.1166818
- Caminiti, R., Ferraina, S., and Johnson, P. B. (1996). The sources of visual information to the primate frontal lobe: a novel role for the superior parietal lobule. *Cereb. Cortex* 6, 319–328. doi: 10.1093/cercor/6.3.319
- Campbell, A. (1905). *Histological Studies on the Localization of Cerebral Function*. London: Cambridge University Press.
- Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. doi: 10.1016/j.neuroimage.2009.12.112
- Cavada, C., and Goldman-Rakic, P. S. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J. Comp. Neurol.* 287, 422–445. doi: 10.1002/cne.902870403

- Cerri, G., Shimazu, H., Maier, M. A., and Lemon, R. N. (2003). Facilitation from ventral premotor cortex of primary motor cortex outputs to macaque hand muscles. *J. Neurophysiol.* 90, 832–842. doi: 10.1152/jn.01026.2002
- Cochin, S., Barthelemy, C., Roux, S., and Martineau, J. (1999). Observation and execution of movement: similarities demonstrated by quantified electroencephalography. *Eur. J. Neurosci.* 11, 1839–1842. doi: 10.1046/j.1460-9568.1999.00598.x
- Colby, C. L. (1998). Action-oriented spatial reference frames in cortex. *Neuron* 20, 15–24. doi: 10.1016/S0896-6273(00)80429-8
- Colby, C. L., Duhamel, J.-R., and Goldberg, M. E. (1993). Ventral intraparietal area of the macaque: anatomic location and visual response properties. *J. Neurophysiol.* 69, 902–914.
- Cooke, D. F., Taylor, C. S. R., Moore, T., and Graziano, M. S. A. (2003). Complex movements evoked by microstimulation of the ventral intraparietal area. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6163–6168. doi: 10.1073/pnas.1031751100
- Coudé, G., Festante, F., Cilia, A., Loiacono, V., Bimbi, M., Fogassi, L., et al. (2013). *The Visual Discharge of Mirror Neurons in Monkey Pmv is Modulated by the Gaze Direction of the Observed Agent*. San Diego, CA: Society for Neuroscience.
- Dancuse, N., Barbay, S., Frost, S. B., Plautz, E. J., Chen, D., Zoubina, E. V., et al. (2005). Extensive cortical rewiring after brain injury. *J. Neurosci.* 25, 10167–10179. doi: 10.1523/JNEUROSCI.3256-05.2005
- Dancuse, N., Barbay, S., Frost, S. B., Zoubina, E. V., Plautz, E. J., Mahnken, J. D., et al. (2006). Effects of small ischemic lesions in the primary motor cortex on neurophysiological organization in ventral premotor cortex. *J. Neurophysiol.* 96, 3506–3511. doi: 10.1152/jn.00792.2006
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180. doi: 10.1007/BF00230027
- Duhamel, J.-R., Colby, C. L., and Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J. Neurophysiol.* 79, 126–136.
- Dum, R. P., and Strick, P. L. (1991). The origin of corticospinal projections from the premotor areas in the frontal lobe. *J. Neurosci.* 11, 667–689.
- Evangelou, M. N., Raos, V., Galletti, C., and Savaki, H. E. (2009). Functional imaging of the parietal cortex during action execution and observation. *Cereb. Cortex* 19, 624–639. doi: 10.1093/cercor/bhn116
- Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* 73, 2608–2611.
- Fagg, A. H., and Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Netw.* 11, 1277–1303. doi: 10.1016/S0893-6080(98)00047-1
- Ferrari, P. F., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714. doi: 10.1046/j.1460-9568.2003.02601.x
- Ferrier, D. (1873). Experimental researches in cerebral physiology and pathology. *J. Anat. Physiol.* 8, 152–155.
- Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771. doi: 10.1038/nature01861
- Fluet, M.-C., Baumann, M. A., and Scherberger, H. (2010). Context-specific grasp movement representation in macaque ventral premotor cortex. *J. Neurosci.* 30, 15175–15184. doi: 10.1523/JNEUROSCI.3343-10.2010
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. doi: 10.1126/science.1106138
- Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., and Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey: a reversible inactivation study. *Brain* 124, 571–586. doi: 10.1093/brain/124.3.571
- Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., and Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *J. Neurophysiol.* 76, 141–157.
- Fritsch, G., and Hitzig, E. (1870). “Ueber die elektrische Erregbarkeit des Grosshirns,” in *The Cerebral Cortex*, ed. W. W. Nowinski (Springfield, IL: Thomas), 73–96.
- Frost, S. B., Barbay, S., Friel, K. M., Plautz, E. J., and Nudo, R. J. (2003). Reorganization of remote cortical regions after ischemic brain injury: a potential substrate for stroke recovery. *J. Neurophysiol.* 89, 3205–3214. doi: 10.1152/jn.01143.2002
- Fujii, N., Hihara, S., and Iriki, A. (2007). Dynamic social adaptation of motion-related neurons in primate parietal cortex. *PLoS ONE* 2: e397. doi: 10.1371/journal.pone.0000397
- Fujii, N., Hihara, S., and Iriki, A. (2008). Social cognition in premotor and parietal cortex. *Soc. Neurosci.* 3, 250–260. doi: 10.1080/17470910701434610
- Fulton, J. (1935). A note on the definition of the “motor” and “premotor” areas. *Brain* 58, 311–316. doi: 10.1093/brain/58.2.311
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi: 10.1093/brain/119.2.593
- Gallese, V., Murata, A., Kaseda, M., Niki, N., and Sakata, H. (1994). Deficit of hand preshaping after muscimol injection in monkey parietal cortex. *Neuroreport* 5, 1525–1529. doi: 10.1097/00001756-199407000-00029
- Gangitano, M., Mottaghy, F. M., and Pascual-Leone, A. (2001). Phase-specific modulation of cortical motor output during movement observation. *Neuroreport* 12, 1489–1492. doi: 10.1097/00001756-200105250-00038
- Gazzola, V., and Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb. Cortex* 19, 1239–1255. doi: 10.1093/cercor/bhn181
- Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., and Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey. I. Somatotopy and the control of proximal movements. *Exp. Brain Res.* 71, 475–490. doi: 10.1007/BF00248741
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., and Luppino, G. (2010). Cortical connections of the macaque caudal ventrolateral prefrontal areas 45A and 45B. *Cereb. Cortex* 20, 141–168. doi: 10.1093/cercor/bhp087
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., and Luppino, G. (2011). Cortical connections of the anterior (F5a) subdivision of the macaque ventral premotor area F5. *Brain Struct. Funct.* 216, 43–65. doi: 10.1007/s00429-010-0293-6
- Gerbella, M., Borra, E., Tonelli, S., Rozzi, S., and Luppino, G. (2013). Connectional heterogeneity of the ventral part of the macaque area 46. *Cereb. Cortex* 23, 967–987. doi: 10.1093/cercor/bhs096
- Geyer, S., Matelli, M., Luppino, G., and Zilles, K. (2000). Functional neuroanatomy of the primate isocortical motor system. *Anat. Embryol. (Berl.)* 202, 443–474. doi: 10.1007/s004290000127
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
- Grafton, S. T., Arbib, M. A., Fadiga, L., and Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Exp. Brain Res.* 112, 103–111. doi: 10.1007/BF00227183
- Graziano, M. S. A. (2006). The organization of behavioral repertoire in motor cortex. *Annu. Rev. Neurosci.* 29, 105–134. doi: 10.1146/annurev.neuro.29.051605.112924
- Grèzes, J., Costes, N., and Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain* 122, 1875–1887. doi: 10.1093/brain/122.10.1875
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15061–15065. doi: 10.1073/pnas.95.25.15061
- He, S. Q., Dum, R. P., and Strick, P. L. (1993). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere. *J. Neurosci.* 13, 952–980.
- He, S. Q., Dum, R. P., and Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *J. Neurosci.* 15, 3284–3306.
- Hepp-Reymond, M. C., Hüsler, E. J., Maier, M. A., and Qi, H. X. (1994). Force-related neuronal activity in two regions of the primate ventral premotor cortex. *Can. J. Physiol. Pharmacol.* 72, 571–579. doi: 10.1139/y94-081
- Hoshi, E., Shima, K., and Tanji, J. (1998). Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex. *J. Neurophysiol.* 80, 3392–3397.
- Hyvärinen, J. (1981). Regional distribution of functions in parietal association area 7 of the monkey. *Brain Res.* 206, 287–303. doi: 10.1016/0006-8993(81)90533-3
- Hyvärinen, J. (1982). Posterior parietal lobe of the primate brain. *Physiol. Rev.* 62, 1060–1129.

- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3:e79. doi: 10.1371/journal.pbio.0030079
- Iriki, A., and Sakura, O. (2008). The neuroscience of primate intellectual evolution: natural selection and passive and intentional niche construction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2229–2241. doi: 10.1098/rstb.2008.2274
- Isa, T., Ohki, Y., Alstermark, B., Pettersson, L.-G., and Sasaki, S. (2007). Direct and indirect corticomotoneuronal pathways and control of hand/arm movements. *Physiology* 22, 145–152. doi: 10.1152/physiol.00045.2006
- Ishida, H., Nakajima, K., Inase, M., and Murata, A. (2010). Shared mapping of own and others' bodies in visuotactile bimodal area of monkey parietal cortex. *J. Cogn. Neurosci.* 22, 83–96. doi: 10.1162/jocn.2009.21185
- Jeannerod, M. (2006). *Motor Cognition: What Actions Tell to the Self*. Oxford: Oxford University Press.
- Jeannerod, M., Arbib, M. A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320. doi: 10.1016/0166-2236(95)93921-J
- Keizer, K., and Kuypers, H. G. (1989). Distribution of corticospinal neurons with collaterals to the lower brain stem reticular formation in monkey (*Macaca fascicularis*). *Exp. Brain Res.* 74, 311–318. doi: 10.1007/BF00248864
- Keysers, C., Kohler, E., Umliltá, M. A., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Exp. Brain Res.* 153, 628–636. doi: 10.1007/s00221-003-1603-5
- Kinoshita, M., Matsui, R., Kato, S., Hasegawa, T., Kasahara, H., Isa, K., et al. (2012). Genetic dissection of the circuit for hand dexterity in primates. *Nature* 487, 235–238. doi: 10.1038/nature11206
- Klein, J. T., Shepherd, S. V., and Platt, M. L. (2009). Social attention and the brain. *Curr. Biol.* 19, R958–R962. doi: 10.1016/j.cub.2009.08.010
- Kohler, E., Keysers, C., Umliltá, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848. doi: 10.1126/science.1070311
- Kraskov, A., Dancause, N., Qualló, M. M., Shepherd, S., and Lemon, R. N. (2009). Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron* 64, 922–930. doi: 10.1016/j.neuron.2009.12.010
- Kuhlman, W. N. (1978). Functional topography of the human mu rhythm. *Electroencephalogr. Clin. Neurophysiol.* 44, 83–93. doi: 10.1016/0013-4694(78)90107-4
- Kurata, K., and Tanji, J. (1986). Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements. *J. Neurosci.* 6, 403–411.
- Leinonen, L., Hyvärinen, J., Nyman, G., and Linnankoski, I. (1979). I. Functional properties of neurons in lateral part of associative area 7 in awake monkeys. *Exp. Brain Res.* 34, 299–320. doi: 10.1007/bf00235675
- Lemon, R. N. (2008). Descending pathways in motor control. *Annu. Rev. Neurosci.* 31, 195–218. doi: 10.1146/annurev.neuro.31.060407.125547
- Lewis, J. W., and Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J. Comp. Neurol.* 428, 112–137. doi: 10.1002/1096-9861(20001204)428:1<112::AID-CNE8>3.0.CO;2-9
- Lu, M. T., Preston, J. B., and Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.* 341, 375–392. doi: 10.1002/cne.903410308
- Luppino, G., Matelli, M., Camarda, R., and Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *J. Comp. Neurol.* 338, 114–140. doi: 10.1002/cne.903380109
- Luppino, G., Murata, A., Govoni, P., and Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp. Brain Res.* 128, 181–187. doi: 10.1007/s002210050833
- Maeda, F., Kleiner-Fisman, G., and Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation. *J. Neurophysiol.* 87, 1329–1335. doi: 10.1152/jn.00773.2000
- Mason, A. H., and Mackenzie, C. L. (2005). Grip forces when passing an object to a partner. *Exp. Brain Res.* 163, 173–187. doi: 10.1007/s00221-004-2157-x
- Matelli, M., Camarda, R., Glickstein, M., and Rizzolatti, G. (1986). Afferent and efferent projections of the inferior area 6 in the macaque monkey. *J. Comp. Neurol.* 251, 281–298. doi: 10.1002/cne.902510302
- Matelli, M., Luppino, G., and Rizzolatti, G. (1985). Patterns of cytochrome oxidase activity in the frontal agranular cortex of the macaque monkey. *Behav. Brain Res.* 18, 125–136. doi: 10.1016/0166-4328(85)90068-3
- Matelli, M., Luppino, G., and Rizzolatti, G. (1991). Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *J. Comp. Neurol.* 311, 445–462. doi: 10.1002/cne.903110402
- Mendres, K., and de Waal, F. B. (2000). Capuchins do cooperate: the advantage of an intuitive task. *Anim. Behav.* 60, 523–529. doi: 10.1006/anbe.2000.1512
- Meulenbroek, R. G. J., Bosga, J., Hulstijn, M., and Miedl, S. (2007). Joint-action coordination in transferring objects. *Exp. Brain Res.* 180, 333–343. doi: 10.1007/s00221-007-0861-z
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Morecraft, R. J., Louie, J. L., Herrick, J. L., and Stilwell-Morecraft, K. S. (2001). Cortical innervation of the facial nucleus in the non-human primate: a new interpretation of the effects of stroke and related subtotal brain trauma on the muscles of facial expression. *Brain* 124, 176–208. doi: 10.1093/brain/124.1.176
- Moschovakis, A. K. (2004). Oculomotor areas of the primate frontal lobes: a transneuronal transfer of rabies virus and [14C]-2-deoxyglucose functional imaging study. *J. Neurosci.* 24, 5726–5740. doi: 10.1523/JNEUROSCI.1223-04.2004
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* 20, 750–756. doi: 10.1016/j.cub.2010.02.045
- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., Rizzolatti, G., et al. (1997). Object representation in the ventral premotor cortex (Area F5) of the monkey. *J. Neurophysiol.* 78, 2226–2230.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., and Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J. Neurophysiol.* 83, 2580–2601.
- Muthukumaraswamy, S. D., and Johnson, B. W. (2004). Changes in rolandic mu rhythm during observation of a precision grip. *Psychophysiology* 41, 152–156. doi: 10.1046/j.1469-8986.2003.00129.x
- Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., et al. (2011). Action observation circuits in the macaque monkey cortex. *J. Neurosci.* 31, 3743–3756. doi: 10.1523/JNEUROSCI.4803-10.2011
- Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., and Orban, G. A. (2005). Observing others: multiple action representation in the frontal lobe. *Science* 310, 332–336. doi: 10.1126/science.1115593
- Neuper, C., Wörtz, M., and Pfurtscheller, G. (2006). ERD/ERS patterns reflecting sensorimotor activation and deactivation. *Prog. Brain Res.* 159, 211–222. doi: 10.1016/S0079-6123(06)59014-4
- Nishitani, N., and Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proc. Natl. Acad. Sci. U.S.A.* 97, 913–918. doi: 10.1073/pnas.97.2.913
- Nishitani, N., and Hari, R. (2002). Viewing lip forms: cortical dynamics. *Neuron* 36, 1211–1220. doi: 10.1016/S0896-6273(02)01089-9
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nat. Neurosci.* 10, 817–818. doi: 10.1038/nn1911
- Nishimura, Y., and Isa, T. (2012). Cortical and subcortical compensatory mechanisms after spinal cord injury in monkeys. *Exp. Neurol.* 235, 152–161. doi: 10.1016/j.expneurol.2011.08.013
- Nishimura, Y., Onoe, H., Morichika, Y., Perfiliev, S., Tsukada, H., and Isa, T. (2007). Time-dependent central compensatory mechanisms of finger dexterity after spinal cord injury. *Science* 318, 1150–1155. doi: 10.1126/science.1147243
- Nudo, R. J. (2007). Postinfarct cortical plasticity and behavioral recovery. *Stroke* 38, 840–845. doi: 10.1161/01.STR.0000247943.12887.d2
- Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., and Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Brain Res. Cogn. Brain Res.* 24, 190–198. doi: 10.1016/j.cogbrainres.2005.01.014
- Oberman, L. M., Pineda, J. A., and Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Soc. Cogn. Affect. Neurosci.* 2, 62–66. doi: 10.1093/scan/0102

- Pani, P., Theys, T., Romero, M., and Janssen, P. (2014). Grasping execution and grasping observation activity of single neurons in the macaque anterior intraparietal area. *J. Cogn. Neurosci.* 26, 1–14. doi: 10.1162/jocn_a_00647
- Passingham, R. E. (2009). How good is the macaque monkey model of the human brain? *Curr. Opin. Neurobiol.* 19, 6–11. doi: 10.1016/j.conb.2009.01.002
- Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., and McCarthy, G. (2003). Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *J. Neurosci.* 23, 6819–6825.
- Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., et al. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage* 14, 749–758. doi: 10.1006/nimg.2001.0872
- Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., et al. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* 146, 87–113.
- Petrides, M., and Pandya, D. N. (1984). Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *J. Comp. Neurol.* 228, 105–116. doi: 10.1002/cne.902280110
- Pfurtscheller, G., and Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neurosci. Lett.* 239, 65–68. doi: 10.1016/S0304-3940(97)00889-6
- Pineda, J. A. (2005). The functional significance of mu rhythms: translating “seeing” and “hearing” into “doing.” *Brain Res. Rev.* 50, 57–68. doi: 10.1016/j.brainresrev.2005.04.005
- Pitti, A., Alirezai, H., and Kuniyoshi, Y. (2009). Cross-modal and scale-free action representations through enaction. *Neural Netw.* 22, 144–154. doi: 10.1016/j.neunet.2009.01.007
- Porter, R., and Lemon, R. N. (1993). *Corticospinal Function and Voluntary Movement*. Oxford: Oxford University Press.
- Prabhu, G., Shimazu, H., Cerri, G., Brochier, T., Spinks, R. L., Maier, M. A., et al. (2009). Modulation of primary motor cortex outputs from ventral premotor cortex during visually guided grasp in the macaque monkey. *J. Physiol.* 587, 1057–1069. doi: 10.1113/jphysiol.2008.165571
- Preuss, T. M., and Goldman-Rakic, P. S. (1989). Connections of the ventral granular frontal cortex of macaques with perisylvian premotor and somatosensory areas: anatomical evidence for somatic representation in primate frontal association cortex. *J. Comp. Neurol.* 282, 293–316. doi: 10.1002/cne.902820210
- Puce, A., Allison, T., Bentin, S., Gore, J. C., and McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* 18, 2188–2199.
- Raos, V., Kilintari, M., and Savaki, H. E. (2014). Viewing a forelimb induces widespread cortical activations. *Neuroimage* 89, 122–142. doi: 10.1016/j.neuroimage.2013.12.010
- Raos, V., Umiltà, M. A., Murata, A., Fogassi, L., and Gallese, V. (2006). Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey. *J. Neurophysiol.* 95, 709–729. doi: 10.1152/jn.00463.2005
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp. Brain Res.* 71, 491–507. doi: 10.1007/BF00248742
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., and Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706. doi: 10.1152/physrev.00009.2013
- Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996a). Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996b). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252. doi: 10.1007/BF00227301
- Rizzolatti, G., and Fogassi, L. (2014). The mirror mechanism: recent findings and perspectives. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130420. doi: 10.1098/rstb.2013.0420
- Rizzolatti, G., Fogassi, L., and Gallese, V. (1997). Parietal cortex: from sight to action. *Curr. Opin. Neurobiol.* 7, 562–567. doi: 10.1016/S0959-4388(97)80037-2
- Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron* 31, 889–901. doi: 10.1016/S0896-6273(01)00423-8
- Rizzolatti, G., Luppino, G., and Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalogr. Clin. Neurophysiol.* 106, 283–296. doi: 10.1016/S0013-4694(98)00022-4
- Rizzolatti, G., and Matelli, M. (2003). Two different streams form the dorsal visual system: anatomy and functions. *Exp. Brain Res.* 153, 146–157. doi: 10.1007/s00221-003-1588-0
- Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–274. doi: 10.1038/nrn2805
- Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G. G., Matelli, M., et al. (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cereb. cortex* 16, 1389–1417. doi: 10.1093/cercor/bhj076
- Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G., and Fogassi, L. (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *Eur. J. Neurosci.* 28, 1569–1588. doi: 10.1111/j.1460-9568.2008.06395.x
- Rozzi, S., Rodà, F., Simone, L., Bimbi, M., Luppino, G., and Fogassi, L. (2011). *Movement-Related Neurons in the Macaque Ventrolateral Prefrontal Cortex*. Washington, DC: Society for Neuroscience.
- Sakata, H., Taira, M., Murata, A., and Mine, S. (1995). Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cereb. cortex* 5, 429–438. doi: 10.1093/cercor/5.5.429
- Saleem, K. S., and Kondo, H. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *J. Comp. Neurol.* 693, 659–693. doi: 10.1002/cne.21577
- Sartori, L., Cavallo, A., Buccioni, G., and Castiello, U. (2011). Corticospinal excitability is specifically modulated by the social dimension of observed actions. *Exp. Brain Res.* 211, 557–568. doi: 10.1007/s00221-011-2650-y
- Sasaki, S., Isa, T., Pettersson, L.-G., Alstermark, B., Naito, K., Yoshimura, K., et al. (2004). Dexterous finger movements in primate without monosynaptic corticomotoneuronal excitation. *J. Neurophysiol.* 92, 3142–3147. doi: 10.1152/jn.00342.2004
- Sebanz, N., and Knoblich, G. (2008). “From mirroring to joint action,” in *Embodied Communication*, eds I. Wachsmuth, M. Lenzen, and G. Knoblich (Oxford: Oxford University Press), 129–150.
- Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Top. Cogn. Sci.* 1, 353–367. doi: 10.1111/j.1756-8765.2009.01024.x
- Shepherd, S. V., Klein, J. T., Deaner, R. O., and Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9489–9494. doi: 10.1073/pnas.0900419106
- Shimazu, H., Maier, M. A., Cerri, G., Kirkwood, P. A., and Lemon, R. N. (2004). Macaque ventral premotor cortex exerts powerful facilitation of motor cortex outputs to upper limb motoneurons. *J. Neurosci.* 24, 1200–1211. doi: 10.1523/JNEUROSCI.4731-03.2004
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., and Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp. Brain Res.* 83, 29–36. doi: 10.1007/BF00232190
- Tanila, H., Carlson, S., Linnankoski, I., Lindroos, F., and Kahila, H. (1992). Functional properties of dorsolateral prefrontal cortical neurons in awake monkey. *Behav. Brain Res.* 47, 169–180. doi: 10.1016/S0166-4328(05)80123-8
- Tanji, J., and Hoshi, E. (2008). Role of the lateral prefrontal cortex in executive behavioral control. *Physiol. Rev.* 88, 37–57. doi: 10.1152/physrev.00014.2007
- Tomasello, M., and Call, J. (1997). *Primate Cognition*. New York, NY: Oxford University Press.
- Tsao, D. Y., and Livingstone, M. S. (2008). Mechanisms of face perception. *Annu. Rev. Neurosci.* 31, 411–437. doi: 10.1146/annurev.neuro.30.051606.094238
- Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., et al. (2008). When pliers become fingers in the monkey motor system. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2209–2213. doi: 10.1073/pnas.0705985105
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing. A neurophysiological study. *Neuron* 31, 155–165. doi: 10.1016/S0896-6273(01)00337-3
- Visalberghi, E., Quarantotti, B. P., and Tranchida, F. (2000). Solving a cooperation task without taking into account the partner's behavior: the case of capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 114, 297–301. doi: 10.1037/0735-7036.114.3.297

- Visco-Comandini, F., Ferrari-Toniolo, S., Satta, E., Papazachariadis, O., Gupta, R., Nalbant, L. E., et al. (2015). Do non-human primates cooperate? Evidences of motor coordination during a joint action task in macaque monkeys. *Cortex*. doi: 10.1016/j.cortex.2015.02.006 [Epub ahead of print]
- Vogt, C., and Vogt, O. (1919). Allgemeine Ergebnisse unserer Hirnforschung. *J. Psychol. Neurol.* 25, 279–461.
- Von Bonin, G., and Bailey, P. (1947). *The Neocortex of Macaca Mulatta*. Urbana, IL: University of Illinois Press.
- Wise, S. P., Boussaoud, D., Johnson, P. B., and Caminiti, R. (1997). Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annu. Rev. Neurosci.* 20, 25–42. doi: 10.1146/annurev.neuro.20.1.25
- Wolpert, D. M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3(Suppl.), 1212–1217. doi: 10.1038/81497
- Woolsey, C., and Settlage, P. (1952). Patterns of localization in precentral and “supplementary” motor areas and their relation to the concept of a premotor area. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* 30, 238–264.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Rozzi and Coudé. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership