

# Automatic imitation of the arm kinematic profile in interacting partners

Alessandro D'Ausilio<sup>1</sup> · Leonardo Badino<sup>1</sup> · Pietro Cipresso<sup>2</sup> · Alice Chirico<sup>3</sup> · Elisabetta Ferrari<sup>1</sup> · Giuseppe Riva<sup>2,3</sup> · Andrea Gaggioli<sup>2,3</sup>

Published online: 31 July 2015

© Marta Olivetti Belardinelli and Springer-Verlag Berlin Heidelberg 2015

**Abstract** Cognitive neuroscience, traditionally focused on individual brains, is just beginning to investigate social cognition through realistic interpersonal interaction. However, quantitative investigation of the dynamical sensorimotor communication among interacting individuals in goal-directed ecological tasks is particularly challenging. Here, we recorded upper-body motion capture of 23 dyads, alternating their leader/follower role, in a tower-building task. Either a strategy of joining efforts or a strategy of independent action could in principle be used. We found that arm reach velocity profiles of participants tended to converge across trials. Automatic imitation of low-level motor control parameters demonstrates that the task is achieved through continuous action coordination as opposed to independent action planning. Moreover, the leader produced more consistent and predictable velocity profiles, suggesting an implicit strategy of signaling to the follower. This study serves as a validation of our joint goal-directed non-verbal task for future applications. In fact, the quantification of human-to-human continuous sensorimotor interaction, in a way that can be predicted and controlled, is probably one of the greatest challenges for the future of human–robot interaction.

**Keywords** Social neuroscience · Joint action · Automatic imitation · Kinematic analysis · Velocity profiles · Tower-building task

## Introduction

Group coordination in ecological tasks is difficult to measure. However, we do that instinctively by sending and receiving socially relevant messages in all our interactions (i.e., hand gestures, facial expressions, and speech). Among all possible communication channels, complex coordinated behavior can arise without the need for verbal communication to happen (Sebanz et al. 2006; Néda et al. 2000; Riley et al. 2011). Although non-verbal communication is clearly an important part of social competence, recent research has mainly focused on minimal interaction to maintain experimental control (D'Ausilio et al. 2015). However, in many everyday social interactions, behavior is continuous, temporally overlapping, and based on whole-body motion (Tognoli et al. 2007). Therefore, group-level behavioral dynamics should be studied as complex, interactive systems in which information transfer is continuous rather than discrete (D'Ausilio et al. 2012; Glowinski et al. 2013; Badino et al. 2014).

In this framework, group-level coordination is grounded upon our tuning to the detection of human biological motion primitives (Johansson 1973; Casile et al. 2010), which in turn drives the implicit motor contagion observed between interacting partners. The phenomenon of motor contagion (or automatic imitation) is the tendency to involuntarily reproduce specific features of observed actions. For instance, participants' movements are automatically contaminated by the velocity profile of a moving dot (Bisio et al. 2010). Such automatic motor contagion is

---

✉ Alessandro D'Ausilio  
alessandro.dausilio@iit.it

<sup>1</sup> RBCS – Robotics, Brain and Cognitive Sciences Department, Italian Institute of Technology (IIT), Via Morego, 30, 16163 Genoa, Italy

<sup>2</sup> Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Verbania, Italy

<sup>3</sup> Psychology Department, Catholic University of Milan, Milan, Italy

reduced when the interacting partner violates the biological laws of motion (Bisio et al. 2014). These findings suggest that a low-level sensory-motor matching mechanism forms the basis of higher levels of social interaction by facilitating group behavioral entrainment (Dumas et al. 2014).

Here, we exploit the concept of automatic imitation of kinematic profiles, in couples of participants playing a joint game. The game consisted in a tower-building task with wooden blocks (tower-building task). The goal of the game was to jointly balance the trade-off between speed and accuracy. Participants were alternating their role of either leader or follower, across trials (see “Methods” for details). We recorded upper-body motion capture with a Vicon system and tested whether automatic imitation of arm velocity profiles was present in each dyad. In principle, the task can be accomplished at different degrees of coordination among participants. The strongest coordination requires the negotiation of motor parameters and thus continuous information transfer. Otherwise, minimal coordination requires only the negotiation of action timing and thus reduced and discrete information transfer. The former end of this continuum would be evidenced by the mutual contagion of low-level motor control parameters. The latter instead would be evident if lags between individual actions are reduced in the absence of motor imitation. Moreover, if automatic kinematic imitation is present, the leader and the follower should differ in this regard. In fact, we predict that the leader should evidence greater kinematic predictability to help the follower build more reliable models of his/her behavior.

## Methods

### Subjects

The study involved 46 participants (23 males and 23 females) who were recruited among the Italian Institute of Technology staff (mean age 29.26,  $SD = 2.92$ ). Procedures were approved by the local ethics committee (ASL-3, Genova), and participants did not receive any compensation.

### Materials

We used three sheets of papers on which were drawn squares (Fig. 1a). Twelve colored cubes (four red, four yellow, and four green) were used to build the tower, and each performance was videotaped.

### Kinematic recordings

Participant movements were measured via a motion-capture system (Vicon system) with nine near-infrared

cameras with acquisition frequency set at 100 Hz. Reflective markers were placed on both shoulders and the dominant arm elbow and wrist, and head (four markers for the first subject and three for the second member of the couple; see Fig. 1c).

### Procedure

The experiment consisted of two phases and lasted about 20 min. In the first phase (i.e., single trial), each member of a couple performed the task alone, to familiarize with it. In the second phase (i.e., couple trials), each couple performed the task together 10 times (see Fig. 1b).

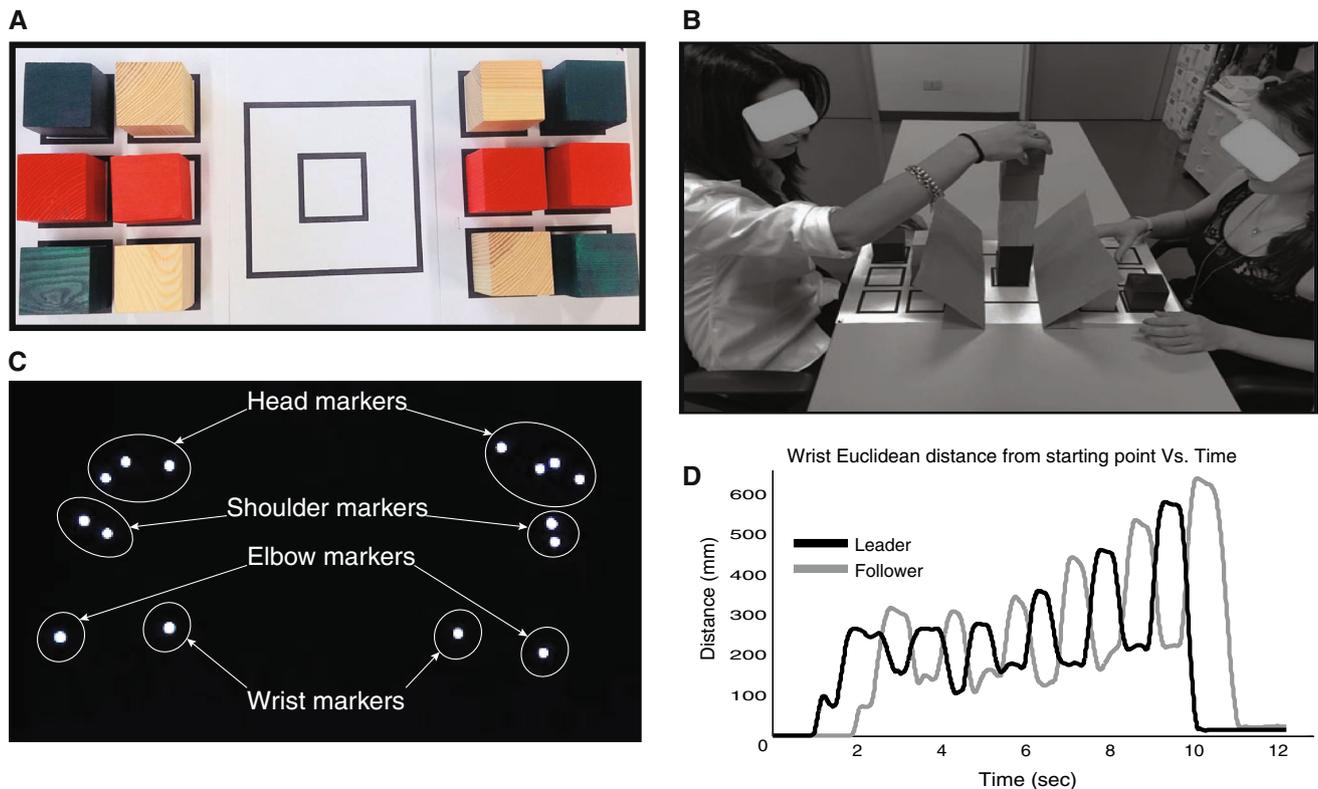
In the single trial, each member of a couple had to build a tower using all the 12 cubes. They were asked to build the tower as quickly as possible, by using just one hand, one cube at a time. In the couple trials, each couple had to build a tower, as quickly as possible, using all the 12 cubes (six each participant; see Fig. 1a). The only constraint concerned the specific turn-taking sequence. The leader had to place the first cube and choose the color of the cubes to place. The follower had to place a cube of the same color chosen by the leader. In each trial, the leader was switched, and thus, each participant acted as a leader for five trials and as a follower for the other five trials.

### Analyses

The trial performance of a dyad was measured as the average normalized interval between the instant in which the leader placed a cube on the top of the tower and the instant in which the follower placed the next cube on top of the tower (henceforth turn-taking time). The turn-taking time was computed by observing the peaks of the Euclidean distance between the wrist position at each time step and the wrist position at time 0 (see Fig. 1d). The turn-taking time was normalized to take into account the increasing distance between the cubes' original position and the top of the tower.

The automatic motor imitation was defined as the maximum cross-correlation value between the velocity profiles over time of the dominant wrist of the two participants. The minimum and maximum lag used to compute cross-correlation was set as the minimum and maximum (not-normalized) turn-taking time, respectively. Cross-correlation was only computed on trials where participants successfully completed the task.

To test the statistical significance of the measured maximum cross-correlation, we ran a bootstrap-based significance test (Efron and Tibshirani 1993). Each bootstrap replication consisted of the same number of time series pairs as the original experiment. To generate a bootstrap (normal) distribution of the cross-correlation (for each



**Fig. 1** **a** Top view of the setup; **b** a dyad executing the task; **c** Vicon point-light display of a dyad executing the task; **d** Euclidean distance between the instantaneous wrist position and the starting point wrist position of a dyad executing the task (color figure online)

trial), we created 10,000 bootstrap replications with replacement, where velocity profile pairs were sampled by coupling the time series of two randomly selected subjects belonging to two different dyads but building the tower at the same trial number. The “same trial number” constraint allows a stricter significance test that takes into account the increasing predictability of the velocity time series over trials which, in turn, can increase cross-correlation values (see “Discussion”). Since cross-correlation computation requires same-length time series, we “stretched” the shortest time series to equal the longest time series by applying resampling. Resampling is a global, uniform scaling technique as opposed to non-uniform techniques such as dynamic time warping which apply local transformations to the time series to adjust local mismatches between two time series and thus can heavily affect cross-correlation computation. At each trial number, we computed the average bootstrap maximum cross-correlation value and its bootstrap-*t* confidence interval. The bootstrap-*t* confidence interval was defined by setting a *p* value ( $p = 0.005$  after Bonferroni correction for multiple comparisons).

An additional analysis was carried out to assess whether the leader exhibits more predictable movements (in terms of velocity profile) than the follower. The predictability of

the leader (and the follower), given the history of her velocity profile within a trial, was defined as the maximum autocorrelation value computed in the minimum and maximum interval times that took the leader (the follower) to place two cubes on top of the tower.

## Results

The actual measured maximum cross-correlation is significant in most trials (except trial 1 and 4) according to the bootstrap-based significance test ( $p = 0.0053$ ,  $p = 0.0001$ ,  $p = 0.0001$ ,  $p = 0.0072$ ,  $p = 0.0001$ ,  $p = 0.0020$ ,  $p = 0.0001$ ,  $p = 0.0002$ ,  $p = 0.0033$ ,  $p = 0.0006$ , where *p* was determined by the position in the vector of bootstrap maximum cross-correlations of the closest value to the actual measured maximum cross-correlation).

The leader mean maximum autocorrelation (mean = 0.6732, SE = 0.0058) is significantly higher than the follower mean maximum autocorrelation (mean = 0.6582, SE = 0.0056) according to a two-tailed *t* test ( $t[410] = 1.9727$ ,  $p = 0.0492$ ). Similarly, the overall mean maximum cross-correlation (mean = 0.6772, SE = 0.0044) is significantly higher than the follower mean maximum autocorrelation ( $t[406] = 2.7757$ ,  $p = 0.0058$ ). There is no

significant difference between the overall mean maximum cross-correlation and the leader mean maximum autocorrelation ( $t[406] = 0.5749$ ,  $p = 0.5657$ ).

## Discussion

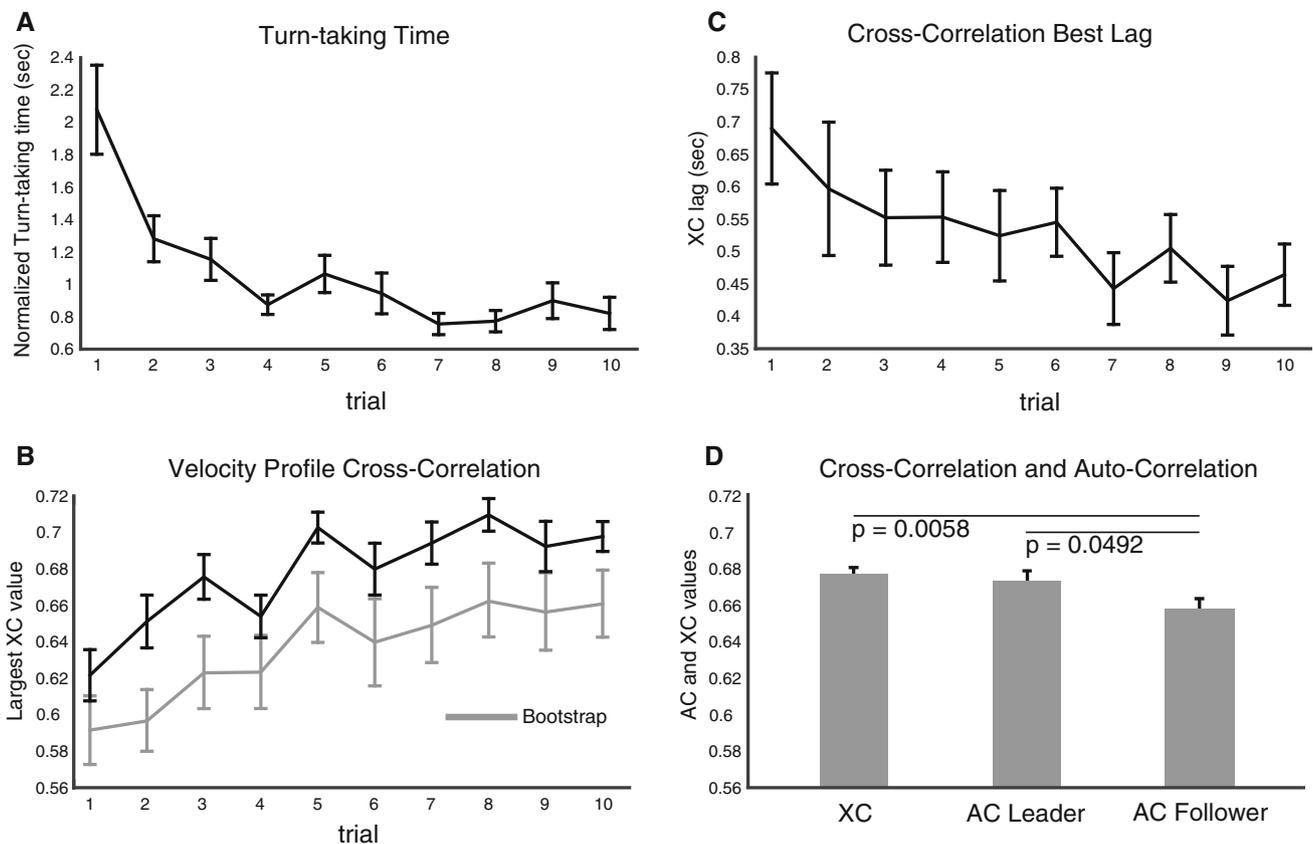
As expected, participants improve their performance across trials. This is shown by the reduction in turn-taking time that reaches a plateau around the seventh trial (Fig. 2a). Concurrently, the increase in kinematic motor imitation (Fig. 2b) means that participants were showing contagion of low-level motor parameters. At the same time, kinematic imitation lags were also reduced (Fig. 2c), suggesting a process of continuous optimization of their imitation strategies, throughout the task.

Interestingly, the leader was also generating velocity profiles that were more predictable than that of the follower (Fig. 2d). High kinematic predictability helps the follower build more efficient models of leader's behavior. The leader may have indeed implicitly put into place a strategy to optimize coordination, increasing behavioral consistency,

and ultimately kinematic readability and predictability. This in turn may explain the better task performance.

This phenomenon is often referred to as signaling (Pezzulo and Dindo 2011; Pezzulo et al. 2013). In fact, when participants have access to different amounts of information, joint action is asymmetric (Schmidt et al. 2011). To cope with such unbalance, the leader is supposed to either make his action more predictable (Vesper et al. 2011) or perform her actions in a way that provides relevant task information to the partner (Sacheli et al. 2013).

The tower-building task we present here has, however, some critical features that set it apart from previous similar studies (i.e., Vesper and Richardson 2014). Subjects did not have to imitate each other nor they had to synchronize the timing of their pick-move-place action. Target position (the tower) was always the same, and thus, there is no added value in signaling its location through the modulation of kinematic parameters. The only motoric constraint was to optimize the turn taking. Therefore, an optimal coordination strategy might have been to simply increase peak velocity of their ballistic movements—not motor imitation.



**Fig. 2** **a** Normalized turn-taking time across trials (*bars* represent standard error of the mean); **b** maximum cross-correlation value of velocity profiles across trials. In *black*, real computed values (*bars* represent the standard error). In *gray*, the bootstrap results (*bars* are the confidence intervals); **c** best lag for the maximum cross-

correlation value (*bars* represent the standard error); **d** comparison between the average across trials of the cross-correlation between participants and autocorrelation of the leader and follower (*bars* represent the standard error)

The only information that the leader might have wanted to signal was the next cube to place (or its color). In our first iteration of the tower-building task, the relative position of colored cubes was kept constant and a screen prevented from seeing object grasping (Fig. 1b). Nevertheless, arm trajectories could still be modulated to offer some hints about the location of the next cube. However, it is important to note that by focusing on velocity profiles, instead of position data, our analyses are independent from this kind of signaling strategies. Importantly, we show that as a by-product of action coordination, participants exhibit a motor contagion that does not seem to be directly functional to task optimization. More interestingly, the fact that the leader generates velocity profiles with higher predictability suggests that motor contagion is not merely a side effect but may offer a significant coordination advantage.

The present work is a first step toward the standardization of the tower-building task and the definition of a set of features and analyses to extract sensorimotor communication. However, there are serious technical issues to be solved, when trying to quantify real social interaction. The first one regards the complexity of capturing whole-body signals in ecological scenarios. State-of-the-art motion-capture technologies already allow accurate recording of whole-body kinematics, although these methods are typically expensive and time-consuming. At the same time, these methods completely neglect the dynamic component of body movement and thus the forces exchanged with the environment. On the other hand, data analysis is the real open issue when it comes to data interpretation. In fact, there is no standard procedure to analyze data or extract quantitative markers of group sensorimotor coordination and information flow between participants. These issues are the real bottleneck to allow the use of these methods in applied scenarios. Future work from our group directly aims at solving both issues.

Finally, we believe that quantifying the sensorimotor markers of real social interaction in ecological contexts may have important applied relevance. For example, diagnosis and rehabilitation of social communication disorders may benefit from standardized tasks and quantitative methods extracting these automatic low-level motor markers of interaction. Indeed, such detailed description of interactive non-verbal behavior may be useful in the classification of patients' subtypes, monitoring progress during rehabilitation or even comparing different protocols.

## References

- Badino L, D'Ausilio A, Glowinski D, Camurri A, Fadiga L (2014) Sensorimotor communication in professional quartets. *Neuropsychologia* 55:98–104
- Bisio A, Stucchi N, Jacono M, Fadiga L, Pozzo T (2010) Automatic versus voluntary motor imitation: effect of visual context and stimulus velocity. *PLoS One* 5(10):e13506
- Bisio A, Sciutti A, Nori F, Metta G, Fadiga L, Sandini G, Pozzo T (2014) Motor contagion during human–human and human–robot interaction. *PLoS One* 9(8):e106172
- Casile A, Dayan E, Caggiano V, Hessler T, Flash T, Giese MA (2010) Neuronal encoding of human kinematic invariants during action observation. *Cereb Cortex* 20(7):1647–1655
- D'Ausilio A, Badino L, Li Y, Tokay S, Craighero L, Canto R, Aloimonos Y, Fadiga L (2012) Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS One* 7(5):e35757
- D'Ausilio A, Novembre G, Fadiga L, Keller PE (2015) What can music tell us about social interaction? *Trends Cogn Sci* 19(3):111–114
- Dumas G, Laroche J, Lehmann A (2014) Your body, my body, our coupling moves our bodies. *Front Hum Neurosci* 8:1004
- Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall, New York
- Glowinski D, Mancini M, Cowie R, Camurri A, Chiorri C, Doherty C (2013) The movements made by performers in a skilled quartet: a distinctive pattern, and the function that it serves. *Front Psychol* 4:841
- Johansson G (1973) Visual perception of biological motion and a model for its analysis. *Percept Psychophys* 14(2):201–211
- Néda Z, Ravasz E, Brechet Y, Vicsek T, Barabási AL (2000) The sound of many hands clapping. *Nature* 403(6772):849–850
- Pezzulo G, Dindo H (2011) What should I do next? using shared representations to solve interaction problems. *Exp Brain Res* 211(3–4):613–630
- Pezzulo G, Donnarumma F, Dindo H (2013) Human sensorimotor communication: a theory of signaling in online social interactions. *PLoS One* 8(11):e79876
- Riley MA, Richardson MJ, Shockley K, Ramenzoni VC (2011) Interpersonal synergies. *Front Psychol* 2:38
- Sacheli LM, Tidoni E, Pavone EF, Aglioti SM, Candidi M (2013) Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Exp Brain Res* 226(4):473–486
- Schmidt RC, Fitzpatrick P, Caron R, Mergeche J (2011) Understanding social motor coordination. *Hum Mov Sci* 30(5):834–845
- Sebanz N, Bekkering H, Knoblich G (2006) Joint action: bodies and minds moving together. *Trends Cogn Sci* 10(2):70–76
- Tognoli E, Lagarde J, DeGuzman GC, Kelso JA (2007) The phi complex as a neuromarker of human social coordination. *Proc Natl Acad Sci USA* 104(19):8190–8195
- Vesper C, Richardson MJ (2014) Strategic communication and behavioral coupling in asymmetric joint action. *Exp Brain Res* 232(9):2945–2956
- Vesper C, van der Wel RP, Knoblich G, Sebanz N (2011) Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Exp Brain Res* 211(3–4):517–530