

# Strategic communication and behavioral coupling in asymmetric joint action

Cordula Vesper · Michael J. Richardson

Received: 17 October 2013 / Accepted: 2 May 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** How is coordination achieved in asymmetric joint actions where co-actors have unequal access to task information? Pairs of participants performed a non-verbal tapping task with the goal of synchronizing taps to different targets. We tested whether ‘Leaders’ knowing the target locations would support ‘Followers’ without this information. Experiment 1 showed that Leaders tapped with higher amplitude that also scaled with specific target distance, thereby emphasizing differences between correct targets and possible alternatives. This strategic communication only occurred when Leaders’ movements were fully visible, but not when they were partially occluded. Full visual information between co-actors also resulted in higher and more stable behavioral coordination than partial vision. Experiment 2 showed that Leaders’ amplitude adaptation facilitated target prediction by independent Observers. We conclude that fully understanding joint action coordination requires both representational (i.e., strategic adaptation) and dynamical systems (i.e., behavioral coupling) accounts.

**Keywords** Joint action · Interpersonal coordination · Signaling · Phase synchronization

---

C. Vesper (✉)  
Department of Cognitive Science, Central European University,  
Frankel Leó út 30-34, Budapest 1023, Hungary  
e-mail: vesperc@ceu.hu

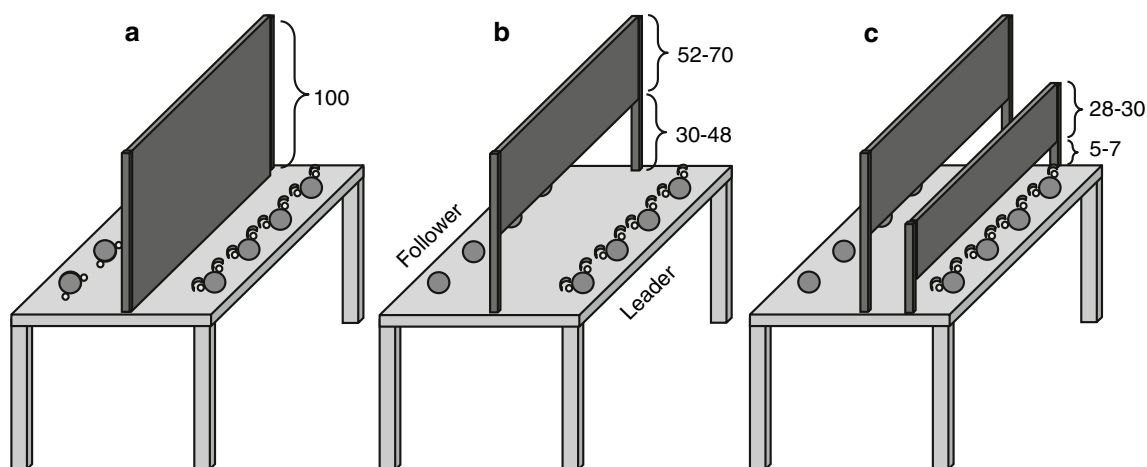
M. J. Richardson  
Center for Cognition, Action and Perception, University  
of Cincinnati, 4150-B Edwards C1, Cincinnati,  
OH 45221-0376, USA

## Introduction

In many joint action tasks, not all participating individuals have access to the same amount or the same type of information. That is, joint actions can be unbalanced (Clark 1996) or asymmetric (Richardson et al. 2011; Schmidt et al. 2011). For example, when carrying a sofa up the stairs, only the person walking forward can see what is coming ahead, whereas the person walking backward has to rely on visual or verbal cues from the partner to successfully make it to the upper floor. Asymmetric joint actions offer an interesting test case for understanding the various levels of interaction between co-actors: In order to achieve the joint action goal (moving the sofa to the top floor), co-actors need to coordinate their actions in real time, thereby relying on shared information (e.g., haptic information transmitted via the sofa), but also facing different amounts of task knowledge and different action roles (e.g., one pushing the other in the right direction and the other being guided).

The effects of asymmetries on coordination are manifold. On the one hand, asymmetries often disrupt coordination between interacting people, especially if information about a task partner’s action performance is lacking. For example, temporal coordination in musical duet performance is impaired when musicians can only hear their partner or only their own playing compared to the more natural situation with bidirectional feedback (Goebel and Palmer 2009). On the other hand, asymmetries can be beneficial for coordination, for instance if they help with negotiating task roles so that performance is more efficient (Brennan et al. 2008; Vesper et al. 2013) and potential collisions between co-actors’ limbs are avoided (Richardson et al. 2011).

In the present study, task knowledge was distributed asymmetrically between co-actors who were instructed to synchronize the endpoints of a sequence of whole-arm



**Fig. 1** Setup used in Experiment 1 along with occluder heights in cm. **a** Individual baseline condition; **b** full vision condition; **c** partial vision condition. Note that in the partial vision condition, the right

occluder did not reach down to the table surface, thereby allowing Followers to see the endpoints of Leaders' trajectories toward the target. Occluder heights were individually adjusted

tapping movements to different targets that were displayed on either side of a touch-screen table (Fig. 1). Thus, the task involved coordinating a spatial (the location of the subsequent target) and a temporal (the timing of tapping onto that target) joint action aspect. Task knowledge was asymmetric because, like in the example of moving a sofa, only one person (*Leader*) received spatial cues about the location of the subsequent tapping target, whereas the other (*Follower*) had to infer where the next movement should be directed based on the *Leader's* movements.

The aim of the present study was to determine how co-actors coordinate taps in this asymmetric task knowledge context. Considering distinct theoretical perspectives on joint action, we hypothesized different aspects of the task to be affected by the asymmetry. First, consistent with a strategic approach to joint action (Vesper et al. 2010), we hypothesized that *Leaders*, who had more task information, would adapt their tapping performance to assist *Followers* in finding the correct target location. Second, in line with a behavioral dynamics account to joint action (Schmidt and Richardson 2008), we hypothesized that co-actors' movements would become synchronized as an effect of monitoring each other's actions and that the patterning and stability of this synchronous coordination would be modulated in accordance with task and information asymmetries. To test these two hypotheses, we manipulated the amount of available visual information between co-actors. In the *full vision* condition, co-actors could see each other's complete movements. In the *partial vision* condition, *Leaders* could see *Followers'* movements, whereas *Followers* could only see the start- and endpoints of *Leaders'* movements. A third *individual baseline* condition was included to compare joint to individual behavior not requiring any coordination. In the following, the two theoretical perspectives on joint

action coordination will be described in more detail. In our opinion, both perspectives are relevant for understanding interpersonal tasks despite their theoretical and methodological differences.

Our first hypothesis was that *Leaders*, who have specific task knowledge, would adapt their movement performance in a way that would help *Followers* in tapping onto the correct target location. This prediction is based on the idea that in order to achieve a joint action outcome, individuals would strategically adapt their own behavior (Vesper et al. 2010). Conceptualizing changes of individual action performance as an effect of their intention to coordinate with someone else presupposes a representational format underlying action planning and performance (Clark 1996; Knoblich et al. 2011; Vesper et al. 2010). For instance, co-actors form representations of each other's tasks (Sebanz et al. 2003; Tsai et al. 2011) and predict and monitor others' actions through internal forward models (de Bruijn et al. 2009; Keller 2008; Loehr et al. 2013; Vesper et al. 2013; Wolpert et al. 2003). Based on such task and motor representations, co-actors can adapt their action performance and thereby support coordination, e.g., by making their actions more salient and predictable (Goebel and Palmer 2009; Vesper et al. 2011) or by performing their actions in a way that provides an uninformed partner with relevant task information (Sacheli et al. 2013). This is referred to as signaling (Pezzulo and Dindo 2011; Pezzulo et al. 2013) and assumes that people can intentionally alter their movements in a way that allows others to read information from it (Becchio et al. 2010; Sartori et al. 2009).

From a representational perspective, asymmetric joint actions will require strategic adaptations from co-actors to overcome the challenges of unequal task knowledge.

Specifically, in the present task, we predicted that Leaders would change the way they perform their tapping movements to make them salient for Followers and to provide Followers with relevant task information. Importantly, by comparing full and partial vision conditions, we manipulated the extent to which Followers could see and therefore read information provided by Leaders. We expected that a communicative signaling strategy would mainly be used in the full vision condition where sufficient visual information was available for Followers and less so in the partial vision condition. No signaling was expected when participants performed the task individually without a joint task goal. A second experiment investigated whether Leaders' action adaptation would indeed support others in acquiring information.

Our second hypothesis was that the task asymmetry (i.e., Leader's task vs. Follower's task) and the degree of available visual movement information would modulate and constrain the patterning and strength of the spontaneous movement coordination that occurred between co-actors. This is in line with a behavioral dynamics perspective on joint action and social movement coordination (Coey et al. 2012; Marsh et al. 2009; Schmidt and Richardson 2008), which postulates that the occurrence and patterning of interpersonal coordination is the result of general dynamical principles, and like the coordination that characterizes many other biological and non-biological systems (e.g., synchronized pendulum clocks, the coordinated activity of social insects, or the synchronized behavior of schools of fish), is emergent and self-organized. One well-studied form of emergent coordination is the in-phase synchrony or entrainment of cyclic movements (Richardson et al. 2007; Schmidt et al. 1990). In-phase synchrony arises if the dynamics of independent systems (conceptualized as oscillators) are coupled (i.e., influence each other). This coupling does not need to be physical and can be informational; for instance, previous research has demonstrated how visual or auditory information alone can result in stable interpersonal rhythmic entrainment (Schmidt and Richardson 2008).

Thus, when two independent individual participants perform the present tapping task together and are able to see each other's movements, it is expected that their movements will no longer be independent, but rather will become coupled to form a single dynamical system or coordinative structure. More specifically, given the previous research on the dynamics of interpersonal rhythmic coordination, we expected that Leaders' and Followers' movements would be coordinated in-phase (i.e., co-actors would perform the same movements at approximately the same time), but that, first, the stability of the observed coordination would be modulated by the amount of available information (less stability in partial than in full vision) and that, second, the degree to which Followers lagged behind

Leaders would be greater in the partial compared to the full vision condition.

Although it is often assumed that the two perspectives on joint action offer opposing explanations for how or why the patterns of joint action occur for various task manipulations or situations, we propose that they are more complementary than conflicting in that they predict different aspects to be affected by task asymmetries. More specifically, both perspectives propose a modification of participant behavior as a consequence of the asymmetries involved. The representational account predicts that the strategic behavioral adaptations of each individual's intentions and the representations of their co-actor's task (Knoblich et al. 2011; Vesper et al. 2010) will be altered by the task asymmetries. The dynamical systems account predicts that the patterning and stability of expected in-phase movement coordination will be altered by the task asymmetries. The complementarity of these predictions is that the former strategic adaptations predicted by a representational account are expected to interact with the latter in-phase coordination predicted by a dynamical systems account by altering the dynamic stabilities of the joint action task in the same way that any physical or informational constraint would (Coey et al. 2012; Schmidt and Richardson 2008). Thus, the two accounts are best conceptualized as complementary explanations of how joint action and interpersonal movement coordination come about. Accordingly, a secondary aim of the present study is to demonstrate how both representational and dynamical approaches play a role in joint action coordination.

## Experiment 1

Pairs of participants were instructed to perform a sequence of coordinated taps together. One participant in each pair was assigned the role of the *Leader* receiving information about the upcoming target location and one the role of the *Follower* without this information. There were three experimental conditions. In the *individual baseline* condition, which was always done first, both participants performed the tapping task independently without any requirement to coordinate their tapping with each other. In the *full vision* condition, the two participants were instructed to perform a sequence of taps while coordinating each tap such that they would touch the table surface on the same target at the same time. Mutual bidirectional vision was available such that each participant could see the other participant's complete tapping movements. In the *partial vision* condition, the same instructions were given to co-actors; however, less visual information was available because Leaders' movements were partially occluded. Followers could only see Leaders' movement start- and endpoints, but not their movements toward the target.

In line with our first hypothesis, we predicted that Leaders in the full vision condition would make use of the available visual modality to support Followers' task performance. In contrast, no such adaptations were expected in the partial vision condition in which Followers could only see a small proportion of Leaders' movements. Based on previous findings (e.g., Sacheli et al. 2013; Sartori et al. 2009), we expected Leaders to exaggerate their movements which should be reflected in higher movement amplitude in full versus partial vision. This should make the movements more salient and thereby make it easier for others to track them. In addition, we expected Leaders to provide Followers with specific task information by communicating the next target location through their movements, helping their partners to more easily perform the correct action (Pezzulo et al. 2013). This adaptation should be specific to the full vision condition in which Followers could make use of this information.

Based on our second hypothesis, we expected that co-actors' movements would be more synchronized in the full compared to the partial vision condition. Synchronization was measured as instantaneous relative phase (IRP) that captures spatiotemporal correspondences in roughly cyclic movement time series. We predicted the mean relative phase angle to be slightly negative indicating that Followers were coordinated in-phase with Leaders, but at a small phase lag. The degree of relative phase shift (degree to which Followers' movements lag behind Leaders) should be smaller with full compared to partial shared vision, indicating a higher proportion of in-phase synchrony between the two co-actors established by more bidirectional coupling in the full vision condition. The available visual information was also expected to increase the stability of coordination. Generally, the more stable interpersonal coordination is the lower is the variability [measured as standard deviation (SD)] of relative phase. Accordingly, we expected that the SD of IRP would be smaller in the full compared to the partial vision condition.

Finally, given that the co-actors' task was to tap at the same time onto the targets, we were also interested in overall coordination success. We expected that the absolute asynchrony between Leaders' and Followers' tapping endpoints (i.e., the moment of touching the table surface) would be smaller with full than with partial vision. Asynchrony and IRP were not analyzed in individual baseline because participants performed the tasks independently.

## Method

### Participants

Twelve women and 20 men of the University of Cincinnati participated in randomly matched pairs (two pairs with

only women, six pairs with only men). They were between 18 and 38 years old ( $M = 22.8$  years,  $SD = 5.2$  years) and right-handed. They received course credit as compensation. The study was approved by the local institutional review board.

### Apparatus and stimuli

Four target areas were displayed by a projector on each side of a custom-made interactive touch-screen table (Fig. 1). Target areas were blue circles ( $\emptyset 7.5$  cm; distance between targets 22 cm) with on either side a smaller red circle ( $\emptyset .5$  cm). To indicate which of the four blue big circles would be the subsequent target, the two small red circles next to it turned green. During the joint conditions, only the Leaders' side had indicators and these were covered by plastic semi-circles ( $\emptyset 1.5$  cm, height 1.5 cm) to prevent the Followers from receiving target information. Depending on the condition, black fabric occluders were set up between participants to restrict visual feedback (exact sizes are displayed in Fig. 1). In the partial vision condition, occluder height was adjusted flexibly to ensure that different participants could see the same details of their co-actor. For this adjustment, Followers indicated how much of the Leaders' movements they could see until a setting was found in which the finger hitting/leaving the target on the table could be seen, but no further part of the movement. During individual baseline performance, participants wore earmuffs.

Six sequences of 274 taps each were prepared in advance. The assignment of sequences to conditions was counterbalanced between participant pairs. All sequences contained a pseudorandom order of targets<sup>1</sup> that were made up of 34 chunks of 8 trials that were permutations of two times all four target locations (for example, one chunk could have the target order 1-3-2-1-3-4-2-4). Chunks were designed to contain no immediate repetition of the same target (e.g., 2-2), no complete series (e.g., 4-3-2-1) and no repetitions of two number patterns (e.g., 3-1-3-1). All six sequences ended with two final taps to random targets.

### Procedure

At the start of the experiment, one participant was randomly assigned the role of *Leader* and the other the role of *Follower*. The roles determined the table position and the amount of perceptual information that participants received in the joint conditions. Participants first independently performed two sequences in the *individual baseline*

<sup>1</sup> Three of the six sequences contained a repeating pattern of length 16; this had no effect on any of the measured variables and was therefore not included as a factor in the statistical analyses.

condition (Fig. 1a). In this condition, the task was identical for Leaders and Followers: Participants were instructed to tap as fast as possible onto the correct target highlighted by the indicators. As soon as the correct target was hit, the next one was highlighted without any delay so that task performance was self-paced. Participants performed the individual baseline condition at the same time in the same room, but they had no visual, auditory or haptic access to each other.

Following the individual baseline condition, participants performed the two joint conditions that differed in the amount of information that Followers received about Leaders' actions. In the *full vision* condition (Fig. 1b), both co-actors could see each other's complete movement trajectories (but not each other's upper body or face). In contrast, in the *partial vision* condition (Fig. 1c), Leaders could see Followers' complete movement trajectories, whereas Followers only saw the start- and end position of Leaders' movements. The task was the same in the two joint conditions. Participants were instructed to quickly tap onto the correct targets at the same time (i.e., hitting the table surface synchronously), while only Leaders would know where the next target would be. Task performance was self-paced by the Leader such that the next target was highlighted when the Leader tapped onto the correct target. Participants did not talk during task performance. The order of the two joint conditions was counterbalanced. Completing one sequence in any of the three conditions took between 4 and 6 min and there were short breaks between sequences. The total duration of the experiment was 45 min.

#### Data collection and analysis

Three-dimensional movement data were recorded with a Polhemus Liberty Latus magnetic motion tracking system with two wireless sensors sampled at a constant frequency of 188 Hz and a .01-cm spatial resolution. For each participant, the sensor was mounted with adhesive tape on the front of the outstretched index and middle finger. The unfiltered movement data were evaluated online to control the experimental procedure: Whenever a correct target hit was detected (in both participants in the individual baseline and in Leaders in the joint conditions; less than .75 cm from the table top), the subsequent target was highlighted. Stimulus presentation and data collection were programmed in C. For the offline analysis, target hits were extracted from the unfiltered raw time series by a semiautomatic MATLAB procedure because Followers' taps were sometimes not registered during the online analysis in the joint conditions as the next target was immediately triggered after the Leaders' taps. The extraction algorithm picked as the movement onset the first instance in the raw time series that fell within all of a combination of three criteria: (1) after a zero

crossing of the overall velocity, (2) within 2 SDs around the respective participant's mean horizontal position for the current target and (3) below 2 SDs of the respective participant's mean amplitude at target.

The data from the first eight and the last two taps were removed, and outlier taps were computed (values above or below the mean tap duration  $\pm 4$  SD). All taps that were outliers in the Leader or the Follower of each pair were removed (1.9 % of all trials), and the following individual and pair performance parameters were calculated: (1) *Tap duration* is the duration between the onset of one tap and the onset of the following tap. (2) *Velocity* is the mean absolute 3-dimensional velocity. (3) *Amplitude* is measured as the maximal vertical extension (with respect to the table top) between one tap and the next. (4) *Amplitude slope* is used to test whether movement amplitude would contain specific information about the target location. To that end, we split up the maximum amplitude values for each absolute target-to-target distance [i.e., first to third target (1-3) had the same absolute distance as 2-4 or 3-1, but not as 1-2 or 3-4] and from these values, regression slopes were computed that indicated how strongly the amplitude increased per distance step, i.e., whether larger absolute distances between the current and the next target go hand in hand with higher-amplitude movement trajectories. (5) *Velocity ratio* is the proportion of peak velocity before and after the maximum amplitude. Values close to 1 indicate a symmetric velocity profile and values larger than 1 mean higher peak velocity before than after the maximum movement amplitude. (6) *Asynchrony* captures pairs' discrete coordination performance, i.e., how well they followed the instruction to tap onto targets at the same time and is measured as absolute mean asynchrony between co-actors' taps. (7) *Instantaneous relative phase (IRP)* is used to investigate continuous coupling between co-actors' movements, i.e., how coordinated they were during target approach where coordination was not part of the instructions. It provides a continuous measure of whether two movement time series (such as Leaders' and Followers' movement trajectories) are synchronized in an in-phase manner and how stable that synchronization is. Essentially, IRP involves determining a relative phase angle for each corresponding pair of data points in the two movement time series (Pikovsky et al. 2001). In the present study, it was computed from co-actors' horizontal (=left/right) movements. From the resultant relative phase time series, circular statistics can be employed to calculate the mean relative phase angle that occurred between the movements over time. The closer the mean relative phase angle is to 0°, the closer the coordination is to perfect in-phase synchronization, with a slight negative or positive mean relative phase angle indicating a lag or lead relationship between the coordinated movements,

**Table 1** Means (and SDs) for all parameters of Experiments 1 and 2

	Individual baseline	Full vision	Partial vision
<i>Leaders (Exp. 1)</i>			
Tap duration (ms)	653.4 (68.6)	1,054.6 (184.2)	1,173.6 (191.9)
Velocity (ms/cm)	.4 (.04)	.25 (.04)	.22 (.04)
Amplitude (cm)	4.5 (1.1)	7.9 (4.6)	4.4 (1.3)
Amplitude slope	.15 (.17)	.85 (.77)	.52 (.57)
Velocity ratio	1.0 (.16)	1.26 (.37)	1.27 (.26)
<i>Followers (Exp. 1)</i>			
Tap duration (ms)	632.1 (96.4)	1,059.8 (185.7)	1,180.9 (189.2)
Velocity (ms/cm)	.42 (.07)	.29 (.06)	.26 (.04)
Amplitude (cm)	5.1 (1.6)	7.4 (4.0)	6.0 (2.7)
Amplitude slope	.37 (.41)	.19 (.43)	.06 (.19)
Velocity ratio	1.09 (.09)	1.08 (.22)	1.03 (.12)
<i>Pair (Exp. 1)</i>			
Asynchrony (ms)		103.6 (33.6)	155.4 (61.8)
IRP (°)		-10.86 (6.06)	-14.55 (6.63)
SD of IRP (°)		6.0 (7.7)	29.5 (7.8)
<i>Observers (Exp. 2)</i>			
IES (ms)	14.0 (3.9)	11.1 (2.5)	12.4 (2.1)
Accuracy (%)	56.8 (10.8)	70.1 (11.6)	66.9 (8.3)
Reaction time (ms)	761.7 (113.0)	757.5 (109.1)	822.5 (103.0)

respectively (see Schmidt and Richardson 2008, for more details). (8) *Variability of IRP (SD of IRP)* is the SD of relative phase around the mean relative phase angle and captures the stability of phase coupling between co-actors' movements.

The parameter values of the two sequences in each condition were averaged for the statistical analysis (performed with IBM SPSS 22). Leaders' and Followers' individual performance was analyzed separately with within-subjects ANOVAs with the single factor condition (individual baseline, full vision, partial vision). We report Greenhouse–Geisser corrected  $p$  values and, where appropriate, Bonferroni-corrected post hoc direct comparisons between conditions. The group performance parameters were analyzed with pair-wise  $t$  tests. All means and SDs of the results are reported in Table 1.

## Results

### Leaders' performance

Tap duration and velocity were significantly influenced by condition,  $F(2,30) = 79.23$ ,  $p < .001$ ,  $\eta_p^2 = .841$ , and  $F(2,30) = 230.49$ ,  $p < .001$ ,  $\eta_p^2 = .939$ . Leaders' taps were longer and slower in the partial than in the full vision condition, both  $p < .01$ , and in the joint conditions than in

the individual baseline, all  $p < .001$ . Leaders' movement amplitudes were used to investigate how much they would modulate their tap performance. As predicted, amplitude was significantly affected by condition,  $F(2,30) = 11.18$ ,  $p < .01$ ,  $\eta_p^2 = .427$  (Fig. 2a). Movements were performed with higher maximum amplitude under full compared to partial vision,  $p < .01$ , and compared to individual baseline performance,  $p < .05$ . Amplitude was not different in the partial vision and individual baseline conditions,  $p = 1.0$ . There was a significant main effect of condition for Leaders' amplitude slope,  $F(2,30) = 11.35$ ,  $p < .001$ ,  $\eta_p^2 = .431$  (Fig. 2c). The slope in the full vision condition,  $p < .01$ , and the partial vision condition,  $p < .05$ , were significantly steeper than in the individual baseline. The amplitude slope in the partial vision condition was descriptively lower than in the full vision condition, but statistically, this was only marginally significant,  $p = .061$ . The velocity ratio was affected by condition,  $F(2,30) = 13.53$ ,  $p < .01$ ,  $\eta_p^2 = .474$ , such that the joint conditions had relatively higher peak velocities before the maximum amplitude was reached than the individual baseline condition, both  $p < .05$ . Full and partial vision were not significantly different,  $p = 1.0$ .

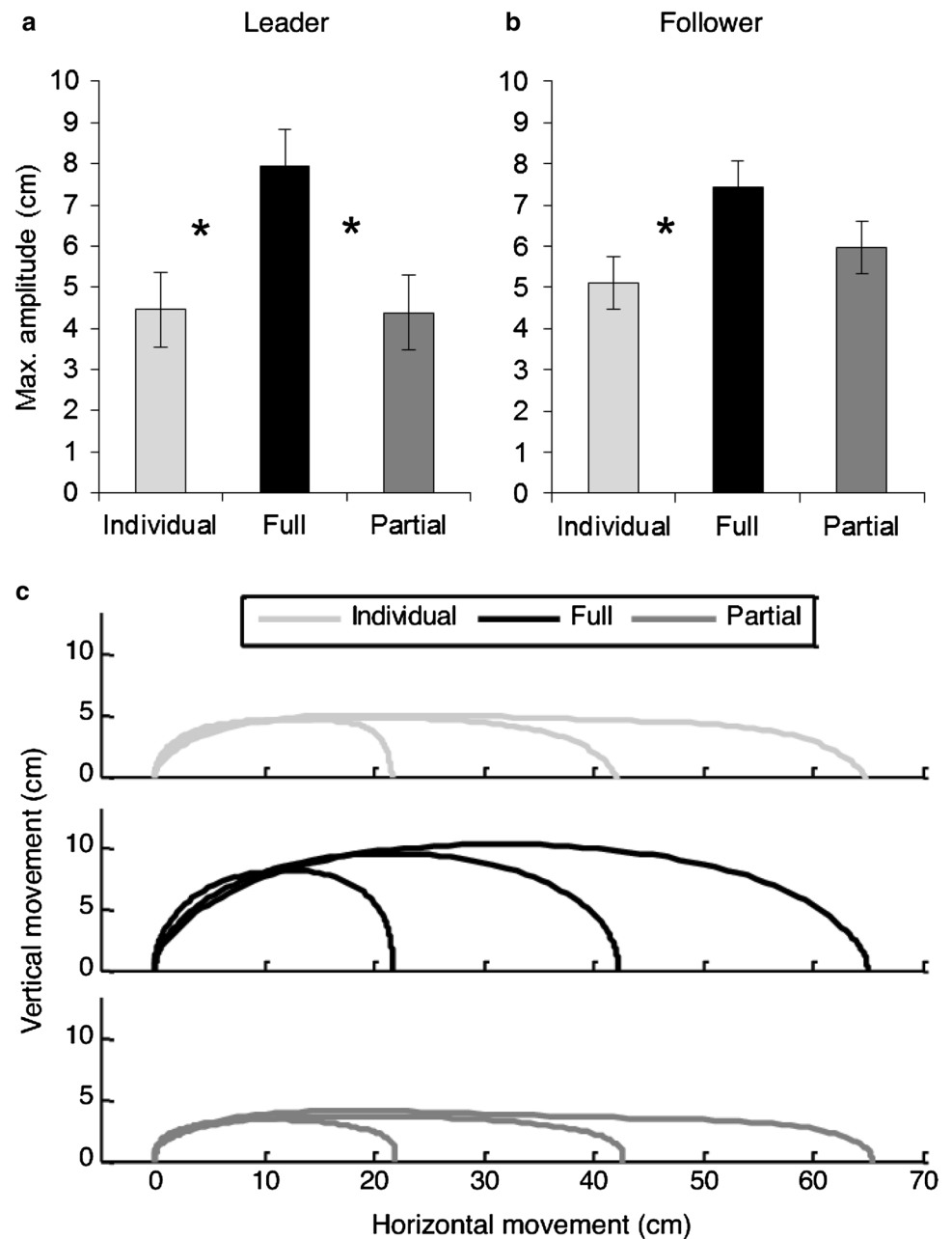
### Followers' performance

Tap duration and mean velocity were affected by condition,  $F(2,30) = 88.64$ ,  $p < .001$ ,  $\eta_p^2 = .855$ , and  $F(2,30) = 69.72$ ,  $p < .001$ ,  $\eta_p^2 = .823$ . Followers' taps were longer and slower in the partial compared to the full vision condition, both  $p < .01$ . Taps in the joint conditions were performed with longer duration and more slowly than in the individual baseline, all  $p < .001$ . For Followers' movement amplitude, there was a main effect of condition,  $F(2,30) = 7.47$ ,  $p < .01$ ,  $\eta_p^2 = .333$  (Fig. 2b). Movements in the full vision condition had significantly higher maximum amplitude than in the individual baseline,  $p < .05$ , but not than in the partial vision condition,  $p > .07$ . Movements performed under partial vision were not significantly different from the individual baseline,  $p > .2$ . Followers' amplitude slope was also influenced by condition,  $F(2,30) = 4.86$ ,  $p < .05$ ,  $\eta_p^2 = .245$ . However, in contrast to the results found for Leaders, slopes in the full vision condition were not significantly different from slopes in the partial vision or individual baseline conditions, both  $p > .4$ . Partial vision was significantly less steep than individual baseline,  $p < .01$ . The velocity ratio was not affected by condition,  $F(2,30) = .61$ ,  $p > .5$ ,  $\eta_p^2 = .039$ .

### Pair performance

As expected, coordination was better, i.e., absolute asynchrony was lower, when co-actors could see each other's

**Fig. 2** **a–b** Leaders' and Followers' maximum movement amplitude (*error bars* within-subjects confidence intervals according to Loftus and Masson 1994; significances are marked with \*). **c** For illustration, one representative example Leader's time- and position-normalized average trajectories show how movement amplitude was specifically adapted in the full vision condition

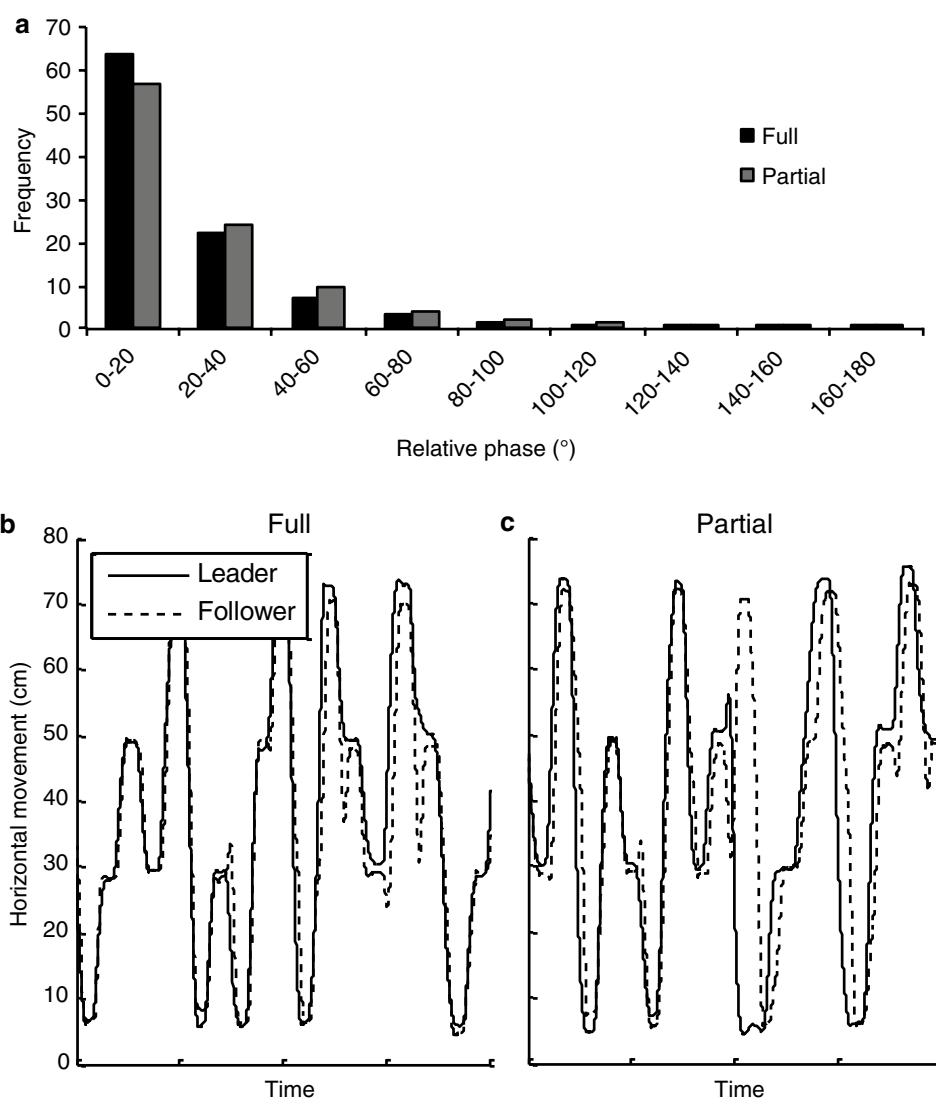


movements (full vision condition) than when Followers could only partially see Leaders' movements (partial vision condition),  $t(15) = -4.17$ ,  $p < .01$ . Moreover, consistent with our predictions, the mean relative phase angle of the two co-actors' movements was closer to  $0^\circ$  in the full vision condition compared to the partial vision condition,  $t(15) = 3.49$ ,  $p < .01$ , indicating that Followers continuously lagged further behind Leaders with partial vision (Fig. 3). The variability of IRP was also significantly influenced by condition,  $t(15) = -4.86$ ,  $p < .001$ , such that movement synchrony was less variable (more stable) in the full vision compared to the partial vision condition.

## Discussion

In line with our first main hypothesis, the current results show that Leaders in the full vision condition provided support for Followers in performing the joint task. Leaders exaggerated their movements by increasing movement amplitude in the full compared to the partial vision condition. This exaggeration of the amplitude made their movements more salient which could have facilitated tracking them for the partner (Vesper et al. 2010). As predicted, Leaders' amplitude adaptation depended on whether Leaders were seen by Followers such that movement amplitude was increased under full vision, but not under partial vision

**Fig. 3** **a** Frequency distribution of IRP angles. The higher the frequency of phase angles close to  $0^\circ$  the more in-phase two time series are. **b** Two example trajectories from the same participant pair showing differences in spatiotemporal overlap in the two joint conditions



or the individual baseline. A similar specificity of information supply is known from studies on verbal conversation showing that interaction partners keep track of each other's knowledge and accordingly adapt what they say (Brennan and Hanna 2009; Lockridge and Brennan 2002). Similarly, exaggerating one's movements during joint tapping can have a communicative function, e.g., by emphasizing important parameters of one's own movement performance like current position, movement direction or velocity (Pezzulo and Dindo 2011; Pezzulo et al. 2013; Sacheli et al. 2013).

The fact that Leaders' movements in the full vision condition had both higher amplitude and were performed faster (shorter duration, higher velocity) than in the partial vision condition confirms that Leaders put in extra effort to support Followers, i.e., their movement adaptations might have created a cost that they were willing to expend (Pezzulo et al. 2013). Moreover, during joint performance, Leaders

had higher peak velocity before reaching the maximum amplitude which could be interpreted as an attempt to provide the critical amplitude information early during the movement. Alternatively, it is also possible that this generally served to increase the amount spent in target approach (the phase after maximum movement amplitude) to support Followers with tracking Leaders' movements.

In addition to increasing their tapping amplitude generally, Leaders provided more specific information about the upcoming target location to Followers. In particular, their movement amplitude contained information about the target location such that amplitude was higher when the next target was relatively far than when the next target was relatively close as indicated by steeper amplitude slopes. Again, this adaptation was strongest when Followers could use the information and much weaker in the partial vision or the individual baseline conditions. It is possible that Leaders also attempted to adapt their movement



performance to support Followers in the partial vision condition, e.g., by moving close to or below the occluder. However, the current results do not provide conclusive evidence for such effort given that amplitudes in partial vision and individual baseline were not different from each other.

Our second hypothesis that co-actors' movement synchronization, measured in terms of IRP, would be higher and more stable in full compared to partial vision was also supported by Experiment 1, indicating closer continuous in-phase synchrony between Leaders and Followers. Thus, when perceptual information was available, co-actors' movements were not only better coordinated at the endpoints (the actual tapping), but also continuously during movement execution. In other words, Leaders' and Followers' movements were informationally coupled based on the continuous exchange of perceptual information between them. This is consistent with earlier findings on interpersonal coordination from a dynamical systems perspective (Schmidt and Richardson 2008; Richardson et al. 2007).

## Experiment 2

A second experiment tested the hypothesis that the strategic amplitude modulation (signaling) found in Experiment 1 contains additional information that others can acquire and which is therefore helpful for coordination by allowing more efficient target prediction for a potential joint action partner (Pezzulo et al. 2013). To this end, a new group of participants (*Observers*) observed partly occluded trajectories that were recorded from Leaders of Experiment 1. Specifically, the beginning of the trajectory until the point of the maximum amplitude toward target approach was shown. Note that this partial trajectory was presented as a static picture so that Observers needed to use the information contained in the shape of the partial trajectory to predict the correct target. Observers' task was to decide as fast as possible which of the possible target locations (excluding the starting location) would most likely be the final position of the trajectory. They received immediate feedback about the accuracy of their choice.

We expected that Leaders' coordination intention in Experiment 1 as well as the availability of visual information to them would be reflected in Leaders' movements and that Observers in Experiment 2 would be able to pick up this information from the displayed trajectories (Becchio et al. 2010). Thus, we hypothesized that Observers would be more efficient at predicting the final target location of movements recorded from the full vision condition than of movements from the partial vision or individual baseline conditions of Experiment 1.

## Method

### Participants

Nine women and seven men from Central European University participated individually. They were between 18 and 24 years old (20.4 years,  $SD = 1.9$  years) and mostly right-handed (two left-handed). Each participant in Experiment 2 was paired with a participant from Experiment 1 such that the data generated by the participant in Experiment 1 was used as a stimulus for the participant in Experiment 2. A small monetary compensation for participation was provided.

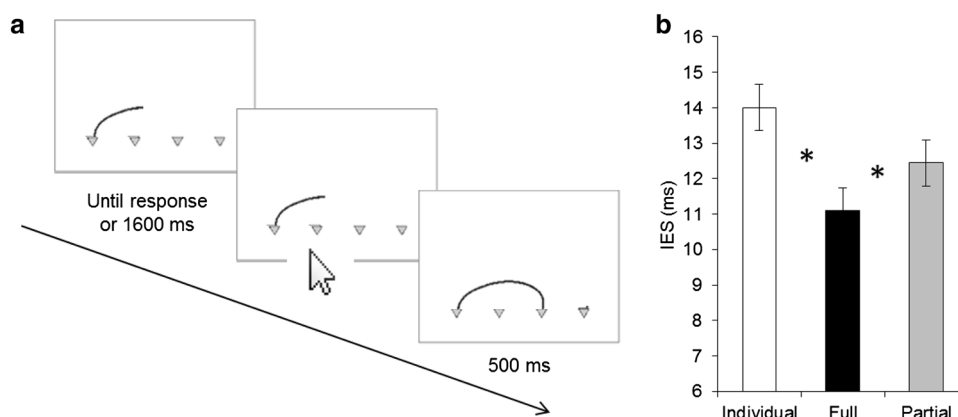
### Apparatus and stimuli

Participants were seated in front of a computer screen (resolution  $1,280 \times 1,024$  px, refresh rate 60 Hz). The movement data recorded from the 16 *Leaders* in Experiment 1 were used as stimuli for *Observers* in Experiment 2. Specifically, in each trial, the horizontal (left–right) and vertical (up–down) dimensions of one tapping movement from the time series of a particular Leader from Experiment 1 were shown (Fig. 4a). Each Observer in Experiment 2 received error-free data from a Leader of Experiment 1. Trials were presented in the same order in which the movements were originally performed, but each sequence was truncated to 256 trials by removing trials in the end. Thus, each Observer in Experiment 2 performed six sequences of 256 trials from one particularly matched Leader from Experiment 1. In the *individual baseline* condition, all trials were taken from the individual baseline condition of Experiment 1 and in the *full vision* and the *partial vision* conditions they came from the full vision and partial vision conditions of Experiment 1, respectively. The partial movement trajectories were displayed as static red lines and showed the start location to the point of highest amplitude (horizontal extensions at amplitude were not significantly different between the three conditions). As feedback, the point of highest amplitude to the movement end location was shown as a blue static line. Four light gray triangles pointing downwards represented the possible target locations. Participants chose one of them in each trial by clicking into the target area with a standard computer mouse. The experimental procedure was controlled using MATLAB 2012a.

### Procedure

Individual participants (*Observers*) were instructed that their task would be to “read” another person's movements and that this other person had participated in an earlier experiment. They would see only a part of the movements

**Fig. 4** Procedure and results of Experiment 2. **a** The order and timing of events during each trial with an example trajectory. **b** Results of the IES (error bars within-subjects confidence intervals according to Loftus and Masson 1994; significances are marked with \*)



recorded from this other person and were to indicate in each trial where the movements would end by clicking onto the respective target. These instructions emphasized the fact that the observed movements came from real human participants and that Observers would make errors when choosing the targets. The task was to try to be as accurate as possible given the highly variable data.

In each trial, the first part of the movement trajectory was shown until Observers responded with a mouse click onto one of the target locations or until a time-out of 1,600 ms (Fig. 4a). Whenever no response was given or the mouse click occurred outside of a target area, a short error tone was played (400 Hz, 150 ms). Finally, the full movement trajectory was shown for 500 ms as feedback before the next trial started. Observers first performed a short training with 20 trials taken from a different Leader of Experiment 1 and then performed six blocks corresponding to three experimental conditions (individual baseline, full vision, partial vision) in random order. They could take short breaks between the blocks. The total duration of the experiment was 60 min.

#### Data analysis

Three parameters were calculated to analyze Observers' prediction ability. (1) *Accuracy* was measured as the percentage of correctly predicted targets in trials in which a response was given (on average, no response was given in .6 % of the trials). (2) *Reaction time* was calculated as the time from start of a trial, i.e., when the visual display with the targets and the partial trajectory was shown, until a response (mouse click) was given and included all correct and incorrect trials. (3) *Inverse efficiency score (IES)* was calculated by dividing reaction time by accuracy to get an overall measure of performance that is independent of potential differences in individual participants' speed/accuracy criterion (Townsend and Ashby 1983). Smaller IES values indicate better prediction. All three parameters

were analyzed with ANOVAs like in Experiment 1, and the means and SDs are also reported in Table 1.

#### Results

As expected, Observers' IES (Fig. 4b) was significantly influenced by condition,  $F(2,30) = 11.15$ ,  $p < .001$ ,  $\eta_p^2 = .426$ . Prediction of the target location was more efficient (i.e., the IES was smaller) in the full compared to the partial vision,  $p < .05$ , and individual baseline conditions,  $p < .01$ . Partial vision and individual baseline conditions were not significantly different from each other,  $p > .1$ . Looking at the component parts of the IES revealed, first, that prediction accuracy was significantly influenced by condition,  $F(2,30) = 15.38$ ,  $p < .001$ ,  $\eta_p^2 = .506$ , such that Observers were better in the joint conditions compared to the individual baseline, both  $p < .01$ . However, accuracy in full and partial vision conditions did not differ significantly from each other,  $p > .4$ . Second, reaction time was affected by condition,  $F(2,30) = 17.2$ ,  $p < .001$ ,  $\eta_p^2 = .534$ , such that Observers were significantly slower in the partial compared to the full vision and individual baseline conditions, both  $p < .001$ . Full vision and individual baseline conditions were equally fast,  $p = 1.0$ . Thus, the effects of efficiency are mainly driven by a difference in accuracy between full vision and individual baseline and by a difference in reaction time between full vision and partial vision.

#### Discussion

Experiment 2 tested the hypothesis that Leaders' strategic amplitude modulation in Experiment 1 supports coordination by allowing another person observing the movements to better predict the targets of those movements. This hypothesis was confirmed. Observers in Experiment 2 were more efficient (taking accuracy and reaction time into account) in predicting Leaders' movements when these came from the full vision condition of Experiment 1 than

if they came from the partial vision or individual baseline conditions. These results suggest that, first, the coordination intention of Leaders in Experiment 1 as well as the availability of visual information between Leaders and Followers was reflected in Leaders' movements and that, second, Observers in Experiment 2 were able to pick up this information from the displayed trajectories. This is consistent with empirical evidence showing that Observers can distinguish different movement goals and social intentions in a performing actor (Becchio et al. 2010; Graf et al. 2007; Sartori et al. 2009).

## Conclusion

The present study aimed at understanding how co-actors achieve joint action coordination when task knowledge is distributed asymmetrically. Based on both representational and dynamical systems perspectives on joint action, we hypothesized different aspects of a joint tapping task with asymmetric knowledge about the tapping targets to be affected by this asymmetry—first, that Leaders having more task information would adapt their movement performance to assist Followers in finding the correct target location and second, that the stability and patterning of continuous synchrony, that would occur between the two co-actors' movements as an effect of monitoring each other's actions, would be modulated by the asymmetries in information and the co-actors' intention.

The present results support these hypotheses. On the one hand, Experiment 1 showed that when full visual information was available, Leaders substantially changed their movement performance, in particular the amplitude of their tapping. Amplitude was higher than at baseline or under partial vision and contained information about the location of subsequent targets. Experiment 2 further demonstrated that Leaders' movement adaptation allowed independent Observers to better predict the goal of a partially shown trajectory. These findings are in line with strategic and signaling approaches to joint action that suggest that co-actors' coordination intentions influence their action performance (Knoblich et al. 2011; Pezzulo et al. 2013; Vesper et al. 2010). On the other hand, Experiment 1 demonstrated how co-actors' movements became synchronized when mutual visual information was available such that their movements were performed in an in-phase manner. In line with a behavioral dynamics perspective on joint action and social movement coordination (Coey et al. 2012; Marsh et al. 2009; Schmidt and Richardson 2008), the in-phase synchronization was also stronger with full compared to partial visual information between Leaders and Followers. Furthermore, a greater phase lag was observed (with Followers further behind Leaders) when only partial vision was

available. That is, not only did the information manipulation influence the stability of the interpersonal coordination (i.e., IRP variability), but also the patterning of the coordination (i.e., mean IRP). An open question at this point is to what degree the increase in coordination stability was due to the amount of visual information per se or due to Leaders' strategic movement adaptations (or a combination of both). Although the results of Experiment 2 suggest that Leaders' movements provided visual information about a subsequent movement target, thereby improving others' prediction efficiency, the influence of this information on interpersonal movement synchronization needs to be further investigated to fully tease this issue apart.

Taken together, the current findings suggest that fully understanding joint action tasks requires both representational and dynamical systems accounts. In the present study, co-actors' action performance was modulated by their intention to perform actions together, but their movements were also coupled via basic dynamic principles. Thus, the two accounts together may well offer a more coherent explanation of joint action and interpersonal movement coordination that target different levels of the interaction. Future work could address situations in which these different levels create a conflict, e.g., if a joint task explicitly requires *not to be continuously synchronized*. All in all, the present study provides important insights about the principles underlying and influencing complex, multi-level joint action performance in real time.

**Acknowledgments** We thank Veronica Romero for her help with data collection and Günther Knoblich for valuable comments. Michael Richardson's effort on this project was partially supported by the National Institutes of Health under Award Number R01GM105045. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

- Becchio C, Sartori L, Castiello U (2010) Toward you: the social side of actions. *Curr Direct Psychol Sci* 19:183–188
- Brennan SE, Hanna JE (2009) Partner-specific adaptation in dialog. *Top Cognit Sci* 1:274–291
- Brennan SE, Chen X, Dickinson CA, Neider MB, Zelinsky GJ (2008) Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition* 106:1465–1477
- Clark HH (1996) Using language. Cambridge University Press, Cambridge
- Coey CA, Varlet M, Richardson MJ (2012) Coordination dynamics in a socially situated nervous system. *Front Hum Neurosci* 6:1–12
- de Bruijn ERA, de Lange FP, von Cramon DY, Ullsperger M (2009) When errors are rewarding. *J Neurosci* 29:12183–12186
- Goebel W, Palmer C (2009) Synchronization of timing and motion among performing musicians. *Music Percept* 26:427–438
- Graf M, Reitzner B, Corves C, Casile A, Giese M, Prinz W (2007) Predicting point-light actions in real-time. *Neuroimage* 36:T22–T32

- Keller PE (2008) Joint action in music performance. In: Morganti F, Carassa A, Riva G (eds) *Enacting intersubjectivity: a cognitive and social perspective on the study of interactions*. IOS Press, Amsterdam, pp 205–221
- Knoblich G, Butterfill S, Sebanz N (2011) Psychological research on joint action: theory and data. In: Ross B (ed) *The psychology of learning and motivation*, vol 54. Academic Press, Burlington, pp 59–101
- Lockridge C, Brennan S (2002) Addressees' needs influence speakers' early syntactic choices. *Psychol Bull Rev* 9:550–557
- Loehr J, Kourtis D, Vesper C, Sebanz N, Knoblich G (2013) Monitoring individual and joint action outcomes in duet music performance. *J Cognit Neurosci* 25:1049–1061
- Loftus GR, Masson MEJ (1994) Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1:476–490
- Marsh KL, Richardson MJ, Schmidt RC (2009) Social connection through joint action and interpersonal coordination. *Top Cognit Sci* 1:320–339
- Pezzulo G, Dindo H (2011) What should I do next? Using shared representations to solve interaction problems. *Exp Brain Res* 211:613–630
- Pezzulo G, Donnarumma F, Dindo H (2013) Human sensorimotor communication: a theory of signaling in online social interaction. *PLoS One* 8:e79876
- Pikovsky A, Rosenblum M, Kurths J (2001) *Synchronization: a universal concept in nonlinear science*. Cambridge University Press, New York
- Richardson MJ, Marsh KL, Isenhower R, Goodman J, Schmidt RC (2007) Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Hum Mov Sci* 26:867–891
- Richardson MJ, Harrison SJ, May R, Kallen RW, Schmidt RC (2011) Self-organized complementary coordination: dynamics of an interpersonal collision-avoidance task. *BIO Web Conf* 1:00075
- Sacheli L, Tidoni E, Pavone E, Aglioti S, Candidi M (2013) Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Exp Brain Res* 226:473–486
- Sartori L, Becchio C, Bara BG, Castiello U (2009) Does the intention to communicate affect action kinematics? *Conscious Cognit* 18:766–772
- Schmidt RC, Richardson MJ (2008) Dynamics of interpersonal coordination. In: Fuchs A, Jirsa VK (eds) *Coordination: neural, behavioral and social dynamics*. Springer, Berlin, pp 281–308
- Schmidt RC, Carello C, Turvey MT (1990) Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *J Exp Psychol Hum Percept Perform* 16:227–247
- Schmidt RC, Fitzpatrick P, Caron R, Mergeche J (2011) Understanding social motor coordination. *Hum Mov Sci* 30:834–845
- Sebanz N, Knoblich G, Prinz W (2003) Representing others' actions: just like one's own? *Cognition* 88:B11–B21
- Townsend JT, Ashby FG (1983) *Stochastic modelling of elementary psychological processes*. Cambridge University Press, New York
- Tsai JCC, Sebanz N, Knoblich G (2011) The GROOP effect: groups mimic group actions. *Cognition* 118:135–140
- Vesper C, Butterfill S, Knoblich G, Sebanz N (2010) A minimal architecture for joint action. *Neural Netw* 23:998–1003
- Vesper C, van der Wel RPRD, Knoblich G, Sebanz N (2011) Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Exp Brain Res* 211:517–530
- Vesper C, van der Wel RPRD, Knoblich G, Sebanz N (2013) Are you ready to jump? Predictive mechanisms in interpersonal coordination. *J Exp Psychol Hum Percept Perform* 39:48–61
- Wolpert DM, Doya K, Kawato M (2003) A unifying computational framework for motor control and interaction. *Philos Trans R Soc Lond B* 358:593–602