Scaling Up Perception–Action Links: Evidence From Synchronization With Individual and Joint Action

Verónica C. Ramenzoni
Radboud University, Nijmegen, The Netherlands

Natalie Sebanz and Guenther Knoblich
Radboud University, Nijmegen, The Netherlands and Central European University, Budapest, Hungary

How do we map joint actions we participate in onto joint actions we observe others performing, such as when a couple dancing tango observes another couple dancing tango? We investigated this question using a task in which participants were instructed to perform individual or joint movements in synchrony with individual or joint movements observed on a computer screen. The observed movements started slowly and then continuously increased in tempo (from 1.75 Hz to 3 Hz). The results showed that, with regard to spatial parameters, joint performance was more accurate when observing joint action than when observing individual action (Experiments 1, 1a, and 1b). Individual performance was more accurate when observing individual action than when observing joint action (Experiments 3 and 4). There were no systematic differences with regard to timing parameters. These results suggest that mechanisms of temporal coordination may be less susceptible to differences between individual and joint action than mechanisms of spatial matching.

Keywords: joint action, synchronization, action imitation, predictive coding

Interactions between groups have an important place in our lives. They are the unit around which a large proportion of our cultural practices are organized, and they are at the core of numerous forms of dancing and sports. Interactions between groups are different from dyadic joint actions, and some aspects of them cannot be reduced to one-to-one interactions. In particular, learning to perform joint actions may require interacting with other groups; ballroom dancers, for instance, learn to dance not only by watching the movements of solo dancers but also by imitating other couples' moves. This likely requires them to pick up on the dancing configurations the observed couple achieves and on how the two dancers coordinate their joint behavior over time. The aim of this article is to investigate how our cognitive system goes about mapping joint actions we participate in onto joint actions we observe others performing. Specifically, we were interested in how perception–action links allow individuals to imitate joint actions respecting both their temporal and spatial structure.

Several mechanisms have been proposed to account for how the cognitive system goes about the continuous dynamic integration of perceptual inputs and motor outputs (for a discussion, see Spivey, 2007). Dynamic approaches propose that perception–action links emerge between interacting individuals in time, as changes in their behavior recur on each other. Representational approaches argue that perception and action are linked directly by generating internal action models that map temporal and spatial parameters of observed actions onto the action system of the perceiver. In the sections below, we discuss the significance of these mechanisms for the study of joint imitation.

Coordination Dynamics

One approach inspired by self-organization theory focuses on the similarity in the timing and variability of online actions across organisms acting together. In this view, perception–action links can emerge between interacting individuals when their behaviors become coordinated in time. Joint behavior between human organisms is conceived as producing interpersonal coordination (also referred to as entrainment, interpersonal synchrony, or interpersonal synchronization), which is regulated by the same coordinative dynamic principles as interactions in a large spectrum of living systems. Interpersonal coordination has been observed for a multiplicity of behaviors, including gestures and mannerisms (Chartrand & Bargh, 1999), facial expressions (Meltzoff & Moore, 1977), and overt limb movements (e.g., Amazeen, Schmidt, & Turvey, 1995; Oullier, de Guzman, Jantzen, & Kelso, 2003; Richardson, Marsh, & Schmidt, 2005; Schmidt, Bienvenu, Fitzpatrick, & Amazeen, 1998; Schmidt, Carello, & Turvey, 1990; Schmidt & Turvey, 1994). Tendencies to coordinate have been found at different time scales, spanning fast milliseconds-long movements, such as postural sway (Shockley, Baker, Richardson, & Fowler, 2007; Shockley, Santana, & Fowler, 2003), to hour-long conversation cycles (Hayes & Cobb, 1982).

Stable interpersonal coordination can emerge spontaneously between two people regardless of whether they make an explicit attempt to coordinate their behavior or share a preexisting common goal (Schmidt & O'Brien, 1997). In a broad sense, spontaneous interpersonal synchronization depends on the availability of visual information about the other person's behavior, specifically, information in the optic array that informs interacting participants about both large and subtle *changes* in behavior (e.g., Schmidt et al., 1998; Schmidt, Carello, & Turvey, 1990). Coordination-dynamics approaches have traditionally downplayed the role of internal representational processes in regulating interactions. In addition, spatial or configurational aspects of the observed actions—which remain *constant* throughout the interaction, such as the spatial configuration of limbs—are thought to play a lesser role in driving interpersonal coordination dynamics. Thus, it is an open question whether task demands that specify the particular aspects of a joint action can affect temporal aspects of performance, such as when imitating multiple actions originating in one or more individuals. The expectation is that insofar as the timing of the actions is the same, the context within which those actions are framed would not change how individuals go about synchronizing to it.

Perception–Action Mapping Models

Representational approaches to joint action emphasize the cognitive capabilities required for successful interaction (i.e., action prediction and understanding). These approaches argue that perception and action are linked directly by generating action models (Haggard, 2005; Wilson & Knoblich, 2005) that follow the same principles as forward models used in motor control (Wolpert & Flanagan, 2009). The proposal is that individuals use their own action system to simulate online the actions they observe in others (for a review, see Pezzulo, Candidi, Dindo, & Barca, 2013) and predict their outcomes (Clark, 2013; Oztop, Wolpert, & Kawato, 2005; Wolpert, Doya, & Kawato, 2003). These models are not fixed or encapsulated. They are constantly updated by experience, they can be used to plan future actions depending on the individual's goals, and they are open to constraints imposed by task demands (Wolpert & Flanagan, 2009).

A growing corpus of behavioral findings supports this proposal, among them evidence showing that perceiving an action can sometimes facilitate, and sometimes interfere with, the performance of concurrent actions (see, e.g., Kilner, Paulignan, & Blakemore, 2003). Internal models are very well suited to explain how we map discrete as well as continuous actions. For the imitation of discrete actions, participants' actions are faster when they are asked to perform actions that are similar to observed actions in terms of movement parameters and/or action goals. Mapping during imitation is regulated by several discrete properties of the observed actions, including the direction and type of the movement (Brass, Bekkering, Wohlschläger, & Prinz, 2000; Brass, Derrfuss, & von Cramon, 2005; Brass, Zysset, & von Cramon, 2001), the effector selected (Belopolsky, Olivers, & Theeuwes, 2008; Stürmer, Aschersleben, & Prinz, 2000), and the characteristics of the motion (i.e., biological vs. nonbiological; Kilner, Hamilton, & Blakemore, 2007; but see Stanley, Gowen, & Miall, 2007).

When mapping continuous events, internal forward models are used to predict events attending to their temporal structure and to imitate the timing and kinematics of observed actions. For instance, studies exploring how the timing of actions affects how they are perceived (cf., Knoblich & Repp, 2009; Repp & Knoblich, 2007) have found that participants can differentiate self-produced versus others' actions based on their temporal structure (i.e., tapping frequency). Internally generated information can also guide action planning and lead to coordinated joint actions in situations in which the timing of the actions is critical for performance (Vesper, van der Wel, Knoblich, & Sebanz, 2013). Moreover, as in the imitation of discrete events, the degree of similarity in the temporal dynamics between perception and production also modulate processing, with greater isomorphism in temporal parameters yielding better action predictions (Flach, Knoblich, & Prinz, 2003; Keller, Knoblich, & Repp, 2007).

Contextual and situational factors also affect mapping in imitation tasks. For instance, when the spatial properties of observed movements and the goals of the observed action are manipulated simultaneously, their effect on imitation depends on what features of the action the participant attends to (Bird, Brindley, Leighton, & Heyes, 2007; Franz, Ford, & Werner, 2007; Knuf, Aschersleben, & Prinz, 2001; Massen & Prinz, 2007, 2009). Perception–action mapping mechanisms allow participants to map more abstract features of actions, such as their goals and the task demands within which they are produced (Wohlschläger, Gattis, & Bekkering, 2003). Mapping mechanisms can then facilitate action selection, thus aiding the system in activating relevant motor commands, depending on the goals and context of the task (Blakemore & Frith, 2005). This implies that mapping mechanisms could operate at a higher order categorical level to predict properties of actions that are not necessarily time locked (Frith & Frith, 2006). One strategy the cognitive system can use to accomplish this is generating different models depending on the task. Tasks that present individual and joint demands would depend on the activation of individual versus joint models, which could result in differences in how motor commands are executed. Research on joint imitation, in which observed and performed actions differ in terms of the *number of actors* involved, but not necessarily in the *number of actions* to imitate, has provided evidence in support of this strategy.

Tsai, Sebanz, and Knoblich (2011) measured RTs when individuals acting alone or together were asked to imitate individually performed or jointly performed actions. The congruency between the number of perceived actors and the number of actors performing the task led to a modulation of imitative performance. Dyads were faster in imitating the unimanual actions of two individuals, and individuals were faster in imitating the bimanual actions of one individual, even when these actions looked exactly the same. The findings in the dyadic condition can only be explained if one assumes that, for jointly acting individuals, the consequences of jointly produced actions have a higher priority in driving each individual's actions than the activation of representations that link each individual's actions to its consequences. In what follows, we refer to representations of joint action outcomes as "we-representations" and to representations of individual action outcomes as "me-representations."

Tsai et al.'s (2011) findings indicate a preference to perform group actions when observing group actions, and a preference to

perform individual action when observing individual actions. This suggests that individuals might activate a joint model to map actions when the task is represented as a joint task, and an individual model when the task is represented as an individual task. In Tsai et al.'s task, however, each observed movement was treated as a discrete event. Group actions typically involve multiple events, originating in more than one individual, that are perceived as related to one another, because of anatomical reason, as in Tsai et al.'s case, or because they are linked in time. It is an open question whether joint models can modulate imitation of continuous real-time performance in which both members of the observed group are continuously acting. In other words, can we-representations help participants to integrate online the observed performance of multiple actions, so that observing joint actions will lead to better imitation of the temporal and spatial structure of events than observing individual actions?

Current Study

For the imitation of joint actions, the spatiotemporal aspects of the task are important in at least three respects. Participants should be able to pick up on how the group they are trying to imitate coordinates within itself ("What are they doing and when?"), they should also be sensitive as to their own role in the interaction, and they should attempt to follow the behaviors that are relevant to them individually ("What am I doing and when?"). Finally, participants should sustain a successful interaction with their partners ("What are we doing and when?"; see Figure 1, Panel A).

For instance, learning to dance a tango as a leader involves not only following the other couple's moves but also making sure that one is imitating how the leader in the couple guides a follower, while simultaneously attending to what one's own partner is doing. All these different relations rely on perception–action links that
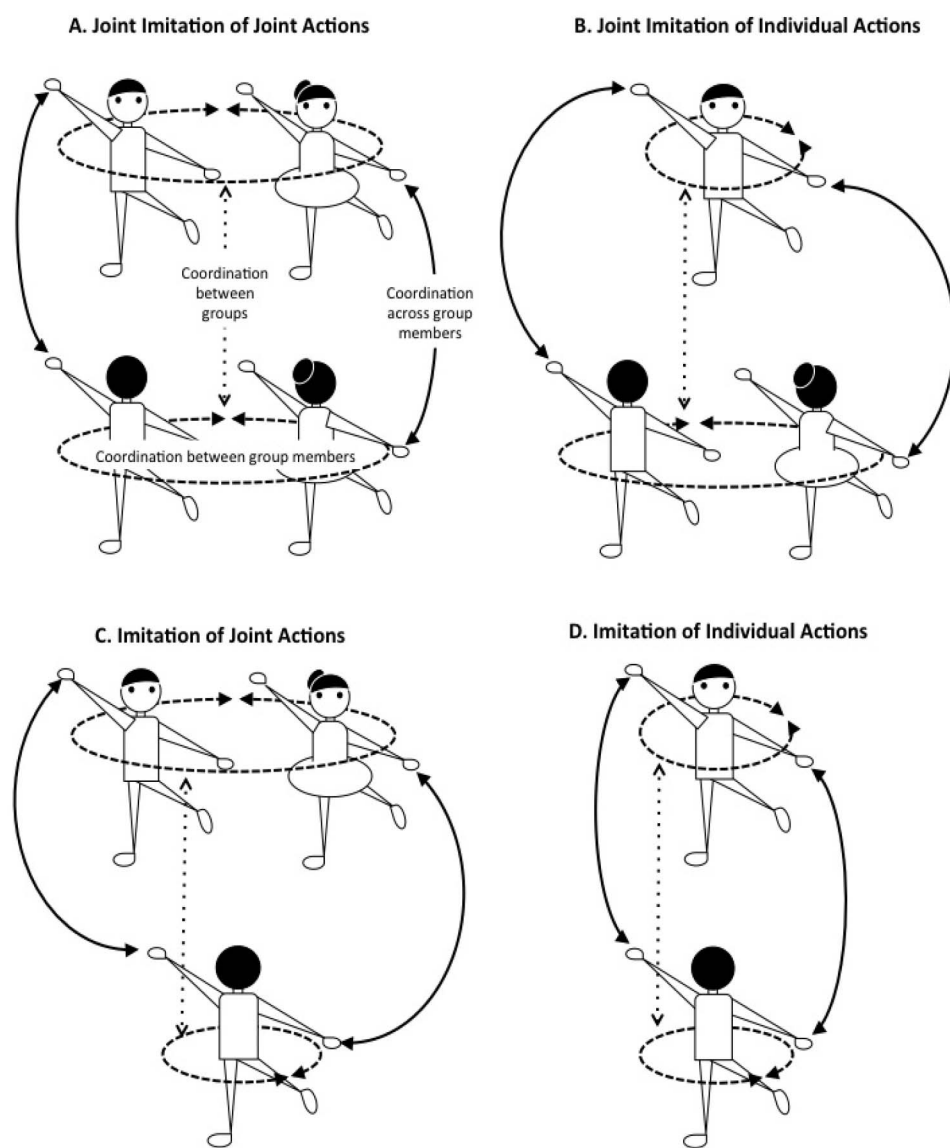

Figure 1. Illustration of the different types of imitation explored in this study. Panel A includes labels for different kinds of coordinative relations involved in joint imitation.

have to be sustained in time. In the present research, we focused on one of these aspects: how each participant in a dyad synchronizes her actions with those of the member of the group he or she is attempting to imitate.

We used a continuous finger-tapping task, a widely used paradigm for the study of synchronization (Luck & Sloboda, 2009; Repp & Keller, 2008; Vorberg & Schulze, 2002). In the typical finger-tapping paradigm, an individual is asked to tap every time she hears or sees a stimulus, most frequently produced by a metronome. We adapted this paradigm for the present purposes by asking participants to tap in synchrony with an observed pair of hands that belonged either to one person or to a dyad. Based on Tsai and colleagues' (2011) findings, we formulated two independent predictions: (a) that people acting together would be more accurate in reproducing spatial and temporal parameters of observed actions when they observed two people in a dyad performing unimanual actions compared with when they observed individuals performing bimanual actions (Experiments 1, 2a, and 2b) and (b) that individuals would be more accurate in reproducing spatial and temporal aspects of the observed actions when they observed individuals performing bimanual actions than when they observed two people performing unimanual actions (Experiments 3 and 4).

We varied the temporal and spatial parameters of the stimuli in the tapping task to assess how different task representations for individual and joint performance constrain synchronization. We studied the effect of temporal task demands by increasing the tempo of the observed movements as each trial progressed. Pure timing measures, such as the percentage of errors participants incur omitting stimuli or producing extra taps, and the average duration of intervals between them (inter-tap-intervals [ITIs]) provide information as to the general success in matching the timing of observed actions. If imitative performance were constrained by participants' ability to reproduce the rhythm of the observed individual or joint movements, it would result in more errors and smaller ITIs for faster movement tempos. We introduced variation in the spatial configuration of the stimuli by displaying spatially compatible or incompatible actions. Because the temporal structure of the observed movements was identical in spatially compatible and incompatible trials, any changes in the participants' ability to match taps can only arise from differences in internal models activated during the observation of individual or joint actions.

Experiment 1

Experiment 1 tested our first prediction: that people acting together would be more accurate in reproducing spatial and temporal parameters of observed actions when they observed two people in a dyad performing unimanual actions compared with when they observed individuals performing bimanual actions. Participants were asked to imitate and synchronize with the tapping of an observed pair of hands that could either belong to an individual or to a group. They performed the task always in pairs and were asked to always respond to the movements of one of the observed hands. We predicted that participants' performance would be more accurate in this task when the observed hands were perceived as belonging to a group. We expected this difference to be evident in participants' ability to match their taps spatially and temporally with the observed movements, which would result in a higher

percentage of correctly matched taps when they believed they were imitating a group compared with an individual. More generally, we predicted that participants would be able to imitate the tapping of the observed movements online, though their ability to synchronize with the observed movements and the stability of their tapping over time would be negatively affected by increments in the observed movements' tempo.

Method

Participants.   A total of 32 participants (18 females, 14 males; mean age = 21. 4 years) volunteered to participate in this study in exchange for either pay or course credit. All participants were right-handed and had normal or corrected-to-normal hearing and vision. Participants were recruited from the participant pool of Radboud University, The Netherlands.

Materials and procedure.   Participants performed a tapping task in pairs, imitating the tapping of observed hands. The stimuli consisted of movies of a pair of hands tapping with their index fingers. To create the stimuli the hands of a male model were photographed resting flat on a dark surface; one picture showed the hand with the left and one with the right index finger in extension. Presenting the index- and left-finger pictures in alternation created apparent motion. The movement tempo at which the pictures alternated was manipulated to show different tapping tempos (see following paragraph). Pictures were edited in Adobe Photoshop CS6 to yield two conditions: a condition that showed a left hand and a right hand in the posture in which they had been photographed (i.e., one person's movements; individual condition), and a condition that showed two right hands (i.e., two people's movements; dyad condition). This was achieved by rotating the left hand from the original photographs along its horizontal axis (see Figure 1, upper panel). The stimuli were displayed on a computer monitor within a rectangular template (subtending 14.25° of visual angle horizontally and 3° vertically) in the middle-upper portion of the screen. The hands were presented at the left and right sides of the template (subtending 4.77° of visual angle horizontally and 3° vertically; see Figure 2, upper panel).

Participants sat next to each other at a table in a well-lit room facing a laptop monitor located 60 cm away at chest level, and with a button box placed between themselves and the computer. Participants were asked to rest their elbows on the table and their right hands to the side of the button box with their fingers on the buttons. They completed four randomized trials: two in the individual condition, in which they saw a left and a right hand, and two in the dyad condition, in which they saw two right hands on the screen (see Figure 1). Each trial lasted 70 s and consisted of seven 10-s intervals (a training phase and six movement tempo conditions) concatenated to create the illusion of a seamless movie of two hands tapping at an ever-increasing speed. The tempo of the observed hands can be described in two ways: in terms of the individual tempo of each hand (within hand), and in terms of the alternation between left-hand and right-hand taps (between hands). Because participants were only responding to one of the hands, we describe the tempo manipulation in terms of the "within hand" tempo. During the training phase, the stimuli moved at a 1.75-Hz movement tempo (105 bits per minute [BPM]); that is, participants saw one of the two hands tap every 571 ms. Of the six interval sequences that followed, the first was identical to the training
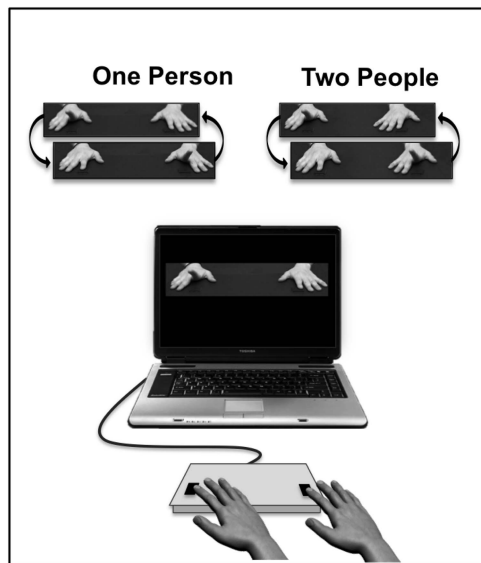
Figure 2. Experimental setup for Experiment 1.

phase and the other five displayed a progressive change in movement tempo in increments of 0.25 Hz (15 BPM). The resulting movement tempo conditions were 1.75 Hz, 2 Hz, 2.25 Hz, 2.5 Hz, 2.75 Hz, and 3 Hz. Expressed in milliseconds, these resulted in six inter-onset-intervals (IOIs): 571 ms, 500 ms, 444 ms, 400 ms, 364 ms, and 333 ms, with a timing accuracy of approximately ± 14 ms.

Participants were asked to press the buttons with their index fingers in response to the stimuli: The left participant was instructed to press her button in response to a tap by the hand on the left of the screen, and the right participant was instructed to press her button in response to a tap by the hand on the right of the screen. They were also given two additional instructions: to keep their eyes focused on the center of the screen and to do their best to keep up with the movement of the hands. Finally, they were advised that some of the trials would show one person and some would show two people. In between trials, participants were given the opportunity to stretch their hands and relax their fingers until they felt ready for the following trial. The experiment took less than 10 min to complete.

Apparatus. The stimuli were presented on a HP Compaq 14.1-in. (43.18-cm) laptop monitor with a horizontal resolution of 1,280 pixels and a vertical resolution of 800 pixels (32-bit color). The monitor was widescreen and antiglare, and the vertical refresh rate was 60 Hz. A button box with two response keys transferring at 38,400 bits per second, which allowed for a sampling rate below 5 ms, was connected to the USB port and was used to register participants' taps. The same laptop computer running Windows Vista recorded taps and the software Presentation (Version 0.90; http://www.neurobs.com) controlled stimulus presentation and the button box.

Dependent measures. Different dependent measures were selected to address distinct aspects of participants' performance. Four synchrony measures assessed how changes in movement tempo affected tapping performance: First, the *tap ratio* provided a global measure of how well participants reproduced the tempo of the observed movements. It was obtained by calculating the ratio

of observed taps over produced taps, averaged over movement tempo conditions. A value smaller than 1 indicates that some taps were omitted, and a value larger than 1 indicates that the number of performed taps was greater than the number of observed taps.

Second, ITIs were used to evaluate, in more detail, how accurately participants reproduced the observed intervals between alternating movements, and the match in interval duration between produced and observed taps. ITIs for each movement tempo condition were calculated as the average time between consecutive taps for each participant individually. ITIs were corrected for skipped stimuli by subtracting the ITI between the onset of skipped stimuli and the temporally closest tap from the final calculation.

Third, mean *asynchrony* evaluated how early or late taps occurred on average relative to the start of each observed movement, but only to observed movements that spatially matched those of the participant (e.g., the ability of the participant sitting on the left to synchronize with the observed movements of the left hand). Mean asynchrony was obtained by subtracting the response onset times from the stimuli onset times, for the stimuli temporally closest to the response and spatially aligned with it. In other words, the temporal distance between each response and the preceding and subsequent stimuli landing times were calculated, and the smaller of the two differences in absolute terms was selected. The mean of all the asynchronies obtained, with the original sign preserved, was then calculated for each interval sequence.

A fourth, related measurement evaluated performance stability, that is, whether participants' degree of asynchrony varied during the trial. Variability of asynchronies was measured using a *coefficient of variation* (Keller & Repp, 2004) that consisted of the mean standard deviation of asynchronies (see the next paragraph) divided by the observed movement tempo.

To test whether participants' belief that they were imitating the movements of a person or the movements of two people affected the spatial alignment between the performed and observed movements, we calculated a measure we will refer to as *percent spatial match*. That is, we assessed how well participants succeeded in aligning their taps spatiotemporally with the stimuli, so that their taps were closer in time to the hand movements on their side of the screen rather than to the movements on the opposite side of the screen. In order to calculate this measure, we first obtained, for each tap, the absolute asynchrony value (irrespective of the sign) for the closest observed movement, regardless of whether it was spatially compatible or incompatible. Each tap bearded a relationship with a spatially compatible and incompatible observed movement. We therefore established how many of the taps had smaller asynchronies to the spatially compatible movement and how many had smaller asynchronies to the spatially incompatible movement. This gave us, for each trial and each participant, a count of compatible and incompatible taps. We calculated the percent spatial match as the percentage of compatible taps over all taps produced in that trial. This measure expresses the percentage of taps that were closest in time as well as spatially compatible with the observed movements.

Error rates were analyzed using a 2×2 ANOVA with movement type (individual vs. dyad) and participant's role (responding to the observed left vs. right hand) as factors. All other dependent measures were submitted to $2 \times 6 \times 2$ repeated measures ANOVAs with movement type, participant's role, and movement tempo (1.75 Hz, 2 Hz, 2.25 Hz, 2.5 Hz, 2.75 Hz, and 3 Hz) as factors. Taps

produced during the initial training phase were excluded from all analyses.

Results

Percent spatial match. An ANOVA on the percent spatial match results showed significant main effects for movement tempo, $F(5, 145) = 28.46$, $p < .001$, $\eta_p^2 = .29$, and movement type, $F(1, 29) = 12.12$, $p = .026$, $\eta_p^2 = .49$. The interaction was not significant. Participants' taps became significantly less matched with the observed movements as their tempo increased, and, over- all, they matched the observed movements better in the dyad compared with the individual condition (see top panel Figure 3).

Error rate. The type of movement observed did not affect error rates. On average, participants responding to the left hand skipped 4.5% of the movements in the dyad condition and 4.45% of the movements in the individual condition, and participants responding to the right hand skipped 6.06% of the movements in the dyad condition and 6.9% of the movements in the individual condition. ANOVAs showed no significant differences for hand used or experimental condition.

ITI. The ANOVA showed a significant main effect for move- ment tempo, $F(5, 75) = 475.48$, $p < .0001$, $\eta_p^2 = .97$. For both movement types, ITIs decreased significantly as the observed movement tempo increased. The ITIs did not differ significantly between movement types and the interaction between movement tempo and movement type was not significant.

Asynchrony. ANOVA on the mean asynchronies showed sig- nificant main effects for movement tempo, $F(5, 75) = 5.39$, $p < .0001$, $\eta_p^2 = .26$, and movement type, $F(1, 15) = 14.50$, $p = .002$, $\eta_p^2 = .48$, and a significant interaction between movement tempo and participants' role in the task, $F(5, 75) = 2.86$, $p = .017$, $\eta_p^2 = .86$. Follow-up ANOVA analyses showed that the tempo manip- ulation affected the size of the asynchronies for the participant responding to the right side movements, $F(5, 75) = 7.36$, $p < .0001$, $\eta_p^2 = .33$, but not for the participant sitting on the left. Asynchronies were larger when observing one person (individual) compared with observing two people (dyad), and decreased sig- nificantly with increments in movement tempo but only for the participant sitting on the right (see Figure 3, bottom panel).

Coefficient of variation. Performance stability changed sig- nificantly with changes in observed movement tempo, $F(5, 75) = 9.33$, $p < .0001$, $\eta_p^2 = .39$, but was not affected by the movement type, and the interaction was not significant. Participants' tapping became more variable as the observed movement tempo increased.

Discussion

Percent spatial match results replicated and extend the findings of Tsai et al. (2011). Pairs of participants working together were more successful in matching their individual taps when the ob- served movements were perceived as belonging to a group com- pared with an individual. This result indicates that compatibility in terms of the number of observed agents and performing agents is critical not only for the imitation of discrete taps in reaction-time tasks but also for the ability to jointly imitate the continuous behavior of groups and individuals.

The observed decrement in movement ITIs indicates that par- ticipants were attempting to imitate the temporal structure of the task. However, as we expected, increments in the tempo affected the participants' ability to imitate the temporal structure of the observed movement. Regardless of whether the observed move- ment was perceived as belonging to a dyad or to an individual, participants' taps were less stable as the tempo increased. Results obtained for the movement asynchronies also indicate that the type of stimuli participants' observed affected imitation of the temporal structure of the task. Mirroring the results obtained for percent spatial match, the asynchronies were larger when the stimuli showed the hands of one individual versus those of a dyad.

Perhaps the most interesting result observed in this experiment is that the participant sitting to the left was affected by the manipulation, even though there were no changes in terms of the individual task he or she was performing. Participants sitting on the left synchronized with exactly the same stimulus in the con- dition in which they perceived a single individual's actions, and in the condition in which they perceived group actions. The fact that the performance of the participant on the left showed the percent spatial matching and asynchrony effects just like the participant on the right can only be attributed to differences in the task-level representations each condition activated. To further explore this finding, we conducted a control experiment (Experiment 2a), in which participants were asked to perform the left side of the task alone. If the observed effect was related to the activation of distinct we-representation', it should disappear when participants no lon- ger belong to a group and are asked to imitate the movements of only one hand.
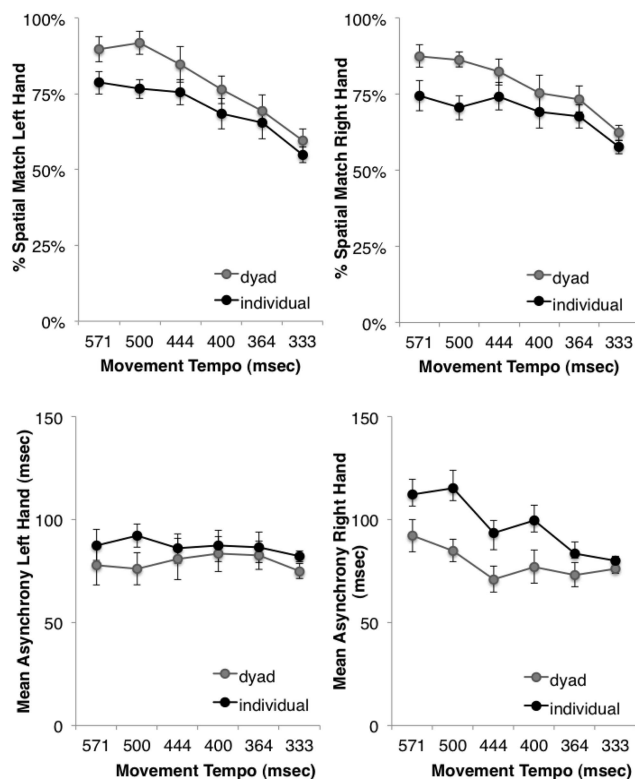

Figure 3. Mean asynchrony and mean percent spatial match results in the individual and dyad condition for the right (right panel) and left hand (left panel) in Experiment 1.

An additional control study (Experiment 2b) explored the effect of spatial compatibility in the alone performance of the participant sitting to the right.

Experiment 2a

In order to further extend the findings of Experiment 1, we tested whether the performance of the participant sitting to the left would be affected by the spatial and temporal structure of the stimuli when working alone. The results of Experiment 1 suggest that changes in imitation arise from the participants' perception of the task as a joint task and are not because of changes in the visual properties of the stimuli. We expected that the effect of observed movement type would disappear when the participants perform the task individually, whereas the effects of movement tempo on ITIs and movement stability would still be present.

Method

Participants. A total of 12 participants (7 females, 5 males; mean age = 22. 8 years) volunteered to participate in this study in exchange for either pay or course credit. All participants were right-handed and had normal or corrected-to-normal hearing and vision. Participants were recruited from the participant pool of Radboud University.

Materials and procedure. The materials and procedure were, for the most part, the same as in Experiment 1, with two important exceptions. Participants performed the task alone instead of jointly, and they were instructed to respond to the hand on the left of the screen. The second change was an increase in the number of trials. Participants completed four randomized trials (two on each movement type) with a 5-min break in between them. The experiment lasted approximately 25 min.

Dependent measures. The same dependent measures as in Experiment 1 were calculated, and the same analyses were employed.

Results

Percent spatial match. The ANOVA on the percent spatial match results showed a significant main effect for movement tempo, $F(5, 11) = 7.95$, $p < .0001$, $\eta_p^2 = .42$. Overall, participants' taps became significantly less matched with the observed movements as their tempo increased (see Figure 4, top left panel). No other effects or interactions were significant.

Error rate. The type of movement observed did not affect error rates. Participants skipped 14.43% of the movements in the dyad condition and 16.82% of the movements in the individual condition. *T*-test comparison showed no significant differences between conditions.

ITI. The ANOVA showed a significant main effects for movement tempo, $F(5, 11) = 15.95$, $p < .0001$, $\eta_p^2 = .59$. For both movement types, ITIs decreased significantly as the observed movement tempo increased. As in Experiment 1, the ITIs did not differ significantly between movement types, and the interaction between movement tempo and movement type was not significant.

Asynchrony. The ANOVA on the mean asynchronies showed a significant main effect for movement tempo, $F(5, 11) = 10.99$, $p < .0001$, $\eta_p^2 = .50$ (see Figure 4, bottom left panel). The average size of asynchronies, however, did not differ significantly between movement types, and the interaction between movement tempo and movement type was not significant.

Coefficient of variation. Performance stability changed significantly with changes in observed movement tempo, $F(5, 11) = 14.26$, $p < .0001$, $\eta_p^2 = .56$. Participants' tapping became more variable as the observed movement tempo increased, but did not change significantly for the two movement types.

Discussion

As predicted, increments in movement tempo led to changes in the participants' ability to keep up with the observed movements. However, none of the performance measures changed in response to changes in the spatial configuration of the stimuli. Together, these results add further support to the findings of Experiment 1. Interpreted in light of the findings of Experiment 1, these results indicate that, when acting alone and responding to a hand that is always in the same configuration, task-level representations or the presence of a partner do not affect imitation.

However, there is one aspect of Experiment 1 that deserves further consideration. It is possible that the changes observed for the participant sitting to the right could have arisen from differences in spatial compatibility between the observed hand and the hand used during the task. Whereas for the participant on the left, the spatial configuration of the hand to the left remained the same across conditions, the participant sitting to the right observed a hand in a mirror position to his or her own in the individual condition (looking like an individual's left hand) and a rotated hand in the dyad condition (looking like an individual's right hand). In order to control for the possibility that the observed results might have been related to the spatial incompatibility experienced by the participant on the right, we ran a control experiment (Experiment 2b) in which single participants performed the role of the participant sitting to the right (i.e., they tapped with their right hand in response to the hand to the right of the screen).

Experiment 2b

The aim of this experiment was to control for the possibility that the observed improvement for the participants who responded to the right hand in Experiment 1 was related to changes in the participants' ability to match spatial properties of the stimuli in both conditions. If the effects of Experiment 1 were replicated when the participant is asked to perform the task alone, it would indicate that other spatial aspects of the task apart from task compatibility might have led to changes in performance. That is, participants were more or less successful in one condition compared with the other because of differences in individual task demands. In contrast, a different pattern of results (i.e., either no differences depending on hand orientation, or an advantage for the hand presented in the mirror position over the rotated hand) would provide additional evidence of the role played by we-representations in the performance of continuous tasks.
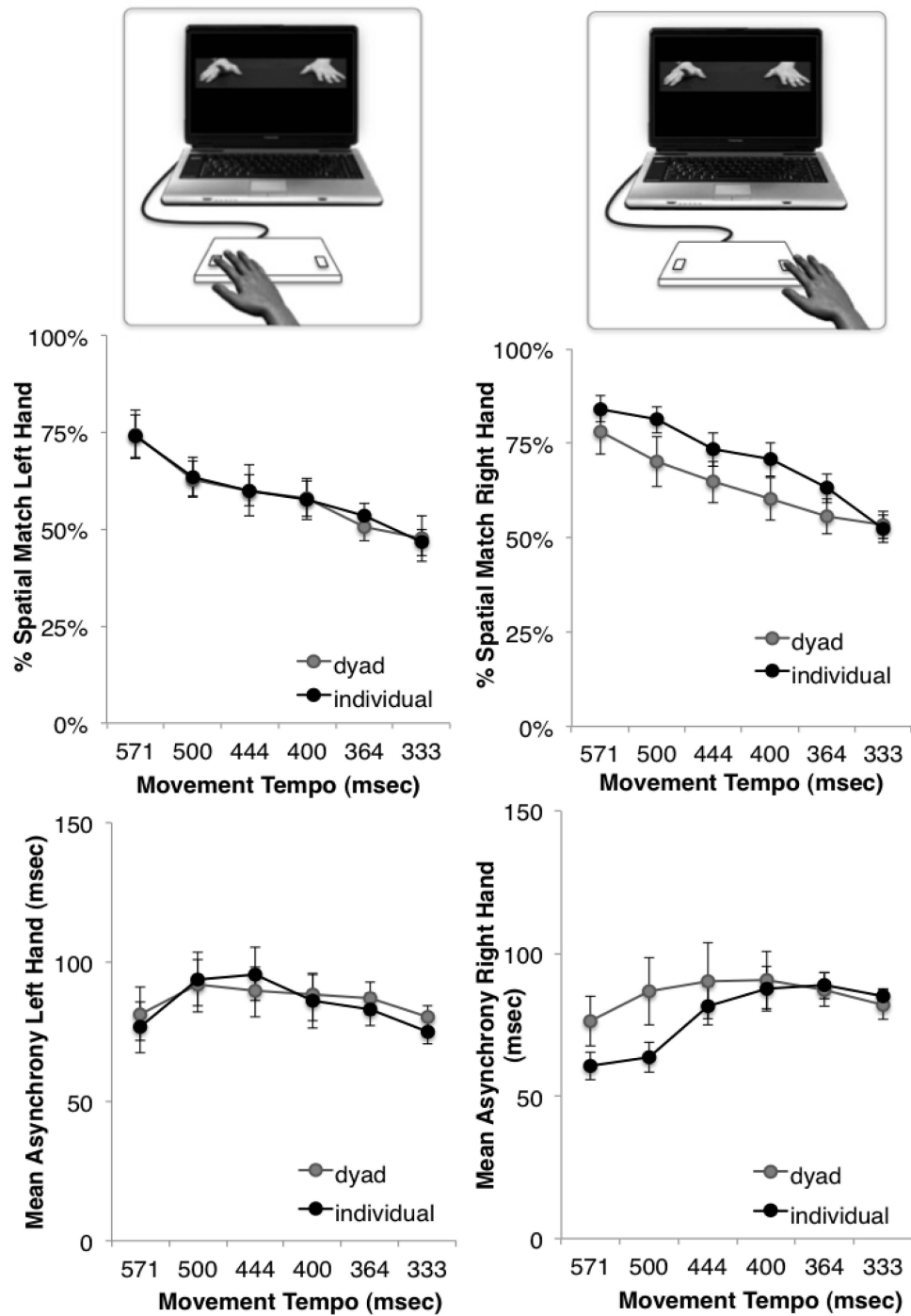
Figure 4. Setup for Experiments 2a (left panel) and 2b (right panel). Mean asynchrony and mean percent spatial match results in the individual and dyad condition for Experiments 2a (left panel) and 2b (right panel).

Method

Participants. A total of 12 participants (6 females, 6 males; mean age = 22. 3 years) volunteered to participate in this study in exchange for either pay or course credit. All participants were right-handed and had normal or corrected-to-normal hearing and vision. Participants were recruited from the participant pool of Radboud University.

Materials and procedure. The materials and procedure were the same as in Experiment 2a, with one exception: Participants responded only to the hand on the right of the screen.

Dependent measures. The same dependent measures and analyses as in Experiment 2a were employed.

Results

Percent spatial match. The ANOVA on the percent spatial match results showed significant main effects for movement tempo, $F(5, 11) = 7.21$, $p < .0001$, $\eta_p^2 = .44$, and only a tendency toward significance for movement type, $F(1, 11) = 3.76$, $p = .084$, $\eta_p^2 = .29$. The interaction was not significant. Participants' taps became significantly less matched with the observed movements as their tempo increased. Moreover, participants showed a larger degree of spatial matching in the individual condition compared with the dyad condition (see Figure 4, top right panel).

Error rate. The type of movement observed did not affect error rates. Participants skipped 13.13% of the movements in the dyad condition and 9.23% of the movements in the individual condition. *T*-test comparison showed no significant differences between conditions.

ITI. The ANOVA showed a significant main effect for movement tempo $F(5, 11) = 106.44$, $p < .0001$, $\eta_p^2 = .92$. For both movement types, ITIs decreased significantly as the observed movement tempo increased, but were not affected by changes in the spatial configuration of the stimuli. The interaction between movement type and movement tempo was not significant.

Asynchrony. ANOVA on the mean asynchronies showed a significant main effect for movement tempo, $F(5, 11) = 13.25$, $p < .0001$, $\eta_p^2 = .60$ (see Figure 4, bottom right panel), but no main effect for movement type, and no significant interaction.

Coefficient of variation. Performance stability changed significantly with changes in observed movement tempo, $F(5, 11) = 16.31$, $p < .0001$, $\eta_p^2 = .64$. Participants' tapping became more variable as the observed movement tempo increased, but was not affected by changes in the spatial configuration of the stimuli. The interaction between movement type and movement tempo was not significant.

Discussion

Participants' ability to match their taps to the observed movements showed a tendency toward significance in the opposite direction from that of Experiment 1. Whereas in Experiment 1, participants had a higher percentage of correct matches when the hand was rotated to look like two individuals' hands compared with one individual's hands, in Experiment 2b, the opposite was true. Participants were more successful in imitating the observed movements of a hand that was in a mirror position to their own.

Taken together with the results of Experiment 1, these findings suggest that the spatial compatibility manipulation affected the participants' ability to match their taps distinctly depending on their task representation. When acting jointly to match joint behavior, participants were better at mapping the temporal structure and matching their movements in time to those of a hand they perceived as belonging to a group compared with an individual. However, when working alone, participants tended to have more success in matching their taps in time to a hand in mirror position compared with a rotated hand. This difference in the pattern of results might have been related to participants focusing on the hand they were responding to and ignoring the relationship between the hands on the screen. In other words, the task-level

representation might have not operated in the same way in both studies. To test for this possibility, we tested individual bimanual performance in Experiment 3, which provided a complementary test of the role of task-level representations in continuous imitation.

Experiment 3

Experiment 3 tested our second prediction: that individuals acting alone would perform a tapping task distinctly when they were asked to imitate the observed movements of what they believe to be a group compared with what they believe to be an individual. Unlike in Experiments 2a and 2b, in which participants performed the task alone unimanually, they were asked to respond to the observed movements of both hands by tapping bimanually. We expected that if task-level representations are indeed activated during performance, results would show the opposite pattern from those of Experiment 1: Participants should be more accurate at spatially and temporally matching their actions to an observed individual's bimanual actions than to an observed dyad's unimanual actions.

Method

Participants. A total of 32 participants (18 females, 14 males; mean age = 21. 4 years) volunteered to participate in this study in exchange for either pay or course credit. All participants were right-handed and had normal or corrected-to-normal hearing and vision. Participants were recruited from the participant pool of Radboud University.

Materials and procedure. The materials and procedure were the same as in Experiment 1, with one exception: Participants performed the task alone and not jointly. They were instructed to tap bimanually mimicking the movements on the screen; they used their left hand to respond to observed left movements and their right hand to respond to the observed right movements. Because the task was bimanual, the relevant tempo participants were instructed to synchronize with corresponded to the interval between left-hand and right-hand taps (between hands). This resulted in a tempo for the training phase of 285 ms and the following six tempo conditions: 285 ms, 250 ms, 220 ms, 200 ms, 180 ms, and 165 ms.

Dependent measures. The same dependent measures were calculated. There was, however, an important difference in how they were obtained. Instead of deriving separate measures for each hand, we calculated them for both hands together. As a result, the only measure that accounted for the matching of observed tap (left or right) and produced tap (left hand or right hand) was percent spatial match. Mean error rates were compared across movement types using a *t* test. All other dependent measures were submitted to 2×2 repeated measures ANOVAs with movement type (individual vs. dyad) and movement tempo as factors.

Results

Percent spatial match. The ANOVA on the percent spatial match showed significant main effects for movement tempo, $F(5, 155) = 7.41$, $p < .001$, $\eta_p^2 = .19$, and movement type, $F(1, 31) = 5.49$, $p = .026$, $\eta_p^2 = .15$, and a significant interaction between movement tempo and movement type, $F(5, 155) = 4.04$, $p = .001$,

$\eta_p^2 = .12$. Pairwise comparisons showed that the participants' taps were significantly more matched with the observed movements in the individual compared with the dyad condition but only for the two slower tempos (1.75 Hz: $t[31] = -2.95$, $p = .006$; 2 Hz: $[31]) = -3.04$, $p = .005$). Participants' ability to match their taps was significantly affected by changes in movement tempo in the individual, $F(5, 155) = 9.65$, $p < .001$, $\eta_p^2 = .24$, but not in the dyad condition (see Figure 5, bottom panel).

Error rate. Error rates were not affected by stimulus condition. On average, participants skipped 13.65% of the movements in the dyad condition and 10.32% of the movements in the individual

Figure 5. Mean asynchrony and mean percent spatial match results averaged over hand in the individual and dyad condition for Experiments 3.

condition. *T*-test comparison showed no significant differences between conditions.

ITI. The ANOVA showed significant main effects for movement tempo, $F(5, 155) = 350.35$, $p < .001$, $\eta_p^2 = .92$, and movement type, $F(1, 31) = 17.99$, $p < .001$, $\eta_p^2 = .37$, and a significant interaction between movement tempo and movement type, $F(5, 155) = 2.67$, $p = .024$, $\eta_p^2 = .08$. Pairwise comparisons showed that ITIs were significantly smaller for the individual compared with the dyad condition in the four middle tempos (2 Hz: $t[31] = 3.84$, $p < .001$; 2.25 Hz: $t[31] = 2.59$, $p = .015$; 2.5 Hz: $t[31] = 2.28$, $p = .030$; 2.75 Hz: $t[31] = 2.09$, $p = .045$), but were not affected by movement type in the slowest and fastest tempos. Separate analyses of the individual and dyad movement types showed significant main effects for movement tempo, $F(5, 155) = 174.81$, $p < .001$, $\eta_p^2 = .85$, and $F(5, 155) = 190.66$, $p < .001$, $\eta_p^2 = .86$, respectively. For both movement types, ITIs decreased significantly with increments in movement tempo.

Asynchrony. The ANOVA on the mean asynchronies showed significant main effects for movement tempo, $F(5, 155) = 95.97$, $p < .001$, $\eta_p^2 = .32$, and movement type, $F(1, 31) = 14.50$, $p = .001$, $\eta_p^2 = .75$, and a significant interaction between movement tempo and movement type, $F(5, 155) = 2.86$, $p = .017$, $\eta_p^2 = .08$. Pairwise comparisons showed that asynchronies were significantly larger for the dyad compared with the individual condition for the two slowest tempi (1.75 Hz: $t[31] = 2.63$, $p = .013$; 2 Hz: $t[31] = 2.27$, $p = .031$) and for the fastest movement tempo (3 Hz: $t[31] = 2.69$, $p = .011$). Separate analyses of the dyad and individual condition showed significant main effects for movement tempo for both conditions, $F(5, 155) = 30.96$, $p < .001$, $\eta_p^2 = .50$, and $F(5, 155) = 86.89$, $p < .001$, $\eta_p^2 = .74$, respectively. Asynchronies decreased significantly with increments in movement tempo (see Figure 5, top panel).

Coefficient of variation. Performance stability changed significantly with changes in movement tempo, $F(5, 155) = 26.55$, $p < .001$, $\eta_p^2 = .46$, but was not affected by the movement type. The interaction was also not significant. Though performance was rather unstable, it became progressively more stable as the trial progressed and dropped slightly for the fastest movement tempo.

Discussion

Participants' ability to match the temporal structure of the stimuli was greater when the observed movements were perceived as belonging to an individual compared with a dyad. The same was true for their ability to map their taps in time to the observed movements. These results confirm our second main prediction and, taken together with the results of Experiment 1, provide consistent evidence for the role played by task-level representations in constraining how groups and individual match observed actions to performed actions. However, these findings need to be qualified with respect to the tempo manipulation. Changes in spatial matching and temporal synchrony abilities were present for the slower movement tempos and disappeared when the tempo increased. This suggests that different task demands interact in a way that when the temporal demands are too high (i.e., the tempo is very fast), the spatial demands of the task might have a smaller impact on imitative behavior. In other words, in attempting to imitate the tempo of the stimuli, participants no longer care to, or are able to, align their responses spatially with the stimuli.
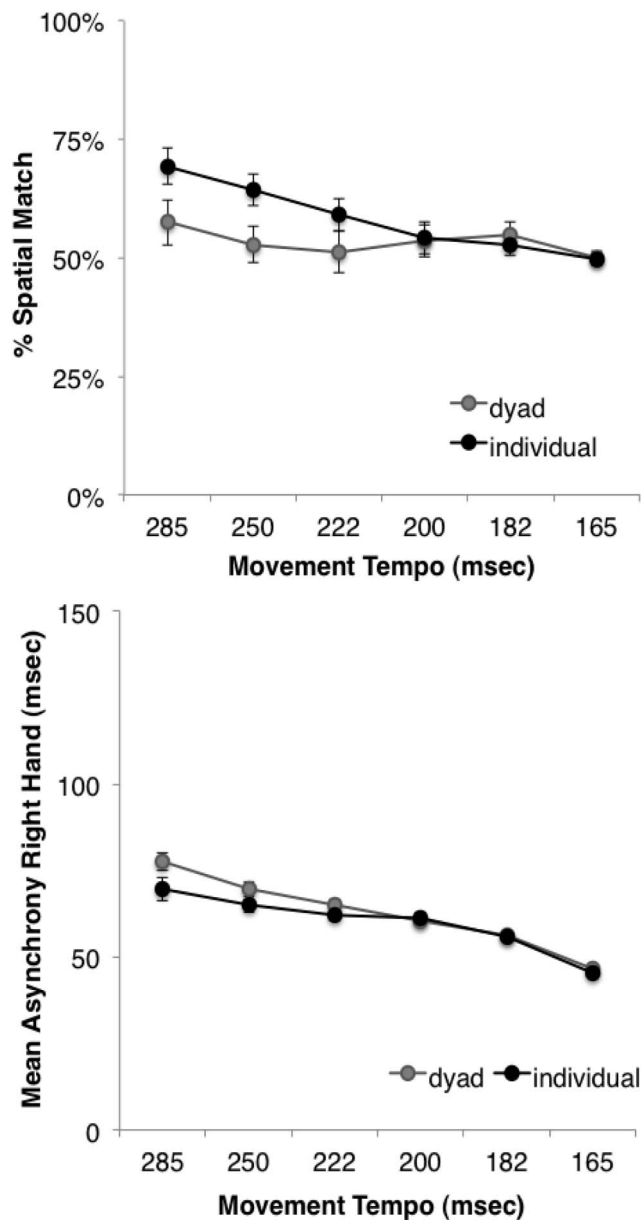
Additional evidence for the role of task-level representations on matching abilities can be found in the ITIs. The effect of movement type on the size of the ITIs for most movement tempos, save for the slowest and the fastest one, suggests that bimanual coordination dynamics are affected by task-level representations. Participants produced ITIs that were smaller and closer to those shown by the stimuli when the hands were perceived as belonging to one individual rather than to a group. This indicates that, to an extent, participants' perception of how the two hands on the screen are coordinated with each other is modulated by how they represent the hands and by whether they ascribe the hand movements to an individual or to a group.

It is possible that the effects observed for both spatial matching and temporal synchrony abilities might have been due merely to the spatial compatibility of the observed hand movements in the individual condition compared with the dyad condition. In other words, performance might have suffered not because of a mismatch at the task level (individual imitating an individual vs. individual imitating a group), but rather because in the dyad condition one of the hands is no longer anatomically aligned with that of the participant. To control for this possibility, in Experiment 4, we presented participants with two hands in the same spatial configuration in both conditions (see Figure 6).

Experiment 4

The goal of this experiment was to control for the possibility that the effects in Experiment 3 had to do with the differences in the visual configuration of the stimuli and not with the crucial

Figure 6. Stimuli used in Experiment 4 (top panel) and percent spatial match averaged over hand in the individual and dyad condition.

movement type manipulation. Hands were now presented in the same spatial configuration in both conditions. For the individual movement condition, the stimuli were the same as in previous experiments and showed the right and left hands of a man. For the condition showing the movements of a dyad, the right hand of a man and the left hand of a woman were shown in the same position as the individual's hands in the individual movements condition. If the effects of movement type on the participants' ability to match the temporal structure and to map the spatial structure of the stimuli were related to the activation of task-level representations, and not to differences in spatial compatibility across conditions, they should be replicated in this experiment.

Method

Participants. A total of 22 participants (11 females, 11 males; mean age = 22. 3 years) volunteered to participate in this study in exchange for either pay or course credit. All participants were right-handed and had normal or corrected-to-normal hearing and vision. Participants were recruited from the participant pool of Radboud University.

Materials and procedure. The materials and procedure were, for the most part, the same as in Experiment 3, with two important exceptions. There was a change in the stimuli used: for the dyad movement condition, the hand to the right of the screen was replaced by a female hand and was presented in the same orientation as in the individual movement condition. To create the new stimuli, a female model was photographed under identical circumstances as in Experiment 1 (see Figure 5). The second change was an increase in the number of trials. Participants completed four randomized trials (two on each movement type), with a 5-min break in between them. The experiment lasted approximately 25 min.

Dependent measures. The same dependent measures as in Experiment 3 were calculated.

Results

Percent spatial match. The ANOVA on the percent spatial match showed a significant main effect for movement tempo, $F(5, 100) = 8.20$, $p < .001$, $\eta_p^2 = .29$, and a significant interaction between movement tempo and movement type, $F(5, 100) = 2.62$, $p = .029$, $\eta_p^2 = .11$. Pairwise comparisons showed that asynchronies were significantly larger for the dyad compared with the individual condition for the two slowest tempi (1.75 Hz: $t[31] = 2.62$, $p = .016$; 2 Hz: $t[31] = 3.12$, $p = .005$). Separate analyses of the individual and dyad condition showed a significant main effect of movement tempo for the dyad, $F(5, 100) = 11.69$, $p < .0001$, $\eta_p^2 = .37$, but not for the individual condition, $F(5, 100) = 0.92$, $p = .47$, $\eta_p^2 = .04$, respectively (see Figure 6).

Error rate. Error rates were not affected by the movement type or by participant's gender. On average, participants skipped 11.58% of the stimuli in the dyad condition and 8.19% of the stimuli in the individual condition. *T*-test comparison showed no significant differences between conditions.

ITI. The ANOVA on the ITIs showed a significant main effect for movement tempo, $F(5, 100) = 106.33$, $p < .0001$, $\eta_p^2 = .84$, no main effect for movement type, and no significant interaction.

Asynchrony. The ANOVA on the asynchronies showed a significant main effect for movement tempo, $F(5, 100) = 19.28$, $p < .0001$, $\eta_p^2 = .49$, no main effect for movement type, and no significant interaction
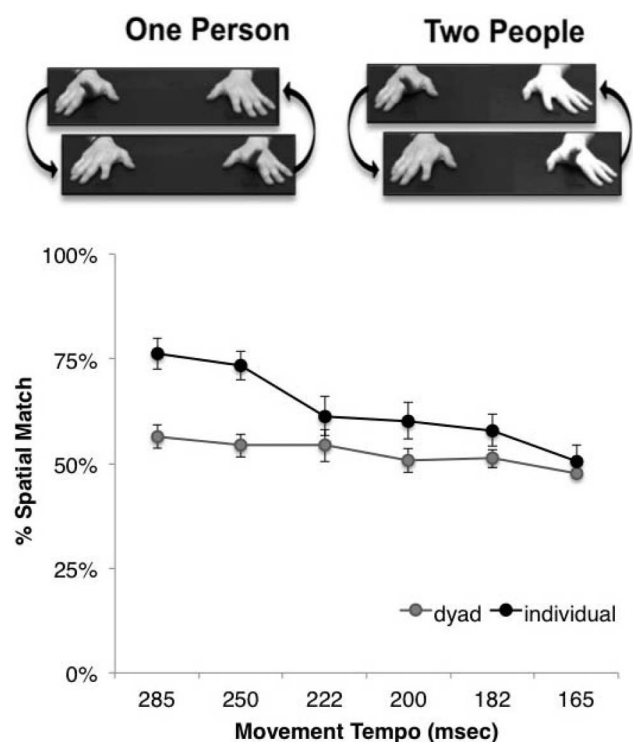
Coefficient of variation. The ANOVA on the coefficient of variation showed a significant main effect for movement tempo, $F(5, 100) = 16.27$, $p < .0001$, $\eta_p^2 = .44$, no main effect for movement type, and no significant interaction.

Discussion

The results of this experiment are in line with those of the previous experiments, and provide further support to the hypothesis that individual's and group's imitation of individual and joint actions is constrained by the activation of distinct task representations. In contrast to previous experiments, Experiment 4 found that the effect of movement type on the participants' ability to map their taps was present for slow movement tempos. One possible explanation for this finding is that the system runs against its own limits for faster tempos and can no longer cope with the distinct demands introduced by different observed movement types.

Moreover, the general effects of group-to-group matching on the participants' ability to match the temporal structure of the observed movements are consistent with the results of Experiments 1 and 3. Timing measures were not affected by the movement type observed and showed the expected decrements in size and stability as the rhythm of the observed movement increased. Taken together with the results of the percent spatial match measure, these findings indicate that in the case of individual imitation the constraints introduced by task-level representations do not impact synchronization. Together, they suggest that the perception–action links that support synchrony processes are not constrained by task-level representations in the same way as the perception–action links that ground mapping processes.

General Discussion

The aim of this study was to investigate the mechanisms that allow individuals to imitate joint actions in a continuous rhythmic task. In particular, we were interested in whether known mechanisms for establishing perception–action links in dyadic interactions (i.e., perception–action mapping) can help us understand interactions that involve more than two individuals. Although we expected the same mechanisms to support imitation across different social scales, we hypothesized that their deployment would differ depending on the social context of the task. Based on previous work by Tsai et al. (2011), we predicted that differences in how individuals represent the task (i.e., as a group or an individual task) would selectively affect their ability to imitate the spatial and temporal structure of observed movements; thus, groups would have an advantage in imitating groups over individuals, and individuals would have an advantage in imitating individuals over groups.

The results of five experiments provide converging evidence that task-level representations change how individuals go about imitating continuous rhythmic movements both when working alone and when working together. When working together, participants were better at mapping the spatial structure of group than individual movements (Experiment 1). The opposite was true for

individuals imitating unimanual (Experiments 2a and 2b) and bimanual (Experiments 3 and 4) movements; regardless of the task, individuals acting alone were always better at mapping the spatial structure of individual movements than group movements. In addition, results showed that task-level representations can affect how well people match the temporal structure of movements they observe others perform depending on whether they perceive them as belonging to one person or a dyad. Taken together, these results indicate that perception–action links supported by internal models are permeated by higher level representations about the task participants are engaged in.

These findings extend those of Tsai et al. (2011) and show that continuous real-time imitation is affected by how participants represent the social parameters of the task. Dyads working together showed a marked improvement in mapping the movements of what they perceived to be a dyad compared with an individual, and individuals showed an improvement in mapping the movements of what they perceived to be an individual over a dyad. These effects were present even when other spatial aspects of the task (hand rotation) were controlled for, so that the movements of the hands look exactly the same across conditions. This demonstrates that participants' continuous performance of joint and individual imitative actions is modulated by whether they represent the actions as joint or individual actions.

The findings of the present study add to our knowledge of how different aspects of a task modulate imitative responses. Consistent with previous work on imitation (e.g., Massen & Prinz, 2007, 2009), our results show that when participants' attention is directed toward one aspect of the stimuli (in our case, the social scale of the task they are observing and performing), the effect of other spatial aspects of the task (i.e., hand orientation) changes. This finding is evident in the results of Experiments 1 and 2b. When asked to perform the imitation task jointly (Experiment 1), participants tapping with their right hands were more successful in mapping their responses to a right hand (rotated along its vertical axes and perceived as belonging to a dyad) compared with a left hand displayed in a mirror position to their own (perceived as belonging to an individual). When working alone, they showed the opposite pattern: They were more successful in mapping their responses to a mirrored compared with a rotated hand. Hence, mapping success seems to depend on how well participants map abstract features of the task, such as the social scale (Wohlschläger et al., 2003).

Our study also addressed the effect of how participants represent the social scale of the task on temporal synchronization. Effects on synchronization were observed when participants worked jointly (Experiment 1) and alone (Experiment 3), and showed the same pattern as the spatial mapping measures: Individuals were better at matching the temporal structure of individual's rhythmic tapping than dyadic rhythmic tapping, and vice versa. These results provide an indication that task representations can affect synchronization mechanisms. This possibility is strengthened by the characteristics of our task: The stimuli were identical in terms of their temporal structure and only varied across conditions in their spatial configuration. Thus, it is unlikely that these results could be explained by stimulus differences or other lower level parameters of the task. Rather, they suggest that emergent synchronization is constrained by top-down representational processes concerning the social nature of the task.

However, effects on synchronization measures were not found consistently across all experiments, suggesting that task representations are one among several parameters modulating synchronization. Increased demands introduced by parameters other than the social scale of the task might have obscured the top-down effect of task representations. Because of the characteristics of the timing task (a sequence of short intervals with tempo increments), a more detailed exploration of the specific synchronization mechanisms that might be at play in this task is difficult. Although, for example, changes in the tempo allowed us to verify that participants were indeed attending to both the temporal and spatial characteristics of the task, task difficulty seems to have driven synchronization abilities to their limit. In addition to the increments in tempo, three aspects of the task made it uncommonly difficult. Participants were asked to (a) synchronize with visual events without scaffolding auditory information, (b) tap at very high frequencies, and (c) perform, in a sense, a dual task: aligning the taps spatially and keeping up with the tempo of the observed hands. The fact that, contrary to common findings in the literature, the asynchronies in all the experiments were positive and large seems to confirm this. It is possible that reducing the temporal difficulty of the task, by allowing participants to tap for a longer period of time at slower tempos, might allow for effects of task representations on synchronization to emerge more clearly.

Further exploration of how social task demands affect synchronization—with a task that allows for synchronization abilities to unfold more cleanly on performance—is necessary to disentangle this issue. Specifically, it would be interesting to explore whether the production and reproduction of a steady tempo for a more prolonged period of time leads to changes in the tapping timing dynamics (indexed, for instance, by changes in the standard deviation of the ITIs) depending on task-level representations. Finally, it is also possible that task representations not only affect the synchronization between performed and observed movements but also the synchronization between the performing hands. The effect of movement type in modulating the relationship between subsequent taps in bimanual performance (Experiment 3) points in that direction.

The effects of task-level representations on performance have intriguing implications for our understanding of joint action. In particular, they pose a theoretical challenge at the root of how predictive coding processes are used to account for joint action: What is predicted in the case of joint imitation, and what are the mechanisms that could link perception–action mappings and task representations (i.e., we-representations)? For individuals engaged in action imitation, the specific mapping linking their own actions to the observed movements should not, in principle, differ when acting alone or jointly. As a consequence, task-level representations do not seem to affect specific mappings but, rather, how these mapping are selected or deployed in the context of performing particular tasks. The participants' understanding of task demands seems to modulate internal model selection and deployment.

It is often the case that we observe multiple actions, and the system is required to select motor commands to respond only to the actions that are relevant to the action goals of the individual. For instance, when observing the actions of two hands, predictions are generated for the motor outcome of both hands' movements. If the task requires multiple responses (e.g., bimanual tapping), predictions are generated for each hand and both motor commands are executed (see Figure 1, Panels C and D). Conversely, when the actions of two hands are observed and instructions require participants to respond to only one hand (see Figure 1, Panels A and B), the motor command generated for the observed movements of the other hand need to be suppressed. Different internal models (linked to joint or individual task demands) might lead to changes in terms of how predictions about observed actions are treated by the system (Novembre, Ticini, Schütz-Bosbach, & Keller, 2012). In order to predict the outcome of multiple actions in the joint case, participants would have to deploy a predictive model to link their own and their partner's actions to their consequences, and ignore the motor commands generated in response to observing the actions that are relevant to the partner's task (see Figure 1, Panel A).

Differences in the compatibility between the observed and perceived movements could prime the selection of motor commands. Individuals acting alone would have an easier time generating motor commands to match multiple observed movements when they perceive them as originating in one agent (see Figure 1, Panel D) and not as belonging to two separate agents (see Figure 1, Panel C). Conversely, individuals acting together would activate an internal model to map joint actions onto joint actions, and observing the actions of an individual would have resulted in a difficulty suppressing motor responses that originate from observing their partner's referent action (see Figure 1, Panel A).

Though this project was not designed to explore this possibility, spontaneous synchronization processes, also referred to as entrainment (see Amazeen et al., 1995; Schmidt & Turvey, 1994; Schmidt et al., 1998), might also play a role in group interactions. Group interactions offer a range of possibilities for entrainment relationships to emerge (e.g., across members of each group, across groups, across an individual and his or her referent in the other group; see Figure 1). An interesting possibility we are currently exploring is that entrainment might help individuals working in a group mimic coupling relationships they observe in other groups.

As in the case of entrainment, mutual adaptation mechanisms might also play a role in continuous imitation of joint action. Mutual adaptation directly impacts synchronization success: The degree to which participants come to synchronize their actions is determined by their ability to mutually change each other's dynamics and adapt to the resulting changes online (Konvalinka, Vuust, Roepstorff, & Frith, 2010). Because of its design, the current study does not allow us to investigate whether interacting groups adapt to each other (as the timing of the stimuli was independent of the participants' performance). This remains an open area of research that deserves further exploration.

Theoretical accounts inspired by self-organization theory could explain some of the specific findings in our study. For instance, the task we used could be conceived as an instance of local (i.e., imitate the temporal and spatial features of the observed movement of a hand) and global (i.e., imitate joint vs. individual movements, alone or jointly with somebody else) processes coexisting within performance (Engstrom & Kelso, 2006). In the spirit of the extended Haken-Kelso-Bunz model of coordination dynamics (Kelso, Delcolle, & Schoner, 1990), the complementarity between these two levels is viewed as evolving within the bounds of a metastable regime. What we have chosen to refer to as representations could correspond to different tendencies within the regime that push the system in different directions depending on task constraints (Kelso, 2012). Task constraints might, hence, impact global and

local processes without necessarily resulting in changes in temporal dynamics (ITIs) or temporal coupling dynamics (asynchronies, coefficient of variation) at the local level. This could explain why we find effects for the spatial matching of taps to stimuli and not for their temporal coupling (as in Experiment 4). Similarly, local changes in task constraints could be signaling a meaningful change in global task constraints, which would lead to a different and opposite performance than that expected by attending only to local changes. For instance, the change in the orientation of the observed hand in Experiments 1 indicates a relevant change in the compatibility between performers and observed models, whereas in Experiment 2b, it only has local implications for the one performing hand.

The notion of we-representations as driving the differences we observed across experiments in this study presents a viable theoretical account for why individuals go about performing imitation tasks differently when they are perceived to be joint versus individual task. However, it is possible that "me" and "we" processes are in continuous interaction and that task constraints might drive performance toward one mode, depending on contextual circumstances. Our results seem to indicate that when the social compatibility demands of the task remain stable, the system adjusts by going about the imitation task differently. It remains to be explored whether transient changes in task constraints might result in differences in performance, such that imitating dancers might alternate between focusing on their own or on the couple's movements. The notion that coordination dynamics lives within a spectrum in which the tendency toward the individual and the collective coexist, and change dynamically to accommodate task constraints, could provide a fertile ground for integrating representation inspired findings, such as ours, and notions originating in self-organization approaches (e.g., Haken-Kelso-Bunz extended model).

The findings of the present study contribute to our understanding of the mechanisms that give rise to perception–action links in different kinds of social interactions. They show that the same mechanisms that account for our ability to imitate others in dyadic interactions play an important role in how we go about engaging and sustaining group interactions. Although focusing on one aspect of group interactions (how an individual group member imitates his or her referent in another group), these results demonstrate that task representations constraining internal models can provide a principled way in which to organize information that would allow us to "scale up" our understanding of joint action to relationships that involve more than two individuals. Further research is needed to test how perception–action mapping and entrainment (and potentially self-organization) mechanisms help us explain other aspects of group interactions, such as group-to-group coordination. However, the current findings suggest that both entrainment and representational mapping mechanisms might change in response to changes in the social nature of the task. It is an open question how these different mechanisms integrate different aspects of group coordination and how they enable us to learn from other by observing them.

References

Amazeen, P. G., Schmidt, R. C., & Turvey, M. T. (1995). Frequency detuning of the phase entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics, 72,* 511–518. doi:10.1007/BF00199893

Belopolsky, A. V., Olivers, C. N. L., & Theeuwes, J. (2008). To point a finger: Attentional and motor consequences of observing pointing movements. *Acta Psychologica, 128,* 56–62. doi:10.1016/j.actpsy.2007.09.012

Bird, G., Brindley, R., Leighton, J., & Heyes, C. (2007). General processes, rather than "goals," explain imitation errors. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1158–1169. doi:10.1037/0096-1523.33.5.1158

Blakemore, S.-J., & Frith, C. D. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia, 43,* 260–267. doi:10.1016/j.neuropsychologia.2004.11.012

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial and imitative cues. *Brain and Cognition, 44,* 124–143. doi:10.1006/brcg.2000.1225

Brass, M., Derrfuss, J., & 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., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. *NeuroImage, 14,* 1416–1423. doi:10.1006/nimg.2001.0944

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893–910. doi:10.1037/0022-3514.76.6.893

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 181–204. doi:10.1017/S0140525X12000477

Engstrom, D., & Kelso, J. A. S. (2006). *The Complementary Nature*. MIT Press, Cambridge, MA.

Flach, R., Knoblich, G., & Prinz, W. (2003). Off-line authorship effects in action perception. *Brain and Cognition, 53,* 503–513. doi:10.1016/S0278-2626(03)00211-2

Franz, E. A., Ford, S., & Werner, S. (2007). Brain and cognitive processes of imitation in bimanual situations: Making inferences about mirror neuron systems. *Brain Research, 1145,* 138–149. doi:10.1016/j.brainres.2007.01.136

Frith, C. D., & Frith, U. (2006). How we predict what other people are going to do. *Brain Research, 1079,* 36–46. doi:10.1016/j.brainres.2005.12.126

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences, 9,* 290–295. doi:10.1016/j.tics.2005.04.012

Hayes, D. P., & Cobb, L. (1982). Cycles of spontaneous conversation under long-term isolation. In M. Davis (Ed.), *Interaction rhythms: Periodicity in communicative behaviour* (pp. 319–340). New York, NY: Human Sciences Press.

Keller, P. E., & Repp, B. H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands. *Quarterly Journal of Experimental Psychology, 57*A, 1085–1101.

Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition: An International Journal, 16,* 102–111. doi:10.1016/j.concog.2005.12.004

Kelso, S. (2012). Multistability and metastability: Understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367,* 906–918.

Kelso, J. A. S., DelColle, J., & Schoner, G. (1990). Action-perception as a pattern formation process. In M. Jeanerod (Ed.), *Attention and Performance XIII* (pp. 139–169). Hillsdale, NJ: Erlbaum.

Kilner, J. M., Hamilton, A. F., & Blakemore, S. J. (2007). Interference effect of observed human movement on action is due to velocity profile

of biological motion. *Social Neuroscience, 2,* 158–166. doi:10.1080/17470910701428190

Kilner, J. M., Paulignan, Y., & Blakemore S.-J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13,* 522–525. doi:10.1016/S0960-9822(03)00165-9

Knoblich, G., & Repp, B. H. (2009). Inferring agency from sound. *Cognition, 111,* 248–262. doi:10.1016/j.cognition.2009.02.007

Knuf, L., Aschersleben, G., & Prinz, W. (2001). An analysis of ideomotor action. *Journal of Experimental Psychology: General, 130,* 779–798. doi:10.1037/0096-3445.130.4.779

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology (2006), 63,* 2220–2230. doi:10.1080/17470218.2010.497843

Luck, G., & Sloboda, J. A. (2009). Spatiotemporal cues for visually-mediated synchronisation. *Music Perception, 26,* 5, 465–473. doi:10.1525/mp.2009.26.5.465

Massen, C., & Prinz, W. (2007). Activation of action rules in action observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 1118–1130. doi:10.1037/0278-7393.33.6.1118

Massen, C., & Prinz, W. (2009). Movements, actions and tool-use actions: An ideomotor approach to imitation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 364,* 2349–2358.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198,* 75–78. doi:10.1126/science.198.4312.75

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cerebral Cortex, 22,* 2894–2903. doi:10.1093/cercor/bhr364

Oullier, O., de Guzman, G. C., Jantzen, K. J., & Kelso, J. A. S. (2003). On context dependence of behavioral variability in inter-personal coordination. *International Journal of Computer Science in Sports, 2,* 126–128.

Oztop, E., Wolpert, D. M., & Kawato, M. (2005). Mental state inference using visual control parameters. *Cognitive Brain Research, 22,* 129–151. doi:10.1016/j.cogbrainres.2004.08.004

Pezzulo, G., Candidi, M., Dindo, H., & Barca, L. (2013). Action simulation in the human brain: Twelve questions. *New Ideas in Psychology, 31,* 270–290.

Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science, 27,* 423–456. doi:10.1016/j.humov.2008.02.016

Repp, B. H., & Knoblich, G. (2007). Toward a psychophysics of agency: Detecting gain and loss of control over auditory action effects. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 469–482. doi:10.1037/0096-1523.33.2.469

Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 62–79. doi:10.1037/0096-1523.31.1.62

Schmidt, R. C., Bienvenu, M., Fitzpatrick, P. A., & Amazeen, P. G. (1998). A comparison of within- and between-person coordination: Coordination breakdowns and coupling strength. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 884–900. doi:10.1037/0096-1523.24.3.884

Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 227–247. doi:10.1037/0096-1523.16.2.227

Schmidt, R. C., & O'Brien, B. (1997). Evaluating the dynamics of nonintentional interpersonal coordination. *Ecological Psychology, 9,* 189–206.

Schmidt, R. C., & Turvey, M. T. (1994). Phase-entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics, 70,* 369–376. doi:10.1007/BF00200334

Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1, 201–208. doi:10.1037/0096-1523.33.1.201

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 2, 326–332. doi:10.1037/0096-1523.29.2.326

Spivey, M. J. (2007). *The continuity of mind*. New York, NY: Oxford University Press

Stanley, J., Gowen, E., & Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 915–926. doi:10.1037/0096-1523.33.4.915

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study on imitation. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 6, 1746–1759. doi:10.1037/0096-1523.26.6.1746

Tsai, J.-C., Sebanz, N., & Knoblich, G. (2011). The GROOP effect: Group mimics group actions. *Cognition, 118,* 1, 135–140. doi:10.1016/j.cognition.2010.10.007

Vesper, C., van der Wel, P. R. D., Knoblich, G., & Sebanz, N. (2013). Are you ready to jump? Predictive mechanisms in interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 48–61. doi:10.1037/a0028066

Vorberg, D., & Schulze, H.-H. (2002). Linear phase-correction in synchronization: Predictions, parameter estimation, and simulations. *Journal of Mathematical Psychology, 46,* 56–87. doi:10.1006/jmps.2001.1375

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131,* 460–473. doi:10.1037/0033-2909.131.3.460

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transitions of the Royal Society of London: B Biological Science, 358,* 501–515.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London*. Series B: Biological Sciences, *358,* 593–602. doi:10.1098/rstb.2002.1238

Wolpert, D. M., & Flanagan, J. R. (2009). Forward models. In T. Bayne, A. Cleermans, & P. Wilken (Eds.), *The Oxford companion to consciousness* (pp. 294–296). Oxford, UK: Oxford University Press.