Running head: Joint Action

Psychological research on joint action: Theory and data

Günther Knoblich1, Stephen Butterfill2, & Natalie Sebanz1

1Centre for Cognition, Donders Institute for Brain, Cognition, & Behaviour, Radboud University Nijmegen, The Netherlands;

2Department of Philosophy, University of Warwick, United Kingdom

Address correspondence to:

Günther Knoblich

Radboud University

Donders Institute for Brain, Cognition, & Behaviour

Centre for Cognition

PO Box 9104

6500 HE Nijmegen

The Netherlands

E-mail: g.knoblich@donders.ru.nl

Abstract

# Introduction

Human life is full of joint actions ranging from a handshake to the performance of a symphony (Clark, 1996). As Woodworth (1939, p. 823) pointed out, in many or all cases of joint action it is not possible to fully understand the individuals’ actions in isolation from each other: ‘Two boys, between them, lift and carry a log which neither could move alone. You cannot speak of either boy as carrying half the log […]. Nor can you speak of either boy as half carrying the log […]. The two boys, coordinating their efforts upon the log, perform a joint action and achieve a result which is not divisible between the component members of this elementary group.’

How, then, can the basic processes enabling people to perform actions together be studied through psychological experiments? What are the perceptual, cognitive, and motor processes that enable individuals to coordinate their actions with others, and how can the seemingly irreducibly interpersonal components of joint actions be characterized? This article provides an overview of current theories and experiments in psychology that have substantially enhanced our understanding of joint action.

Generally, a joint action is a social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment (Sebanz, Bekkering, & Knoblich, 2006). Coordinating one’s actions with others to achieve a joint outcome, such as lifting a basket together and placing it on a table, seems to require some kind of interlocking of individuals’ behaviour, motor commands, action plans, perceptions, or intentions. Early approaches to joint action originate in philosophers’ interest in the nature of joint intentionality. These approaches specify representational systems that enable the planning of joint actions.

Philosophers generally agree that joint actions are actions done with shared intentions: what distinguishes joint from individual actions is that the joint ones involve a shared intention, and shared intentions are essential for understanding coordination in joint action. This conceals deep disagreement on what shared intentions are. Some hold that shared intentions differ from individual intentions with respect to the attitude involved (Searle, 1990 [2002]). Others have explored the notion that shared intentions differ with respect to their subjects, which are plural (Gilbert, 1992), or that they differ from individual intentions in the way they arise, namely through team reasoning (Gold & Sugden, 2007), or that shared intentions involve distinctive obligations or commitments to others (Gilbert, 1992; Roth, 2004). Opposing all such views, Michael Bratman (1992, 2009) argues that shared intentions can be realised by multiple ordinary individual intentions and other attitudes whose contents interlock in a distinctive way.

The philosophical work on joint intentionality has guided research on language use where language is conceived of as a form of joint action (Brennan & Hanna, 2009; Clark, 1996). Focusing on common perceptions, common knowledge, and communicative signals, this approach situates joint planning in particular environments and particular interaction histories. For instance, the analysis of joint actions such as assembling furniture together or playing a piano duet has revealed how speech is used to pre-specify who will do what and to agree on the specifics of the joint performance (Clark, 2005). Studies addressing how people solve spatial coordination problems have demonstrated that humans readily invent new symbol systems to coordinate their actions if conventional communication is not an option (Galantucci, 2009).

The philosophical work on joint intentionality has also inspired groundbreaking research on the phylogenetic and ontogenetic roots of joint action and social understanding (Call, 2009; Carpenter, 2009; Tomasello, 2009). Melis, Hare, and Tomasello (2006) found that chimpanzees understand when they need to elicit the help of a conspecific to retrieve food and select the best collaborators to support their actions. This indicates that humans are not the only species to possess a representational system to support the planning of joint actions. However, it seems that humans are especially prone (‘have a special motivation’, Tomasello et al., 2005) to engage in joint action and to help others to achieve their goals (Brownell, Ramani, & Zervas, 2006). For instance, one-year-old infants perform actions to help adults attain their goals (Warneken & Tomasello, 2007) and gesture helpfully to provide relevant information (Liszkowski, Carpenter, & Tomasello, 2008). By three years children understand forms of joint action involving commitment of the individual partners (Graefenhain, Behne, Carpenter, & Tomasello, 2009).

Research on perception, action, and cognitive control has focused on the nuts and bolts of joint action addressing the perceptual, cognitive, and motor mechanisms of planning and coordination. Ecological psychologists have studied rhythmic joint actions in order to determine whether dynamical principles of intrapersonal coordination scale up to the interpersonal case (Marsh, Richardson, & Schmidt, 2009). This research has shown that in many cases the movement of limbs belonging to different people follows the same mathematical principles as the movement of an individual’s limbs (e.g., Schmidt, Carello, & Turvey, 1990). Cognitive psychologists have studied how co-actors represent each other’s tasks and how the ability to predict each other’s actions supports coordination in real time (Sebanz, Bekkering, & Knoblich, 2006). The results of this research suggest that specific perceptual, motor, and cognitive processes support joint action (Knoblich & Sebanz, 2008; Semin & Cacioppo, 2007) and that the needs of joint action shape individual perception, action, and cognition (Knoblich & Sebanz, 2006).

The present article provides a review of recent joint action research with a focus on the nuts and bolts of joint action. In order to place this research in a wider Cognitive Science context we begin by outlining a set of processes supporting interpersonal coordination during joint action.

# Emergent and planned coordination

We distinguish between two types of coordination that can occur during joint action, planned coordination and emergent coordination. In planned coordination agents’ behaviour is driven by representations that specify the desired outcomes of joint action and the agent’s own part in achieving these outcomes. How much is specified about other agents’ tasks, perceptions, and knowledge may vary greatly. Typically agents engaged in planned coordination will consider others’ motives, thoughts, or perspectives. But in the limiting case they will simply wait for particular conditions beyond her control to be met (Vesper, Butterfill, Knoblich, & Sebanz, in press).

By contrast, emergent coordination (EC) is independent of any joint plans and of any awareness agents might have that reaching their goals depends on conditions beyond their control (which may all be entirely absent). In emergent coordination, coordinated behaviour occurs due to perception-action couplings that make multiple individuals act in similar ways and due to the ways in which several agents may process perceptual and motor cues in the same way as each other. Two separate agents may start to act as a single coordinated entity (Marsh et al., 2009; Spivey, 2007) because common processes in the individual agents are driven by the same cues and motor routines.

## Emergent coordination

Emergent coordination can occur spontaneously between individuals who have no plan to perform actions together as well as during planned joint actions. For instance, pedestrians often fall into the same walking patterns (Van Ulzen et al., 2008) and people engaged in conversation synchronize their body sway (Shockley, Santana, & Fowler, 2003) and mimic each other’s mannerisms (Chartrand & Bargh, 1999). In all of these instances of emergent coordination similar behaviours occur spontaneously in two agents. Because these similarities don’t seem instrumental for either individual goals or joint goals, emergent coordination has sometimes been portrayed as a single process (Semin & Cacioppo, 2007). However, if (as we believe) emergent coordination is a key facilitator of joint action then it is essential to distinguish different processes that give rise to emergent coordination according to different dimensions that drive the coordination. We will distinguish between four such processes, 1) entrainment, 2) common affordances, 3) perception action matching, and 4) action simulation.

### *Entrainment*

Entrainment is perhaps the most widely studied social motor coordination process (Schmidt, Fitzpatrick, Caron, & Mergeche, in press). For instance, two people in rocking chairs involuntarily synchronize their rocking frequencies (Richardson, Marsh, Eisenhower, Goodman, & Schmidt, 2007), and audiences in theatres tend to clap in unison (Neda, Ravasz, Brechet, Vicsek, & Barabasi, 2000). Entrainment is a process that leads to temporal coordination of two actors’ behaviour, in particular, synchronization, even in the absence of a direct mechanical coupling. In dynamical systems research interpersonal entrainment is often considered as a particular instance of the coupling of rhythmic oscillators (Schmidt & Richardson, 2008) that is frequently observed in mechanical as well as biological systems. However, there are now first indications that social coordination constitutes a specific process that can be identified through particular markers of brain activity (Tognoli, Lagarde, DeGuzman, & Kelso, 2007).

### *Common Affordances*

Whereas entrainment occurs in the direct interaction between agents, common object affordances provide the basis for a further dynamical process of emergent coordination. Object affordances (Gibson, 1977) previously discussed as the ‘funktionale Toenung’ of objects (von Uexkuell, 1920) specify the action opportunities that an object provides for an agent with a particular action repertoire. For instance a chair ‘invites’ sitting down on it. When two agents have similar action repertoires and perceive the same object, they are likely to engage in similar actions because the object affords the same action for both of them. This can lead to emergent coordination when agents perceive the same objects at the same time. Examples for objects with a common affordance that may induce emergent coordination include the arrival of a bus, an apple falling from a tree, and a bench in the park.

### *Perception-action matching: Common action representations*

A third process that can lead to emergent coordination is the matching of observed actions onto the observer’s own action repertoire. Such a matching can lead to mimicry of observed actions because perceiving a particular action activates corresponding representations that also guide the actions of the observer. Common representations in perception and action have been postulated in extensions (Hommel, Muesseler, Aschersleben, & Prinz, 2001; Jeannerod, 1999; Prinz, 1997) of ideomotor theories of voluntary action control (James, 1890) and have received neurophysiological support from single-cell studies in monkeys and brain imaging studies in humans (Rizolatti & Sinigalia, 2010). In monkeys and humans the matching is based on the similarity in actor-object relations. For instance, seeing someone grasp a grape activates grasping actions directed at small, round objects. In humans, the matching can also be based on similarity in intransitive movements that are not directed at objects. For instance, observing someone dancing will activate corresponding action representations if one knows how to dance (Calvo-Merino, Cross). The perception action match can lead to emergent coordination because it induces the same action tendencies in different agents who observe each other’s actions (Knoblich & Sebanz, 2008).

### *Action simulation: Common predictive models*

The fourth process of emergent coordination is closely related to the perception action matching described above. Once a match between observed and performed actions is established it enables the observer to apply predictive models in his or her motor system to accurately predict the timing and outcomes of observed actions. This processes is often referred to as action simulation (Sebanz & Knoblich, 2009) because it uses internal models guiding an agent’s own actions to predict other agents’ actions in real time (Wolpert, Doya, & Kawato, 2003). To illustrate, a basketball player observing a shot will be able to accurately predict whether the shot will be a hit or a miss (Aglioti, Cesari, & Romani, & Urgesi, 2008). Action simulation can lead to emergent coordination because it induces the same expectations about the unfolding of actions in different actors and thus induces similar action tendencies for future actions (Knoblich & Sebanz, 2008).

## Planned coordination

In order to perform joint actions, such as playing a piano duet or lifting a heavy log, planned coordination is required. In planned coordination agents plan their own actions in relation to joint action outcomes or in relation to others’ actions whereas planning is absent or confined to the agent’s own actions in emergent coordination. The extent to which other agents’ tasks, perceptions, and knowledge are taken into account during planning of joint actions may vary greatly. Minimally, planned coordination requires a plan that specifies the joint action outcome (j) and one’s own part in a joint action (me). The other’s part can remain unspecified as captured by the formula ‘j = me + x’ for the minimal joint action plan (Vesper, Butterfill, Knoblich, & Sebanz, in press). Starting with minimal representational requirements (Clark, 1997) allows one to address a wide range of joint actions (Tollefsen, 2005) that do not involve the sorts of detailed representation of other agents and their plans that have been postulated in philosophical approaches to joint action (Bratman, 1993; Roth, 2004; Tomasello et al., 2005). Given our focus on the nuts and bolts of joint action, assuming such detailed representations are always necessary seem unhelpfully restrictive. Among the many processes contributing to planned coordination we will focus on joint task representations and joint perceptions.

### *Joint task representations*

In minimal cases of joint action, actors represent an outcome that they are not going to achieve alone and the task they need to perform themselves. Very often, though, joint action involves representations of the other agents who are actually and potentially involved. For instance, a chimpanzee who can only get food from a tray with the help of a con-specific may select one among several potential helpers according to how useful each is likely to be (Melis, Hare, & Tomasello, 2006). This chimpanzee needs to represent the goal to obtain food and their own task of pulling a rope but need not have detailed representations of the con-specific’s actions. Often, however, representations of others’ tasks are more detailed, specifying the actions others are going to perform. This is demonstrated by children (Carpenter, 2009) and adults’ proneness to represent specifics of others’ actions and tasks (Sebanz, Knoblich, & Prinz, 2005).

Joint task representations provide control structures that allow agents to flexibly engage in planned coordination. Joint task representations do not only specify in advance the individual parts each agent (me and you in the simplest case) is going to perform but they also govern monitoring and prediction processes that enable interpersonal coordination in real time (Knoblich & Jordan, 2002; Pacherie & Dokic, 2006). For instance, two soccer players of one team, where one player is specialized on crosses and the other is specialized on headers, will monitor and predict each other’s running paths in the light of their individual tasks and in the light of the common goal of scoring.

### *Joint perceptions*

Planned coordination can be improved by including another’s perceptions into one’s representation of the other’s task. This can consist in taking the other’s perspective in situations where co-actors’ perspectives on a jointly perceived environment differ such as when two actors sit face to face looking at objects to be assembled. Or it can consist in inferring what a co-actor can or cannot perceive in situations where perceptual access to objects in the environment differs between co-actors (Brennan & Hannah, 2009; Keysar et al., 2009). Although it is debated how prone agents are to co-representing each other’s perceptions, there is evidence that at least some aspects of another’s perspective are computed even when doing so hinders one’s own performance (Samson et al, in press). Co-represented perceptions might be highly useful for planned coordination in helping to establish perceptual common ground between actors (Clark, 1996), in enabling one to adapt one’s own task, and in facilitating monitoring of the other’s task.

# Evidence (7000 words, 28 pages)

## Emergent coordination

### *Entrainment and affordance*

For a long time, social psychologists (???) and developmental psychologists (Trevarthen???) have recognized the importance of rhythmic behaviour in social interaction. Building on this earlier work psychologists subscribing to a dynamical systems perspective have proposed that entrainment is best understood as a self-organizing process that occurs in coupled oscillators (???). The claim is that just as two clocks hanging on the same wall tend to synchronize because they are mechanically coupled (Huygens, ???) individuals may become automatically coupled through perceiving the same visual, auditory, or haptic information. This hypothesis was tested in experiments that determine whether people fall into synchrony even though they try to keep their own speed (Schmidt & O’Brien, 1997). Such experiments provide converging evidence that people cannot resist falling into synchronous behaviour with others.

Schmidt and O’Brien (1997) were perhaps the first to study explicitly whether interpersonal entrainment between two people would occur even when both try to resist entrainment and try to maintain their own speed. In their experiment two persons sitting side by side moved a hand held pendulum. One person used her left hand and the other person used her right hand so that the two pendulums were located in between the two persons. On each trial both individuals started out moving their own pendulum in a speed that felt comfortable to them. Importantly, they were asked to look straight and thus did no see each other or the pendulums during this phase. For the second half of each trial both individuals were asked to “maintain their preferred tempo from the first half of the trial while looking at the other participant’s moving pendulum” (Schmidt & O’ Brien, 1997).

Two results showed that participants could not resist falling into synchrony with each other. First, cross-spectral coherence, a sensitive measure of the correlation between the timing of the two individual movements, was higher during the second half of the trial than during the first half of the trial. Second, the relative phase between the two movements was much more frequently close to 0° and 180° during the second half of the trial than during the first half of the trial. Especially the later result shows that the two individuals could not resist falling into synchrony, either in-phase (0°, same synchronized turning points for the pendulum) or anti-phase (180°, different synchronized turning points for the pendulum).

Richardson and colleagues (2007, Experiment 2) obtained further evidence to support this claim. They asked two individuals to rock in a rocking chair at their preferred tempos either while looking at each other or while looking away from each other. The results demonstrated that individuals could not resist interpersonal entrainment even if the ‘natural rocking frequencies’ (eigenfrequencies) of the two rocking chairs they were rocking in differed. Other than in the pendulum studies participants were only drawn into in-phase coordination (same synchronized turning front and back) and not into anti-phase coordination suggesting that interpersonal entrainment varies depending on the specific body parts and the specific objects being moved.

A recent study by Oullier and colleagues (2008) addressed whether effects of unintended coordination occur for tapping movements. Similarly as in the previously described experiments two individuals were instructed to tap a comfortable tempo either while looking at each other’s tapping movements or while tapping with eyes closed. Auditory signals indicated when participants should open or close their eyes. Again, the analysis of relative phase revealed that participants strongly tended to fall into synchrony (in-phase only). Surprisingly, two participants stayed entrained with each other even when they closed their eyes again after having seen each other’s movements. This finding conflicts with the view that interpersonal entrainment can be reduced to a coupling between oscillators because this would predict that each individual should return to their preferred tempo. Accordingly, Oullier et al. (2008) propose that a ‘social memory’ keeps participants synchronized when the visual input supporting the coupling is absent.

Tognoli and colleagues (2007) adapted the tapping task described above to investigate whether there are specific neural markers of interpersonal entrainment in the human electroencephalogram (EEG). In particular, they simultaneously recorded EEG from two people who were looking or not looking at each other’s tapping (see above). The results demonstrated that two oscillatory EEG components in the range between 10 and 11 Hz, Phi1 and Phi2, specifically occur during interpersonal entrainment. Whereas Phi1 activation decreased with increasing coordination Phi2 activation increased with increasing activation. It remains to be seen whether Phi1 and Phi2 can be established as a general marker of interpersonal entrainment that can be generalized across different experimental settings.

Issartel, Marin, & Cadopi, (2007) demonstrated that behavioural effects of interpersonal entrainment can even be obtained when participants are asked to freely move their forearms while explicitly instructed to ignore each other’s peripherally observed movements. Although under this instruction participants did not engage in the joint rhythmic movements characterizing the studies above, their individual motor signatures (preferred movement frequencies) became more similar when they peripherally observed each other’s movements. This demonstrates that individuals cannot resist subtle interpersonal entrainment effects for ‘freely chosen’ movements that look random and independent to an observer.

Harrison & Richardson (2010) investigated whether the same principles govern the rhythmic movement of four limbs within and across organisms (Jeka, Kelso, & Kiemel, 1993). Two participants were asked to walk around at a certain distance from one another. The participant walking behind could either see the other participant or she was mechanically connected to the participant by a big foam cube, or both. The results showed that when the two participants were only visually or mechanically coupled they fell into a coordinated walking pattern that very much resembled a horse pace. When they were visually and mechanically coupled they fell into a walking pattern that very much resembled a horse trot. These findings suggest that the same stable multi-limb coordination patterns can emerge within and across organisms (cf. Mechsner & Knoblich, 2004).

Finally, the mechanisms of interpersonal entrainment have also been investigated in situations that involve more than two persons. One famous example are the dynamics of the transformation of tumultuous applause into orderly and rhythmic clapping studied by Néda, Ravasz, Brechet, Vicsek, and Barabási (2000). They demonstrated that applauding audiences fall into synchrony by slowing down their own spontaneous clapping to roughly half its initial speed. Interestingly, the slower tempo required for synchronization of large groups considerably reduces the loudness of the applause so that synchronization disappears again to increase noise levels.

Whereas the studies described above provide ample evidence for interpersonal entrainment, the role of affordances as a mechanism for emergent coordination has so far not been addressed (Knoblich & Sebanz, 2008). Although object affordances have been studied extensively in experiments on individual perception (Jones 2003; Tucker & Ellis, 1998, Yoon, Humphreys, & Riddoch, in press), we are not aware of psychological experiments addressing the role of affordances in emergent coordination. Such experiments would need to establish that similar action affordances induced by ‘usable’ objects help actors to coordinate their actions. Such benefits should be particularly strong when actors have the same experience with the particular use of objects because coordination should profit from the increased similarity in actor-object relations that results from frequently using objects in the same way. Some researchers have started to explore how the presence of another person creates affordances for acting together (Richardson et al. 2007, 2008). This reflects an interaction between planned and emergent coordination and will be discussed below.

### *Perception action matching and action simulation*

Perception action matching is a further mechanism that can lead to emergent coordination. Whereas processes of entrainment can explain why two actors’ rhythmic actions get aligned, perception action matching can explain why individuals tend to perform similar actions (Heyes & Brass, ???) or actions that lead to similar perceptual consequences (Hommel et al., 2001; Prinz, 1997) while observing each other. Accordingly, studies on entrainment tend to address situations where people perform the same or very similar movements (but see Richardson, Campbell, & Schmidt, in press). However, visual and auditory entrainment should occur regardless of action similarity if two actions are performed at the same frequencies. The studies on perception action matching, mimicry, and action simulation described below tend to exclusively focus on the similarity between observed actions and performed actions and neglect the role of timing.

Several studies have demonstrated that observing a particular movement in another person leads to an automatic activation of the same movement in the observer (Brass, Bekkering, & Prinz, 2001; Longo et al. ???). In Brass and colleagues’ (2001) experiment participants observed a video of either a lifting movement or a tapping movement on the computer screen. In one block, participants were instructed to respond to any movement on the screen with a tapping movement. In a second block participants were instructed to respond to any movement on the screen with a lifting movement. Although participants knew exactly what to do in each trial participants were faster performing a lifting movement when they observed a lifting movement and faster performing a tapping movement when they observed a tapping movement.

Stuermer, Aschersleben, and Prinz (2000) found similar results for manual movements. Participants observed videos of a hand that performed a spreading movement or a grasping movement from a neutral position. They were instructed to react to a colour patch that occurred at the same time as the movement onset or shortly after movement onset. For one colour the response consisted in a spreading movement and for the other colour the response consisted in a grasping movement. Thus the observed hand movement was irrelevant for the response. Nevertheless, responses were faster when they corresponded to the observed movement providing further evidence for the assumption that observing a movement activates the same movement in the observer’s motor repertoire.

In another study, Kilner, Paulignan, and Blakemore (2003) asked participants to perform vertical or horizontal arm movements while observing vertical or horizontal movements of a human actor or of a robot. They found that the participants’ arm movements became more variable if they did not correspond with the observed human movement than when they corresponded with the observed human movement. The correspondence effect was not present when robot movements were observed. This finding suggests that perception action matching occurs only if the kinematics of an observed movement are similar to the kinematics the observer would produce. Richardson, Campbell, & Schmidt (in press) have proposed an alternative explanation for this finding that is based on entrainment mechanisms.

The studies described so far all involved simple movements that were not directed at objects. However, in animal research, the paradigmatic case for a close perception action match consists in movements that serve to manipulate objects (Rizzolatti & Sinigaglia, 2010). Thus several studies have investigated whether a perception action match occurs when an observer perceives another person performing object-directed actions. Castiello and colleagues (2002) performed a study where participants observed a person grasping a small or a large object and were subsequently asked to grasp a small or large object themselves. If participants performed the same action they had observed before they were faster in initiating their action and more effective in optimizing motor parameters such a grip aperture (see also Edwards, Humphreys, & Castiello, 2003). Similarly, Bach and Tipper (2007) found that observing a person kicking a ball facilitated foot responses whereas observing a person typing on a keyboard facilitated finger responses. Griffiths and Tipper (2009) demonstrated that an observer does not only match the type of action observed but also specific kinematic parameters an observed actor adopts to avoid obstacles to reaching (Jax & Rosenbaum, 2007; van der Wel, Fleckenstein, Jax, & Rosenbaum, 2007).

Further studies have demonstrated that perception action matching can induce motor predictions in the observer (Hamilton, Wolpert, & Frith, 2004; Wilson & Knoblich, 2005; Wolpert, Doya, & Kawato, 2003). This has been tested varying the similarity between observed actions and the observer’s action repertoire. Assuming that this similarity is higher when people perceive their own previous actions than when people perceive somebody else’s previous actions Knoblich and Flach (2001) hypothesized that people should be better able to predict the landing position of a dart when observing their own throwing movement than when observing somebody else’s movement. The results confirmed the prediction and supported the assumption that perception action matching can trigger motor predictions in the observer (see Knoblich, Seigerschmidt, Flach, & Prinz, 2002, for similar results in the handwriting domain). Converging evidence for this conclusion has been obtained in a study that compared professional basketball players’ and basketball reporters’ ability to predict the outcome of basketball shots (Aglioti et al., ???). The hypothesis in this study was that basketball players’ expertise would allow them to more accurately predict whether a particular throwing movement would be a hit or miss and this hypothesis was confirmed.

Finally, it has been demonstrated that perception action matching can influence attention. A study by Flanagan and Johansson (2003) demonstrates that perception action matching can result in predictive eye movements implying that an observer allocates attention to location or objects that the observed actor is expected to manipulate next. Flanagan and Johansson recorded eye movements of a person who moved a stack of objects from one location to another and compared them to the eye movements of people who observed a video recording of these actions. The results showed that the gaze behaviour of participants observing the performance were highly similar to the gaze behaviour of the person who had carried out the original action. These results suggest that perception action matching does not only activate corresponding hand actions in the observer but also mimics processes of eye-hand coordination in the observed actor, in particular, the well known temporal order ‘eye precedes hand’.

Findings on inhibition of return across people show that perception action matching can induce inhibition of return mechanisms for locations an observed actor attended to (Welsh et al., 2005). Inhibition of return refers to the phenomenon that it takes individuals more time to detect a target when it appears in the same location as another stimulus presented shortly before the target. Welsh and colleagues (2005; 2007) demonstrated that observing a person

respond to a target in a particular location slowed down an observer’s response to a target appearing at the same location. This between-person inhibition of return effect suggests that inhibitory attention processes can be triggered by mere action observation. The results of a recent study by Frischen, Loach, and Tipper (2009) also support this conclusion. This study demonstrated that observing another’s actions triggered inhibitory attention processes of negative priming. Interestingly, in the observation condition the inhibitory processes followed the observed actor’s spatial reference frame and not the observer’s spatial reference frame.

## Emergent coordination during joint action

In the studies reviewed in the previous section emergent coordination occurred despite the fact that individuals were instructed to ignore each other’s actions and or at least were not given reason to attend to each other~~, although participants were asked to avoid it~~. This section reviews studies where emergent coordination was studied in the context of joint action including conversation. In all of these studies, two individuals showed emergent coordination of behaviour that was apparently not necessary for achieving the goal of the joint action. This includes emergent coordination of movements such as body sway (Fowler, Richardson, Marsh, & Shockley, 2008) as well as emergent coordination of eye movements between speaker and listener in a conversation (Richardson, Dale, & Shockley, 2008).

### *Entrainment and affordance (500 words)*

Although temporal coordination of speech patterns and body movements during conversation were the subject of many observational studies (Condon, 1976; Kendon, 1970; Wachsmuth, Lenzen, & Knoblich, 2008), the systematic experimental study of interpersonal entrainment during joint action and conversation is quite new. Richardson, Marsh, and Schmidt (2005) investigated interpersonal coordination of pendulum swinging while participants solved a puzzle task. Participants were asked to swing hand held pendulums while jointly solving a puzzle. Two factors were varied: Participants either saw or did not see each other and they either talked or did not talk to each other. Interpersonal entrainment occurred when participants perceived each other’s movements implying that verbal interaction alone was not sufficient to produce a coupling between the individuals. However, the lack of interpersonal entrainment in the verbal interaction condition may be due to the dual task character of the study. Rhythmic pendulum movements and verbal rhythms may not have been sufficiently similar to produce an interpersonal entrainment of manual movements through speech.

Studies on interpersonal entrainment of body sway during conversation suggest that talking to each other can be sufficient to produce interpersonal entrainment of body sway (Fowler et al., 2008), automatic movements that serve to keep a stable body posture. Shockley, Santana, and Fowler (2003) asked two individuals to find subtle differences between two cartoon pictures either of which could only be seen by one of them. Participants were either facing each other or looking away from each other. The surprising finding was that talking to each other was sufficient to create interpersonal entrainment of body sway as evidenced by a higher rate of recurrence in a cross-recurrence analysis. This analysis allows one to discover similarities in temporal patterns across different time series (Shockley, Butwill, Zbilut, and Webber, 2002). In a recent study Shockley, Baker, Richardson and Fowler (2007) extended these results by showing that particular properties of the conversation such as dyadic speaking rate and similarity in stress patterns give rise to acoustically mediated entrainment of body sway. Stoffregen, Giveans, Villard, Yank, and Shockley (2009) have identified further factors that modulate the entrainment of body sway such as the rigidity of the surface people are standing on.

However, body sway is not the only type of movement that gets entrained during conversation. Two studies demonstrated that there is also acoustically mediated emergent coordination between the eye movements of speakers and listeners. Richardson and Dale (2005) recorded eye movements from speakers describing stories from an American sitcom they were highly familiar with while looking at the main characters. The verbal utterances were replayed to listeners (new participants) who were also familiar with the same sitcom while their eye movements on the same display of the main characters were recorded. Cross-recurrence analysis was used to determine overlap in the temporal patterns of speaker and listener. This analysis showed that verbal utterances were sufficient to produce emergent coordination between the eye movements of the speaker and the listener. In other words, verbal messages lead to an overlap in the temporal rhythm between a speaker and a listener thereby aligning attention. Similar results were obtained in a setting where speaker and listener were engaged in a real time dialogue (Richardson, Dale, & Kirkham, 2007).

### *Perception action matching and action simulation*

Studies on non-conscious mimicry during dialogue have also revealed emergent coordination based on perception-action matching. These studies demonstrate that observing the actions and mannerisms of a conversation partner leads individuals to perform the same movements without being aware of mimicking their partner. Chartrand and Bargh (1999) provided a demonstration of this ‘chameleon effect’ by asking participants to take turns with another participant sitting next to them at describing photographs. The other participant was actually a confederate who engaged in particular mannerisms such as shaking her foot or rubbing her face. Video analyses showed that participants mimicked the confederate, rubbing their face more often when the confederate rubbed her face and shaking their foot when they observed their partner shaking her foot. Participants were not aware of their partner’s mannerisms and did not deliberately try to mimic them. This suggests that perceiving an action triggers corresponding action representations in the observer, which leads to overt mimicry in the context of conversation. The extent to which people mimic others’ actions depends on individual characteristics, including the tendency to take others’ perspective (Chartrand & Bargh, 1999), and the tendency to rely on contextual information and to feel close to others (for a review see van Baaren, Janssen, Chartrand, & Dijksterhuis, 2009).

## Consequences of emergent coordination

It may seem puzzling that emerging coordination occurs independently of individuals’ joint goals and sometimes even interferes with individual action planning such as when people can’t help falling into the same rhythm or mimicking observed actions. Several recent studies suggest that emergent coordination has various effects that may serve a number of different psychological functions. These effects include increased affiliation and liking of a partner, increased willingness to cooperate with the partner, and increased understanding of the meanings a partner intends to convey in a conversation.

It has long been suggested that the rapport between individuals is reflected in the synchrony of their body movements (e.g., Bernieri, 1988). However, only recently, studies have provided conclusive evidence for a causal link between entrainment and rapport. Miles, Nind, and Macrae (2009) demonstrated that people judge the connectedness of individuals in a dyad based on the perceived synchrony of their movements. Participants saw or heard pairs of walkers walking in a more or less synchronized manner, and rated their degree of rapport. The results showed that participants attributed the highest levels of rapport to those pairs of walkers that displayed in-phase or anti-phase coordination, and assigned the lowest levels of rapport to walkers displaying phase relationships that were far from in-phase or anti-phase. Thus, the most stable patterns of entrainment were clearly perceived as reflecting a close connection between individuals, regardless of whether information about the walkers’ synchrony was conveyed through visual or auditory information.

A study by Hove and Risen (2009) demonstrates that the link between entrainment and liking does not only apply to the perception of others’ rapport but also holds for individuals engaged in coordinated interaction. Participants performed a tapping task next to another person, synchronizing their finger movement with a visual or auditory signal. Each of the two individuals responded to separate signals so that the target tempo for their tapping could be more or less similar. Even though participants knew that the signals determined to which extent they and their task partner were synchronized, those who had been more in synchrony with their partner subsequently reported liking her more.

Entrainment also seems to boost people’s willingness to cooperate with group members (Wiltermuth & Heath, 2009). Participants who walked in step in groups of three subsequently made more cooperative choices in a standard experimental economy game than participants who had not walked in step. Those who had engaged in synchronized walking also reported feeling more connected and trusting each other more. The same was true for groups of three singing in synchrony. Of particular interest was the finding that following synchronous group actions, contributions to the public good persisted over time, whereas such contributions declined quickly in groups that had not engaged in synchronous behaviour. These findings suggest that by increasing group cohesion, synchronous group action serves to increase altruistic behaviour.

Improvements of joint action performance following entrainment, as well as gains in understanding due to entrainment during conversation provide further demonstrations of the benefits of entrainment. In a study by Valdesolo, Ouyang and DeSteno (2009), a group of participants rocked in rocking chairs next to each other, which allowed them to entrain, while another group sat back to back to avoid entrainment effects. Participants who had rocked in synchrony were better at an individually performed perceptual sensitivity task that required judging the speed of an occluded object, compared to participants who had rocked back to back. Interestingly, the increased perceptual sensitivity induced by the synchronized rocking may explain why synchronized dyads also performed better in a subsequent joint action task that involved steering a ball through a labyrinth together. These findings provide a first indication that entrainment may have effects on the quality of subsequent joint action performance.

The study by Richardson and Dale (2005) discussed above provides evidence that emergent coordination between the eye movements of a speaker and listener can aid understanding. The degree to which the gaze between speakers and listeners overlapped as the listener looked at a picture of the main characters in a sitcom while the speaker retold an episode looking at the same picture, predicted how many comprehension questions listeners subsequently answered correctly. A second experiment provided evidence that gaze coordination and comprehension are causally connected. While looking at the picture, listeners’ attention was drawn to particular locations at particular times by flashes of light. This made it possible to make listeners’ gaze pattern highly similar to the speaker’s. Compared to a condition where the flashes appeared at random times, participants indeed responded to comprehension questions more readily when their gaze had been drawn to the locations coinciding with the speaker’s gaze fixations.

### *Perception action match and action simulation*

Liking

Viele viele: Chartrand and others;

Pro-social/cooperation

???

## Planned coordination

### *Joint task representations*

#### Basic processes.

In the minimal case of planned coordination, agents represent the outcome to be achieved ~~by coordinating their actions with another~~ and they represent their own task. Psychological experiments performed with the aim of investigating how individual task performance is modulated by co-actors’ tasks have shed light on the question of when and how others’ tasks are represented. Although representing a co-actor’s task may not always be necessary, the findings of these experiments consistently suggest that humans form task representations that specify not only their own part, but also the part to be performed by the co-actor. Moreover, the findings suggest that task representations entailing a specification of the individual tasks each agent is going to perform govern stimulus processing (Heed, Habets, Sebanz, & Knoblich, 2010), action monitoring (Schuch & Tipper, 2007), control (Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Jing, Hung, Tzeng, 2006) and prediction (Ramnani & Miall, 2004) processes during the ensuing interaction.

A first study (Sebanz, Knoblich, & Prinz, 2003) investigated whether a response selection conflict between two action alternatives (a right and a left button press) that is known to occur within individuals, is also observed across individuals in a social setting where one’s own task does not require taking the co-actor’s actions or tasks into account.

Participants responded to pictures of a red or green ring presented on an index finger pointing left or right. When participants performed the two-choice task alone, they responded to red stimuli by pressing a left button and to green stimuli by pressing a right button. Although the pointing direction of the index finger was irrelevant, participants responded faster to red stimuli when the finger pointed left than when it pointed right, and vice versa for green stimuli. This spatial compatibility effect demonstrates that the irrelevant spatial information of the stimulus elicited a response conflict when the finger pointed to the side opposite to the button that had to be pressed.

The social version of this task tested whether the same response conflict would occur across individuals. One participant responded to red stimuli by pressing a single button in front of her. Next to this participant sat another participant responding only to green stimuli with her own button. Thus, each task could be performed without taking the co-actor’s task into account. Nevertheless, a response selection conflict was observed, with participants responding more slowly when the finger pointed at their co-actor. A control condition showed that this interference did not occur when another person merely sat next to participants. The findings of this first study suggest that participants did not ignore their co-actor. Instead, they represented the action to be executed by the co-actor so that a similar conflict in action selection occurred regardless of whether they were in charge of both actions or whether they performed the task together.

These findings were replicated in a study where participants responded to odd and even numbers with left or right key presses (Atmaca, Sebanz, Prinz, & Knoblich, 2008). The numbers ranged from two to nine and number magnitude was always irrelevant. It is well established that when individuals perform the parity task alone, as a two-choice task, left key presses are faster in response to small numbers and right key presses are faster for large numbers. This effect of number magnitude on parity judgments (the so-called ‘SNARC’ effect) has been explained by the assumption that the perception of numbers automatically activates a magnitude representation on a mental number line going from the left to the right (Dehaene, 1997). Atmaca and colleagues showed that the same effect occurs when two people sitting next to each other perform the task together, so that one responds only to even numbers and the other only to odd numbers. Participants sitting on the left were faster when responding to small numbers, and participants sitting on the right were faster when responding to larger numbers. This suggests that, as in the compatibility task described above, participants represented their own action alternative in relation to the co-actor’s actions.

These findings of “co-representation” have raised many questions. In particular, there has been considerable debate regarding the mechanisms underlying the observed co-representation effects. Guagnano and colleagues (Guagnano, Rusconi, & Umilta, in press) tested whether co-representation only occurs when two actors perform complementary tasks, taking turns in responding, or whether co-representation effects occur even when the two agents’ tasks are completely independent. In their study, the two participants in a pair performed independent detection tasks, one responding to red stimuli and the other to blue stimuli. The stimuli appeared either on the side on which the response should be made (compatible) or on the opposite side (incompatible). On 80% of the trials the stimuli required a response from both participants to avoid turn taking and to make the two tasks maximally independent.

A compatibility effect was observed when the two participants were sitting close to each other, within arm reach (in so-called peripersonal space). However, the compatibility effect vanished when the co-actors were sitting outside of each other’s peripersonal space. Based on these findings, Guagnano and colleagues suggested that if a co-actor is sufficiently close, the co-actor provides a spatial reference point for coding the location of one’s own action. Instead of representing the specifics of the other’s task, then, co-actors might simply use the other as a spatial reference. However, it remains to be specified how exactly such a spatial reference is established. Welsh (2009) reported similar compatibility effects when participants sitting next to each other crossed their hands and when they performed the same tasks with hands uncrossed. This finding suggests that if spatial coding is taking place, it can be flexibly based on the position of one’s body relative to the other’s body, or on the position of one’s hand relative to the other’s hand.

A recent study (Heed, Habets, Sebanz, & Knoblich, 2010) demonstrates that both the spatial relation between co-actors and task representations specifying the co-actor’s part play a role. In this study, participants held cubes that emitted tactile stimulation on top (index finger) or at the bottom (thumb). Their task was to indicate via a foot response at which location the tactile stimulation had occurred. A light appeared on top or at the bottom of the cube (congruent or incongruent with the tactile stimulation) and was irrelevant for the task. When participants performed this task alone, responses to tactile stimuli were faster when the irrelevant light appeared in the same location (e.g., touch and light at the bottom) compared to when the light and the tactile stimulation appeared in opposite locations (e.g., touch at the bottom, light on top). It was tested whether this cross-modal congruency effect is modulated when a co-actor performs a task involving the lights.

The results showed that the person responding to tactile stimuli could ignore the irrelevant light much better when her co-actor responded to the location of the light. This effect only occurred, however, when the person responding to lights was sitting in the peripersonal space of the person responding to tactile stimuli, and when she responded to all lights. The reduction in cross-modal congruency was not observed when the person in charge of lights responded only to one of two different coloured lights. This finding indicates that the weight assigned to the visual modality was changed when the partner’s task covered all visual events in peripersonal space. Thus, a representation of the other’s task modulated stimulus processing provided that the co-actors were in a particular spatial relation to each other.

Stimulating further debate about the mechanisms underlying co-representation, some studies indicate that for others’ actions to be included in one’s own action plan, they must be visible (Welsh, Higgings, Ray, & Weeks, 2007) and of a biological nature (Tsai & Brass, 2007), whereas other findings suggest that co-representation occurs even when people merely believe that they are acting together with another person (Ruys & Aarts, in press; Tsai, Kuo, Hung, & Tzeng, 2008).

Using the social compatibility task described at the beginning of this section, Tsai and colleagues told participants that they were going to perform the task (e.g., responding to red stimuli) together with a person in another room (responding to green stimuli) or with a computer program (the computer taking care of green stimuli). They found a compatibility effect when people believed that they were performing the task together with another person but not when they believed that they were performing the task with the computer. This indicates that the actual task performance is constrained by task representations formed in advance. Importantly, in the studies that found co-representation effects with invisible co-actors, participants constantly received (mock) feedback about the other’s actions (Ruys & Aarts, in press; Tsai et al., 2008). This feedback may be necessary to maintain a representation of the other’s task.

Electrophysiological and brain imaging methods have been used to specifically investigate the processes occurring when a co-actor does not need to act herself and awaits the other’s response. These studies have revealed two main findings. First, individuals seem to generate predictions of the other’s actions based on their representation of the other’s task (Ramnani & Miall, 2004). For instance, if two participants have been instructed to perform particular actions in response to certain colour cues, seeing a colour cue that specifies the action to be performed by a co-actor elicits activation in brain areas associated with mental state attribution, which may reflect an ongoing prediction process.

Second, electrophysiological evidence suggests that acting together requires the recruitment of control processes to ensure that one does not act when it is the other’s turn. A positive event-related potential occurring 300-500ms post- stimulus was significantly more pronounced when participants needed to inhibit an action because it was their co-actor’s turn compared to when they needed to inhibit an action because it was nobody’s turn to act (De Bruin, Miedl, & Bekkering, 2008; Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Jing, Hung, Tzeng, 2006).

Effects of joint task representations on action control have also been demonstrated by studies investigating the observation of errors (Bates, Patel, & Liddle, 2005; Schuch & Tipper, 2007; van Schie, Mars, Coles, & Bekkering, 2004). To investigate whether similar inhibitory processes occur in the person trying to stop an action and in an observer watching her co-actor stop an action, Schuch and Tipper asked participants to respond to targets as quickly as possible, but to stop if a stop signal was presented shortly after the target. It is well known that participants respond slower on the trial following a stop signal, both if they have successfully stopped and if they have made an error. The results showed that participants were not only slower after they had stopped or made an error themselves, but also after their co-actor had done so. This indicates that control processes governing one’s own actions are also active during co-actor’s performance, even if the co-actor’s performance is irrelevant for one’s own task.

The study by Ramnani and Miall (2004) mentioned above is also important in that it shows that completely arbitrary task rules are co-represented. In many other co-representation experiments, the stimuli had a spatial dimension, leading to an overlap in perceptual features of the responses to be made by the two co-actors and perceptual features of the stimuli (e.g., Sebanz et al., 2003). In contrast, in the study by Ramnani and Miall, the stimuli did not refer to the co-actors in any such way. This allows one to conclude that co-actors anticipated each other’s actions based on their task representation only. Converging results were obtained in a study where participants responded to a spatial stimulus feature next to a co-actor responding to a certain stimulus colour (Sebanz, Knoblich, & Prinz, 2005). Whereas some of the stimuli required a response from one participant only, others required a response from both participants. The results showed that participants responding to the spatial stimulus feature were slower when the stimulus colour indicated that it was also the other’s turn to respond. This suggests that the (arbitrary) task rule specifying responses to colour was represented by the individuals responding to the spatial stimulus feature.

Recently, researchers have started to address whether and how jointly practiced task rules modulate subsequent performance of another joint task (Milanese, Iani, & Rubichi, in press). It is known from studies of individual performance that when participants respond to stimuli on the left with a right key press, and to stimuli on the right with a left key press, they subsequently show a reduced or even reversed spatial compatibility effect in a task where they need to respond to colour and have to ignore the spatial position of the stimuli. That is, whereas participants would normally find it easier to make a left key press when a stimulus appears on the left, after the practice, participants find it easier to make a left key press when a stimulus appears on the right.

Milanese and colleagues used this transfer effect to demonstrate that joint practice modulates subsequent joint performance of a compatibility task in just the same way. Interestingly, transfer effects also occurred when participants first performed the practice alone followed by the joint compatibility task, but not when they first performed the practice together, and then performed the compatibility task alone. This indicates that the representations guiding joint task performance may in fact be quite different from the representations guiding individual performance, with transfer occurring more easily from the individual to the joint case. Transfer studies may thus constitute a useful new way to study the nature of joint task representations.

#### Interpersonal factors.

A related line of research has investigated how social factors, such as characteristics of the co-actors and the nature of the interaction context, modulate the tendency to take each other’s tasks into account. To investigate possible links between impairments in mental state attribution and co-representation, individuals with autism were asked to perform variants of the spatial compatibility task described above (Sebanz, Knoblich, Stumpf, & Prinz, 2005). They showed similar co-representation effects as a matched control group of typical adolescents and adults, indicating that deficits in understanding particular mental states such as beliefs do not necessarily affect the tendency to represent the rules specifying a co-actor’s task. However, recent findings provide a more nuanced view (Ruys & Aarts, in press).

Using an auditory version of the joint compatibility paradigm where participants believed that they were interacting with another person, Ruys and Aarts found that individuals who were good at inferring others’ mental states took the co-actor’s task into account regardless of the interaction context, whereas individuals who were less able to infer others’ mental states showed signs of co-representation only in a competitive context. The ability to infer others’ mental states was measured using the mind in the eyes test which requires selecting one of four terms that best describes the emotional state of different pairs of eyes (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). These results indicate that limitations in the ability to infer others’ mental states, which may come with less of a propensity to do so, may also come with a decreased tendency to co-represent others’ tasks.

The comparison between competitive and cooperative interaction contexts is of general interest in that it may reveal effects of particular prior intentions on task performance. Such effects have been demonstrated by studies comparing the performance of a grasping movement in a collaborative context, where participants reach for and place a wooden block on a table to build a tower together with a co-actor, and in a competitive context, where participants intend to place the block down sooner than their partner (Georgiou, Becchio, Glover, & Castiello, 2006; Becchio, Sartori, Bulgheroni, & Castiello, 2007). The kinematics differed already during the initial reach-to-grasp action, with longer movement duration, a higher movement path of the wrist, and a later time of opening the hand to grasp (maximum grip aperture) during cooperation than during competition. Control conditions where participants act alone under different instructions suggest that these effects are not simply due to movement speed, but instead reflect the intention underlying the grasping movement.

Finally, affect also seems to play a role in task co-representation. Using the joint compatibility task, Hommel and colleagues (Hommel, Colzato, & van den Wildenberg, 2009) found co-representation effects when participants acted together with a confederate who was friendly and cooperative, but not when they acted with a confederate who was intimidating and competitive. This result suggests that co-representation only occurs in positive relationships. However, it is also possible that mood is a key factor. When participants were presented with movies to induce a positive, negative, or neutral affective state before performing the joint compatibility task, co-representation effects occurred only following positive and neutral mood induction (Kuhbandner, Pekrun, & Maier, in press).

### *Joint perceptions*

Actors can adjust their actions to facilitate coordination if they are able to assess what their partner perceives at a particular moment in time. This may involve directing one’s own attention depending on where the other is looking, taking the other’s perspective, or inferring what the co-actor can or cannot perceive in situations where perceptual access to objects in the environment differs between co-actors.

A study by Brennan and colleagues demonstrates that co-actors are able to distribute a common search space by directing their attention depending on where the other is looking (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008). The task was to find the letter ‘O’ among lots of Qs on a computer screen. Participants were instructed to indicate the presence or absence of the O as quickly as possible by pressing one of two buttons. The two participants in a pair sat in different rooms and wore a head-mounted eye tracker each. This made it possible to indicate the current gaze position of a searcher to her partner, who saw it displayed as a cursor on her screen. Both participants thus could see where their partner was looking. Joint search performance was much better than individual performance. Interestingly, joint performance was best when participants were not allowed to talk to each other. These findings suggest that being able to see where their partner was looking allowed people to divide the search space in an efficient manner.

However, co-actors do not always have access to the same visual input. When they are in different spatial locations, they may have different perspectives on the same scene, and only some objects but not others may be visible to both. There has been considerable debate as to how such differences could be overcome in communication. On the one hand, it has been proposed that people make partner-specific adaptations based on what they assume to be common knowledge (Brennan & Hanna, 2009). On the other hand, it has been argued that mental state inferences play only a limited role because they are time consuming and cognitively demanding, whereas processes of emergent coordination may be highly useful for achieving coordination (Shintel & Keysar, 2009).

In the light of this debate, recent findings by Samson and colleagues are particular relevant (Samson, Apperly, Braithwaite, Andrews, & Scott, in press). Participants were asked to judge their own or another’s visual perspective in situations where the two perspectives were the same or different. They found that even when taking the other’s perspective interfered with their own task performance because the two perspectives differed, participants computed the other’s perspective. This indicates that there may be fast and effortless ways of visual perspective taking that could support planned coordination.

A study on effects of a co-actor’s perspective on mental rotation (Boeckler, Knoblich, & Sebanz, under review) provides further evidence that people take another’s perspective into account even when this is not required. Participants in a pair sat opposite each other and took turns in performing a mental rotation task with pictures of hands. While one participant performed the task, the participant opposite her either closed her eyes or looked at the pictures. When the co-actor looked at the stimuli to be rotated, participants were slowed down when small rotations were required (large rotations from the other’s perspective) and were speeded up when larger rotations were required (smaller rotations from the other’s perspective). These findings indicate that the other’s perspective could not be ignored. Joint attention triggered a switch to processing the stimuli within the other’s, allocentric, reference frame. Although this effect of ~~effortless~~ perspective taking might seem to constitute an instance of emergent coordination, it is important to keep in mind that the perspective taking was observed in a planned coordination context where participants knew about each other’s tasks.

## Synergies between planned and emergent coordination

## Although recent research has enhanced our understanding of the component mechanisms of emergent and planned coordination, it is still not well understood how planned coordination and emergent coordination work together in order to enable effective joint action. However, numerous studies indicate that planning joint actions taps into several different mechanisms of emergent coordination recruiting the functionality of these fast and parallel mechanisms in the service of planned interpersonal coordination.

### *Planning and Entrainment*

Entrainment is not only observed in situations where individuals do not plan to coordinate with each other but also when individuals plan to coordinate their movements in order to obtain the joint goal of producing a particular movement pattern. For instance, repetitive rhythmic movements do not seem to require elaborate planning ~~except producing a particular visual pattern~~. Accordingly, proponents of a dynamical systems approach to cognition have stressed the importance of entrainment in planned coordinated movement. In particular, they have studied whether movement coordination across individuals follows the same principles that govern the coordination of limbs within individuals.

Schmidt, Carello, and Turvey (1990) asked two persons who sat side by side to rhythmically swing their outer legs (left leg of person sitting on the left and right leg of person sitting on the right) at the same pace as a metronome that indicated different tempos. In the symmetric in-phase condition, one or two participants performed synchronous forward and backward movements with their legs, flexing and extending them at the same time. In the parallel anti-phase condition participants performed synchronous leg movements but now one participant extended the leg while the other flexed the leg and vice versa.

The main finding was that the dynamical interpersonal coupling between the movements followed several predictions of the HKB equation (Haken, Kelso, & Bunz, 1985) that was originally developed as a quantitative model of inter-limb coordination within a person. In particular, participants found it easier to perform symmetric movements than parallel movements and tended to switch from parallel into symmetric mode, especially at high movement speeds. These results show that the same quantitative relationship holds in planned within-person and across-person coordination of simple rhythmic movements.

Further studies where participants were asked to swing hand held pendulums at different speeds (Schmidt & Turvey, 1994; Schmidt, Beinvenu, Fitzpatrick, & Amazeen, 1998) or to jointly move a rocking board (Bosga, Meulenbroek, & Cuijpers, 2007) also supported this general conclusion. However, Schmidt and colleagues (1998) noted that planned interpersonal coupling of rhythmic movements was weaker and broke down more easily than intrapersonal coupling of rhythmic movements. This could be an indication that interpersonal coordination involves planning structures that are different from the one-person case, modulating stimulus-driven couplings in the two-person case.

One such difference has been addressed in a recent developmental experiment that addressed drumming in young children. Children aged 2.5 years on average deviated more from their preferred drumming tempo when they drummed with an interaction partner than when they drummed with a mechanical device producing the same rhythmic intervals as the interaction partner (Kirschner & Tomasello, 2009). In fact, their drumming in the social interaction condition was in a timing range that was not spontaneously performed by the children. This indicates that children (and adults) entrain more when they are engaged in social interactions. Thus top-down influences of joint planning may modulate the extent to which entrainment occurs.

Another interesting question with regard to the relation between planning and entrainment is how musicians manage to coordinate in joint performances. Goebl and Palmer (2009) investigated which cues pianists use to coordinate their performances. Not surprisingly, visual and auditory feedback was important for synchronization indicating a crucial role for entrainment and online error correction mechanisms (Repp & Keller, 2004). When auditory feedback was absent, pianists produced ostensive cues for each other by finger movements. This indicates that, at crucial points of the musical performance, ‘communicative’ coordination mechanisms ensure that performers’ joint plans stay aligned.

Studies on the entrainment of eye movements between speaker and listener provide further evidence for top-down modulation of the entrainment processes through common knowledge and joint plans. Richardson, Dale, and Kirkham (2007) recorded eye movements from two conversants who were engaged in a real time dialogue about a Dali painting. Before starting the conversation they either received the same or different information about Dali’s art. Cross-recurrence analysis revealed that the eye movements of two speakers who shared common knowledge were more tightly temporally coordinated. In a further study Richardson, Dale, and Tomlinson (2009) demonstrated that the coordination of eye gaze is not only modulated through common knowledge but also by sharing or not sharing a visual scene or believing that the conversation partner has or does not have access to the visual scene.

### *Planning and Affordance (200 words)*

Another interaction between planned and emergent coordination combines joint planning and dynamic actor-object relations (affordances). Perceiving the affordance of an object for oneself (would I be able to lift this object myself?), its affordance for another person (could this person lift this object?), or its affordance for the group (could we lift this object together, given the affordance the object has for me and the other) may provide the basis for deciding whether one should plan an individual action or a joint action. Richardson, Marsh, and Baron (2007, Experiment 4) investigated this hypothesis in an experiment where they asked two individuals to lift planks of different length from a conveyor belt. Participants were free to decide to lift particular planks alone or together and were required to make their decision on the fly as the plank passed by on the conveyor belt.

The results demonstrated that the decision to engage in joint action or individual action systematically depended on the ratio between plank length and the groups’ joint arm-span. Moreover, the transition from individual action to joint action followed the same dynamic principles as the transition from uni-manual to bimanual action in individual plan lifting (Richardson et al., 2007, Experiment 2). Importantly, participants with a longer armspan took into account the shorter armspan of their partner by choosing joint action more frequently than predicted by their individual armspan. Thus the results provide clear evidence that affordances play an important role in deciding whether to perform a joint action or an individual action with an object.

As study by Mottet and colleagues (2001, Experiment 2) provides a further indication that individuals assess task environments in terms of their joint action capabilities. Mottet and colleagues asked participants to jointly perform rhythmical movements as fast as possible. One person moved a pointer in order to move between two targets that varied in size and were separated by different distances. The other person could move the targets making the task easier for the person moving the pointer. The results showed that the combined movements of both persons followed Fitts’s law just as when one person performed the whole task bimanually. Fitts’s law predicts quantitatively the extent to which increases in target size and decreases in movement amplitude (distance between targets) allow for faster movements between two targets. Thus Mottet and colleagues’ study provides evidence that actors can jointly optimize performance to particular object sizes and particular distances between objects.

### *Planning and perception action matching/action simulation BEAREN*

Knuf et al., 2001; Haeberle et al., ???; Sebanz & Shiffrar, 2007 -> intentional induction in ideomotor movements

Liepelt -> top down manipulation of perception action match by attribution of intention

Miall -> Kilner paradigm top down manipulation of perception action match by attribution of intention

Van Schie, Waterschoot, & Bekkering, 2008 -> top down modulation from joint plan

Newman-Norlund, van Schie, van Zuijlen, & Bekkering (2007) -> top down modulation from joint plan

Sartori, Becchio, Bulgheroni, & Castiello (2009) -> social request affects action execution online -> speculation: specific set of conventional gestures activate frequently used joint action scripts (may also apply to Van Schie and Newman-Norlund).

Simulation:

Becchio, Sartori, Bulgheroni, & Castiello (2007b) Consciousness and Cognition, movement kinematics depending on whether people plan a joint action -> prediction of other’s part affects execution of own movement

Becchio, Satori, Bulgheroni, & Castiello (2007): partner acted in line with expectation (cooperative in cooperative context; competitive in competitive context) or displayed opposite intention (through face, posture); a partner displaying the intention to compete rendered the agent’s action more competitive. The opposite effect emerged when competing with a partner displaying the intention to cooperate: the kinematic pattern of the agent became similar to a cooperative pattern.

separate monitoring and prediction of action outcomes (Keller, Knoblich, Repp, Knoblich & Jordan, 2003; Vesper et al., in prep.). Possibility that only joint and not individual action outcomes are monitored.

Bosga et al., 2007, Motor Control: Joint lifting task (could also be in entrainment section)

Knoblich & Jordan (2003); Keller, Knoblich, & Repp (2007); Vesper/Schuboe??

Konvalinka, Vuust, Roepstorff, & Frith, in press: no leader-follower pattern; rather, constant prediction and adaptation to each other (“hyper-following”)

**Discussion**

Similarity of joint action planning/coordination processes to what is needed in ‘executive’ functions when a person performs multiple tasks. Can couples serve as a model for executive functions? Is multitasking internalized joint action?

MUSIC

Clayton et al., 2004 and Congado studies

## Unit of analysis: 100 words

Hutchins, 1995, Cockpit

-> point back to Allport citation in the beginning.

## Commitment: 200 words

Gilbert 1989, 1990

Developmental: Graefenhain, Behne, Carpenter, & Tomasello (2009).

## Competition/Dilemma vs joint action: 100 words

Game theory and competitive situations (Braun, Ortega, Wolpert 2009)

## Robotics

### *Robotics, Neuromodels: Erlhagen, who else? 100 words*

### *Robotics, Joint work spaces, Breazeal 100 words*

### *Robotics, Experience: De Jaegher, DiPaolo, & Gallagher (in press?) 100 words*

Contextual, enabling, constitutive role of social interaction in social cognition

### *Robotics, Haptic coordination 500 words*

Groten et al.,

Feth et al., Proceedings,

Auvray et al.,

Reed et al.

Ruddle et al.

Salinas & Zhai

Streubner

End with Titanic?

# References

Calvo-Merino, B., Glaser, D., Grèzes, J., Passingham, R., & Haggard, P. (2005). Action

observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral*

*Cortex, 15,* 1243-1249.

Castiello, U., Lusher, D., Mari, M., Edwards, M., & Humphreys, G. W. (2002). Observing a human or a robotic hand grasping an object: differential motor priming effects. In Common Mechanisms in Perception and Action. Series: Attention and Performance XIX (pp. 315-333). Oxford: Oxford University Press.

Jeka, J. J., Kelso, J. A. S., & Kiemel, T. (1993). Spontaneous transitions and symmetry: Pattern dynamics in human four-limb coordination. *Human Movement Science, 12*, 627-651.

Schmidt, R.C., Fitzpatrick, P., Caron, R., & Mergeche, J. (in press). Understanding social motor coordination. Human Movement Science.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of Interpersonal Coordination. In A. Fuchs & V. Jirsa (Eds.) Coordination: Neural, Behavioral and Social Dynamics. Springer.

Yoon, Humphreys, & Riddoch (in press). The paired-object affordance effect. *Journal of Experimental Psychology: Human Perception and Performance.*

Bratman, M. (1992). Shared Cooperative Activity. *The Philosophical Review, 101*(2), 327-341.

Bratman, M. (2009). Modest Sociality and the Distinctiveness of Intention. *Philosophical Studies, 144*(1), 149-165.

Gilbert, M. (1992). *On Social Facts*. Princeton, NJ: Princeton University Press.

Gold, N., & Sugden, R. (2007). Collective Intentions and Team Agency. *Journal of Philosophy, 104*(3), 109-137.

Kutz, C. (2000). Acting Together. *Philosophy and Phenomenological Research, 61*(1), 1-31.

Roth, A. S. (2004). Shared Agency and Contralateral Commitments. *The Philosophical Review, 113*(3), 359-410.

Searle, J. R. (1990 [2002]). Collective Intentions and Actions. In *Consciousness and Language* (pp. 90-105). Cambridge: Cambridge University Press.

# Acknowledgements

We thank Brian Ross for …