

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/162706>

Please be advised that this information was generated on 2017-12-07 and may be subject to change.



Joint Action: Mental Representations, Shared Information and General Mechanisms for Coordinating with Others

Cordula Vesper¹, Ekaterina Abramova², Judith Bütepage³, Francesca Ciardo⁴, Benjamin Crossey⁵, Alfred Effenberg⁶, Dayana Hristova⁷, April Karlinsky⁸, Luke McEllin¹, Sari R. R. Nijssen⁹, Laura Schmitz¹ and Basil Wahn^{10*}

¹ Department of Cognitive Science, Central European University (CEU), Budapest, Hungary, ² Faculty of Philosophy, Theology and Religious Studies and Donders Center for Cognition, Radboud University Nijmegen, Nijmegen, Netherlands, ³ School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden, ⁴ Department of Communication and Economics, University of Modena and Reggio Emilia (UNIMORE), Reggio Emilia, Italy, ⁵ School of Psychology, University of Birmingham, Birmingham, UK, ⁶ Institute of Sports Science, Leibniz University of Hannover, Hannover, Germany, ⁷ Faculty of Psychology, University of Vienna, Vienna, Austria, ⁸ School of Kinesiology, University of British Columbia (UBC), Vancouver, BC, Canada, ⁹ Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands, ¹⁰ Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany

OPEN ACCESS

Edited by:

Anna M. Borghi,
University of Bologna, Italy

Reviewed by:

Patric Bach,
Plymouth University, UK
Luisa Sartori,
Università di Roma, Italy

*Correspondence:

Basil Wahn
bwahn@uos.de

Specialty section:

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

Received: 22 August 2016

Accepted: 15 December 2016

Published: 04 January 2017

Citation:

Vesper C, Abramova E, Bütepage J, Ciardo F, Crossey B, Effenberg A, Hristova D, Karlinsky A, McEllin L, Nijssen SRR, Schmitz L and Wahn B (2017) Joint Action: Mental Representations, Shared Information and General Mechanisms for Coordinating with Others. *Front. Psychol.* 7:2039. doi: 10.3389/fpsyg.2016.02039

In joint action, multiple people coordinate their actions to perform a task together. This often requires precise temporal and spatial coordination. How do co-actors achieve this? How do they coordinate their actions toward a shared task goal? Here, we provide an overview of the mental representations involved in joint action, discuss how co-actors share sensorimotor information and what general mechanisms support coordination with others. By deliberately extending the review to aspects such as the cultural context in which a joint action takes place, we pay tribute to the complex and variable nature of this social phenomenon.

Keywords: joint action, social interaction, action prediction, joint attention, culture, sensorimotor communication, coordination

ACTING IN A SOCIAL WORLD

People rarely act in isolation; instead, they constantly interact with and coordinate their actions with the people around them. Examples of such ‘joint actions’ range from carrying a sofa with multiple people (**Figure 1**), building a toy brick tower with a child, playing basketball, to performing a musical duet. Accordingly, an often-used definition describes joint action as “any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment” (Sebanz et al., 2006, p. 70). Especially in light of a long research tradition that focused on the psychological (neuro)cognitive and perceptual processes of individuals, it is crucial to realize the importance of studying these processes in the social environment in which they typically occur.

COORDINATION MECHANISMS

Agents involved in joint action make use of different mechanisms to support their coordination. These include forming representations about, and keeping track of, a joint action goal and the specific to-be-performed tasks. Some coordination mechanisms depend on sensorimotor information shared between co-actors, thereby making joint attention, prediction, non-verbal communication, or the sharing of emotional states possible. General mechanisms, which depend

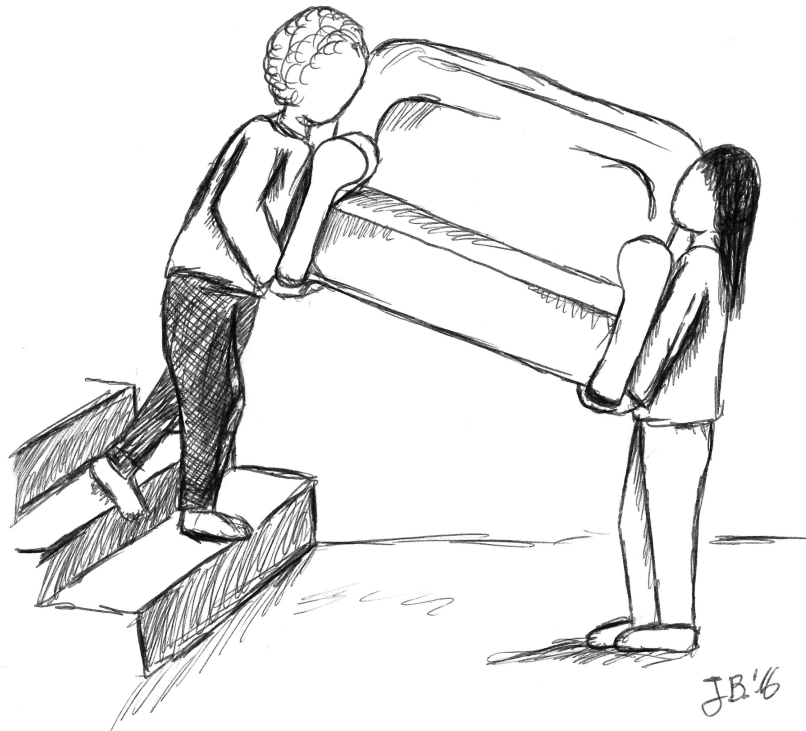


FIGURE 1 | Two people carrying a heavy sofa together face the challenge of coordinating their actions in a temporally and spatially precise manner.

to a lesser degree on sharing information online, influence and support coordination such as when co-actors rely on ‘coordination smoothers’ or conventions to act together. This article provides an overview of these coordination mechanisms (Table 1) and their role for joint action. We focus on intentional real-time joint action, in which adult co-actors share a physical space and where coordination may require high temporal precision. It therefore complements work on rhythmic or unintentional coordination (Repp and Keller, 2004; Schmidt and Richardson, 2008), verbal communication (Clark, 1996; Brennan et al., 2010), strategic cooperation (Schelling, 1960), and coordination in more temporally or spatially remote joint tasks.

MENTAL REPRESENTATIONS IN JOINT ACTION

The following subsections discuss the mental representations underlying joint action, such as representing and monitoring the joint action goal and agents’ specific tasks.

Joint Action Goal

To successfully perform a joint action, actors need to plan their own action in relation to the desired outcome and/or their co-actor’s actions. For example, Kourtis et al. (2014) showed that the neural signature of action planning is modulated when one’s own action is part of a joint plan, e.g., when clinking glasses with another person, compared to performing a corresponding

TABLE 1 | Overview of different coordination mechanisms supporting joint action, along with a set of examples.

| Coordination mechanism | Example |
|---|---|
| Mental representations in joint action | |
| Joint action goal | Relocating a sofa by lifting and moving it together |
| Task (co-)representation | Carrying a sofa forward or backward |
| Monitoring | Noticing errors in a co-actor’s performance |
| Sharing sensorimotor information | |
| Joint attention and shared gaze | Being mutually aware of an obstacle in the way |
| Sensorimotor prediction | Predicting a co-actor’s movement direction |
| Sensorimotor communication | Pushing a co-actor into a certain direction |
| Haptic coupling | Feeling a co-actor pushing the sofa |
| Multisensory processing | Integrating information from different senses |
| Emotion understanding and expression | Realizing how exhausted a co-actor is |
| General mechanisms supporting coordination | |
| Coordination smoothers | Distributing the task of moving forward or backward |
| Affordances | Being constrained by available space and a co-actor’s physical strength |
| Conventions and culture | Appreciating rules about who carries more weight |

solo or bimanual action. According to a minimal joint action account, agents intending to perform a coordinated action with others minimally represent the joint action goal and

the fact that they will achieve this goal with others (Vesper et al., 2010). This does not presuppose high-level interlocking mental representations (cf. Bratman, 1992), and might therefore form the basis for joint action in young children (Butterfill, 2012). Moreover, joint action goals influence the acquisition of new skills: after learning to play melodies in a joint action context (i.e., duets), piano novices played better when later coordinating toward the shared action goal (the duet) compared to their own action goal (the melody; Loehr and Vesper, 2016). Further evidence for the role of goal representations in joint action comes from work on complementary action (Sartori and Betti, 2015) showing that, in contrast to an imitation context, performance of an action is facilitated if the goal is to complement someone else (van Schie et al., 2008; Poljac et al., 2009).

Task (Co-)representation

In many joint actions, detailed knowledge about another's task is available and people tend to co-represent these tasks, even if detrimental to their own action performance. For instance, one might be influenced by others' stimulus-response rules in reaction-time tasks (Sebanz et al., 2003, 2005) or automatically memorize word list items relevant only to another person (Eskenazi et al., 2014). When acting together with others toward a joint goal, representing a co-actor's task can be beneficial as it enables agents to predict others' actions and to integrate them into their own action plan. For example, knowledge about a co-actor's task can be useful if access to online perceptual information about the co-actor's action is unavailable such that monitoring the co-actor's unfolding action and continuously adjusting in an appropriate manner is not possible. This was shown in a study where dyads coordinated forward jumps of different distances such that they would land at the same time (Vesper et al., 2013). Although co-actors could not see or hear each other's actions, task knowledge about the distance of each other's jump was sufficient to predict the partner's timing and adjust their own jumping accordingly.

Monitoring

While performing a joint task, co-actors typically monitor their task progress to determine whether the current state of the joint action and the desired action outcome are aligned (Vesper et al., 2010; Keller, 2012). For example, one might keep track of how far a jointly carried sofa has been moved and whether all task partners are equally contributing to lifting the weight. Monitoring is useful to detect mistakes or unexpected outcomes in one's own or one's partner's performance, enabling one to quickly react and adapt accordingly. Performance monitoring in social contexts involves specific processes and brain structures such as brain areas involved in mentalizing and perspective-taking, e.g., medial prefrontal cortex (van Schie et al., 2004; Newman-Norlund et al., 2009; Radke et al., 2011). Findings from an EEG experiment with expert musicians (Loehr et al., 2013) indicate that the neural signature associated with the detection of unexpected musical outcomes is similar irrespective of whether an auditory deviation arises from one's own or the partner's action. This

suggests that co-actors monitor the actions toward the overall joint goal in addition to their own individually controlled part.

SHARING SENSORIMOTOR INFORMATION

The following subsections provide an overview of different ways in which co-actors share sensorimotor information to support joint action through joint attention, prediction, non-verbal communication, or sharing emotions.

Joint Attention and Shared Gaze

Others' eye movements are an important source of information about what others see and about their internal states (Tomasello et al., 2005). For example, when jointly moving a sofa, co-actors may use mutual gaze to infer whether everyone is aware of a potential obstacle that is in their way (e.g., a curious dog). Joint attention relies on co-actors' ability to monitor each other's gaze and attentional states (Emery, 2000). For instance, when synchronizing actions, co-actors divide attention between locations relevant for their own and for their co-actor's goal (Kourtis et al., 2014; see Böckler et al., 2012; Ciardo et al., 2016 for similar results using different tasks), and sharing gaze affects object processing by making attended objects motorically and emotionally more relevant (Becchio et al., 2008; Innocenti et al., 2012; Scorolli et al., 2014). Moreover, in a joint search task, co-actors who mutually received information about each other's gaze location via different sensory modalities (i.e., vision, audition, and touch) searched faster than without such information (Brennan et al., 2008; Wahn et al., 2015). Together, these findings demonstrate the important role of gaze information for joint action.

Sensorimotor Prediction

Predicting others' actions and their perceptual consequences is often important for joint action. When moving a sofa together with someone, individuals need to predict what the other is going to do next in order to adapt their own action and thereby facilitate coordination. It has been postulated that action prediction relies on individuals' own motor plans and goals such that when an interaction partner's actions are observed, this activates representations of corresponding perceptual and motor programs in the perceiver (Prinz, 1997; Blakemore and Decety, 2001; Wolpert et al., 2003; Wilson and Knoblich, 2005; Catmur et al., 2007). At a functional level, action prediction can be explained in terms of internal forward models that generate expectations about the sensory consequences of partner-generated actions based on an individual's own motor experience. At a neurophysiological level, the mirror system (Rizzolatti and Sinigaglia, 2010) provides a plausible mechanism linking action observation, imagination, and representation of others' actions with motor performance. Although motor prediction has mostly been studied in action observation, some evidence demonstrates that it supports joint action by allowing precise temporal coordination (Vesper

et al., 2013, 2014) and that it is modulated by own action experience. For instance, Tomeo et al. (2012) found that expert soccer players, compared to novices, more effectively predict the direction of a kick from another person's body kinematics (see Aglioti et al., 2008; Mulligan et al., 2016, for similar results with basketball and dart players). Action prediction also affects perception (Springer et al., 2011) as predictions based on knowing another person's task can bias how their subsequent actions are perceived (Hudson et al., 2016a,b). Due to the overlap of own and others' sensorimotor representations, additional processes are needed to keep a distinction between self and other (Novembre et al., 2012; Sowden and Catmur, 2015) and to inhibit the tendency to automatically imitate another's (incongruent) action (Ubaldo et al., 2015).

Sensorimotor Communication

In some joint actions, it is useful to not only gather information about other people but to actively provide others with information about one's own actions. Accordingly, co-actors might adjust the kinematic features of their action (e.g., velocity or movement height) in order to make their own actions easier to predict for another person. Thus, 'sensorimotor communication' is characterized by having both an instrumental (e.g., pushing a sofa) and a communicative goal (e.g., informing a partner about one's movement direction). This facilitates action prediction by disambiguating different motor intentions for the observer (Pezzulo et al., 2013), thereby relying on people's ability to detect even subtle kinematic cues (Sartori et al., 2011). Studies on sensorimotor communication typically involve tasks where a 'leader' participant has information about an aspect of a joint task that a 'follower' participant lacks and so the follower has to rely on the leader's action cues to act appropriately. For example, leaders exaggerated the height of their movements to allow followers to more easily recognize the intended action target (Vesper and Richardson, 2014). Similarly, leaders communicate the end-point of a grasping action with the help of exaggerated kinematic parameters, such as wrist height and grip aperture (Sacheli et al., 2013).

Haptic Coupling

Information about another person's action might also be provided through the tactile channel. For instance, jointly carrying a sofa allows mutual exchange of force information, revealing co-actors' movement direction or speed. Accordingly, dyads who performed a joint pole-balancing task enhanced the force feedback between each other to support smooth interaction (van der Wel et al., 2011). Generally, touch can function as an information channel when joint action partners are in physical contact with each other. The ability to decode signals such as emotional cues (Hertenstein et al., 2009) from close physical interaction with their parents is a crucial aspect of children's development, establishing and regulating social encounters (Feldman et al., 2003). Mother-infant tactile communication, gaze, and emotional vocalization are found in all cultures and societies, although cross-cultural research revealed that touch

plays a more important role for communication during play and learning in traditional compared to Western societies (Richter, 1995). Moreover, tactile communication is integral to cultural practices such as dance and martial arts (Kimmel, 2009).

Multisensory Processing

Information processing in joint action is not limited to only one sensory modality: when carrying a sofa together, visual, auditory, and haptic sensory input is available, facilitating, e.g., the prediction of a partner's change in movement direction. A recent study provides support for the flexibility of multisensory processing: using a 'sonification' technique, in which kinematic movement parameters are transformed into sound, it was shown that 'sonified' forces and movement amplitudes on a rowing ergometer provide sufficient information for listeners to predict a virtual boat's velocity and to reliably discriminate own actions from those of other persons (Schmitz and Effenberg, 2012). Humans are also able to integrate redundant information from multiple sensory modalities, thereby enhancing the reliability and precision of perception (Ernst and Banks, 2002; Wahn and König, 2015, 2016). For instance, whilst the mirror system is mostly understood as a visual system sensitive to biological motion information, it is actually also tuned to auditory (Kohler et al., 2002; Bidet-Caulet et al., 2005) and audiovisual information (Lahav et al., 2007). Neuroimaging evidence shows enhanced activation of most parts of the action observation system (medial and superior temporal sulcus, inferior parietal cortex, premotor regions, and subcortical structures) when observing agents' convergent compared to divergent audiovisual movement patterns (Schmitz et al., 2013).

Emotion Understanding and Expression

Sharing emotions with others provides motivational cues helpful to initiate and continue joint tasks and to facilitate coordination (Michael, 2011). Humans are capable of reading others' affective states from body movements, body posture, gestures, facial expressions, and action performance, possibly via activation of the observer's corresponding states (Bastiaansen et al., 2009; Borgomaneri et al., 2012). A two-system model of emotional body language (de Gelder, 2006) distinguishes between automatic, reflexes-based manifestations of an emotional message and more deliberate emotional expression based on reflection and decision-making. Together, these efficiently provide information about others' emotional states and help establish and maintain joint action. For example, having an uncooperative co-actor affected participants' own response times (Hommel et al., 2009), suggesting that people adjust their own behavior according to the perceived affective states of others. Emotional body language also plays a major role in art improvisations, such as contact improvisation dance (Smith, 2014). Since improvisers explicitly use input from their partners to develop their movement interaction, this dance form allows performers to display and experiment with inner states and emotional body

language, which, in turn, influences the overall joint action outcome.

GENERAL MECHANISMS SUPPORTING COORDINATION

The following subsections introduce coordination mechanisms that depend to a lesser degree on shared online information but influence and support joint action more generally.

Coordination Smoothers

When shared perceptual information is scarce or unavailable, ‘coordination smoothers’ (Vesper et al., 2010) support joint actions. One example is reducing the temporal variability of one’s own actions, first identified in dyads who synchronized the timing of key presses in a reaction time task (Vesper et al., 2011). Co-actors’ responses were overall faster and less variable in joint compared to individual performance and variability reduction effectively improved coordination. A further coordination smoother is the distribution of tasks between joint action partners. In order to facilitate coordination, co-actors who have a relatively easier task might adapt their actions in a different way than those with a more difficult task (Vesper et al., 2013; Skewes et al., 2015). For example, if a door needs to be opened while carrying a sofa to another room, it will be done by the actor who is closer to the door while the other will momentarily take over more weight to provide support.

Affordances

Affordances are action possibilities available to an agent in an environment (Gibson, 1979). In the context of joint action, information comes from the co-actor’s body or movements and from the objects in the environment in which the joint action takes place. On the one hand, ‘affordances for another person’ specify co-actors’ action possibilities provided by their particular abilities and the environment. For example, based on the perceived relation between chair height and an actor’s leg length, observers can distinguish between maximum and preferred sitting heights of actors of different body height (Stoffregen et al., 1999). Such information is useful in understanding other agents (see Bach et al., 2014, for a review on affordance in action observation) but can also help to efficiently complement their behavior. On the other hand, ‘affordances for joint action’ (or ‘joint affordances’) concern actions available to multiple agents together. For example, when dyads lifted wooden planks alone or together, they transitioned between these two modes based on a relational measure (the ratio of plank length and both persons’ mean arm span; Isenhower et al., 2010). Social affordances might be directly perceived given that the information is publicly available. Therefore, learning to perceive affordances for others might be a natural consequence of learning to perceive affordances for oneself (Mark, 2007). Consequently, own capabilities and experiences play a role in perceiving affordances for others (Ramenzoni et al., 2008), possibly

by activating one’s own motor system (Costantini et al., 2011).

Conventions and Culture

Cultural and societal norms play a major role in regulating behaviors, social encounters, and cooperation in groups by providing conventions that can reliably guide individual behavior. Generally, culture and conventions depend on establishing and maintaining common ground between the members of a group through shared experiences (Clark, 1996). Culture is both a product of large-scale joint actions, such as celebrations or protests, and it profoundly shapes how people approach joint action in small-scale interpersonal encounters. For example, if a person of a higher social rank performs a joint task with their direct subordinate (e.g., an employer carries the sofa together with an employee), coordination might be influenced by the pre-existing power relation, the established culture (e.g., favoring hierarchical or egalitarian communication; Cheon et al., 2011) and the particular situational context (e.g., formal or informal). Joint actions involving people from different cultural backgrounds are an interesting test case for studying cooperation that is not regulated by the framework of a single culture. Different cultures might promote conflicting approaches to communication, decision making, and coordination (Boyd and Richerson, 2009) and consider different amounts of personal space, gaze, or tactile communication appropriate (Gudykunst et al., 1988). For instance, people from East Asia would typically bow for a formal greeting, whereas European people would shake hands. This cultural difference may result in a failure to perform the planned joint action of greeting properly. Strategies to avoid such unsuccessful coordination, e.g., adopting the partner’s cultural technique or establishing a new ‘third-culture’ way, might be used in a variety of joint tasks.

CONCLUSION

The aim of this article was to provide an overview of the major cognitive, sensorimotor, affective, and cultural processes supporting joint action. Given the extent of the phenomena (from moving a sofa to playing in a musical ensemble) as well as the variety of coordination mechanisms underlying joint action (as introduced in this review), we postulate that research on joint action needs to acknowledge the complex and variable nature of this social phenomenon. Consequently, future psychological, cognitive, and neuroscientific research might (1) integrate different lines of research in ecologically valid tasks, (2) specify the relative contribution of particular coordination mechanisms and contextual factors, and (3) set the grounds for an overarching framework that explains how co-actors plan and perform joint actions.

AUTHOR CONTRIBUTIONS

All authors contributed to the writing. CV developed the article structure and performed final editing of the text.

FUNDING

This work was supported by the European Union through the H2020 FET Proactive project socSMCs (GA no. 641321).

REFERENCES

- Aglioti, S. M., Cesari, P., Romani, M., and Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nat. Neurosci.* 11, 1109–1116. doi: 10.1038/nn.2182
- Bach, P., Nicholson, T., and Hudson, M. (2014). The affordance-matching hypothesis: how objects guide action understanding and prediction. *Front. Hum. Neurosci.* 8:254. doi: 10.3389/fnhum.2014.00254
- Bastiaansen, J. A., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2391–2404. doi: 10.1098/rstb.2009.0058
- Becchio, C., Bertone, C., and Castiello, U. (2008). How the gaze of others influences object processing. *Trends Cogn. Sci.* 12, 254–258. doi: 10.1016/j.tics.2008.04.005
- Bidet-Caulet, A., Voisin, J., Bertrand, O., and Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage* 28, 132–139. doi: 10.1016/j.neuroimage.2005.06.018
- Blakemore, S.-J., and Decety, J. (2001). From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* 2, 561–567.
- Böckler, A., Knoblich, G., and Sebanz, N. (2012). Effects of a coactor's focus of attention on task performance. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1404–1415. doi: 10.1037/a0027523
- Borgomaneri, S., Gazzola, V., and Avenanti, A. (2012). Motor mapping of implied actions during perception of emotional body language. *Brain Stimul.* 5, 70–76. doi: 10.1016/j.brs.2012.03.011
- Boyd, R., and Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3281–3288. doi: 10.1098/rstb.2009.0134
- Bratman, M. E. (1992). Shared cooperative activity. *Philos. Rev.* 101, 327–341. doi: 10.2307/2185537
- Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., and Zelinsky, G. J. (2008). Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition* 106, 1465–1477. doi: 10.1016/j.cognition.2007.05.012
- Brennan, S. E., Galati, A., and Kuhlen, A. K. (2010). Two minds, one dialog: coordinating speaking and understanding. *Psychol. Learn. Motiv. Adv. Res. Theory* 53, 301–344. doi: 10.1016/S0079-7421(10)53008-1
- Butterfill, S. (2012). Joint action and development. *Philos. Q.* 62, 23–47. doi: 10.1111/j.1467-9213.2011.00005.x
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531. doi: 10.1016/j.cub.2007.08.006
- Cheon, B. K., Im, D. M., Harada, T., Kim, J. S., Mathur, V. A., Scimeca, J. M., et al. (2011). Cultural influences on neural basis of intergroup empathy. *Neuroimage* 57, 642–650. doi: 10.1016/j.neuroimage.2011.04.031
- Ciaro, F., Lugli, L., Nicoletti, R., Rubichi, S., and Iani, C. (2016). Action-space coding in social contexts. *Sci. Rep.* 6:22673. doi: 10.1038/srep22673
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Costantini, M., Committeri, G., and Sinigaglia, C. (2011). Ready both to your and to my hands: mapping the action space of others. *PLoS ONE* 6:e17923. doi: 10.1371/journal.pone.0017923
- de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nat. Rev. Neurosci.* 7, 242–249. doi: 10.1038/nrn1872
- Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* 24, 581–604. doi: 10.1016/S0149-7634(00)00025-7
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Eskenazi, T., Doerrfeld, A., Logan, G. D., Knoblich, G., and Sebanz, N. (2014). Your words are my words: effects of acting together on encoding. *Q. J. Exp. Psychol.* 66, 1026–1034. doi: 10.1080/17470218.2012.725058

ACKNOWLEDGMENT

We thank Natalie Sebanz for helpful comments on an earlier version of the manuscript.

- Feldman, R., Weller, A., Sirota, L., and Eidelman, A. I. (2003). Testing a family intervention hypothesis: the contribution of mother-infant skin-to-skin contact (kangaroo care) to family interaction, proximity, and touch. *J. Fam. Psychol.* 17, 94–107. doi: 10.1037/0893-3200.17.1.94
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton, Mifflin and Company.
- Gudykunst, W. B., Ting-Toomey, S., and Chua, E. (1988). *Culture and Interpersonal Communication*. Newbury Park, CA: Sage Publications.
- Hertenstein, M. J., Holmes, R., McCullough, M., and Keltner, D. (2009). The communication of emotion via touch. *Emotion* 9, 566–573. doi: 10.1037/a0016108
- Hommel, B., Colzato, L. S., and Van Den Wildenberg, W. P. (2009). How social are task representations? *Psychol. Sci.* 20, 794–798. doi: 10.1111/j.1467-9280.2009.02367.x
- Hudson, M., Nicholson, T., Ellis, R., and Bach, P. (2016a). I see what you say: prior knowledge of other's goals automatically biases the perception of their actions. *Cognition* 146, 245–250. doi: 10.1016/j.cognition.2015.09.021
- Hudson, M., Nicholson, T., Simpson, W. A., Ellis, R., and Bach, P. (2016b). One step ahead: the perceived kinematics of others' actions are biased toward expected goals. *J. Exp. Psychol. Gen.* 145, 1–7. doi: 10.1037/xge0000126
- Innocenti, A., De Stefani, E., Bernardi, N. F., Campione, G. C., and Gentilucci, M. (2012). Gaze direction and request gesture in social interactions. *PLoS ONE* 7:e36390. doi: 10.1371/journal.pone.0036390
- Ishihara, R. W., Richardson, M. J., Carello, C., Baron, R. M., and Marsh, K. L. (2010). 'Affording cooperation: embodied constraints, dynamics, and action-scaled invariance in joint lifting. *Psychon. Bull. Rev.* 17, 342–347. doi: 10.3758/PBR.17.3.342
- Keller, P. E. (2012). Mental imagery in music performance: underlying mechanisms and potential benefits. *Ann. N. Y. Acad. Sci.* 1252, 206–213. doi: 10.1111/j.1749-6632.2011.06439.x
- Kimmel, M. (2009). Intersubjectivity at close quarters: how dancers of Tango Argentino use imagery for interaction and improvisation. *Cogn. Semiot.* 4, 76–124. doi: 10.1515/cogsem.2009.4.1.76
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848. doi: 10.1126/science.1070311
- Kourtit, D., Knoblich, G., Wozniak, M., and Sebanz, N. (2014). Attention allocation and task representation during joint action planning. *J. Cogn. Neurosci.* 26, 2275–2286. doi: 10.1162/jocn_a_00634
- Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action Representation of Sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* 27, 308–314. doi: 10.1523/JNEUROSCI.4822-06.2007
- Loehr, J. D., Kourtit, D., Vesper, C., Sebanz, N., and Knoblich, G. (2013). Monitoring individual and joint action outcomes in duet music performance. *J. Cogn. Neurosci.* 25, 1049–1061. doi: 10.1162/jocn_a_00388
- Loehr, J. D., and Vesper, C. (2016). The sound of you and me: novices represent shared goals in joint action. *Q. J. Exp. Psychol.* 69, 535–547. doi: 10.1080/17470218.2015.1061029
- Mark, L. S. (2007). Perceiving the actions of other people. *Ecol. Psychol.* 19, 107–136.
- Michael, J. (2011). Shared emotions and joint action. *Rev. Philos. Psychol.* 2, 355–373. doi: 10.1007/s13164-011-0055-2
- Mulligan, D., Lohse, K. R., and Hodges, N. J. (2016). An action-incongruent secondary task modulates prediction accuracy in experienced performers: evidence for motor simulation. *Psychol. Res.* 80, 496–509. doi: 10.1007/s00426-015-0672-y
- Newman-Norlund, R. D., Ganesh, S., van Schie, H. T., Bruijn, E. R. A. D., and Bekkering, H. (2009). Self-identification and empathy modulate error-related brain activity during the observation of penalty shots between friend and foe. *Soc. Cogn. Affect. Neurosci.* 4, 10–22. doi: 10.1093/scan/nsn028

- Novembre, G., Ticini, L. F., Schütz-Bosbach, S., and Keller, P. E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cereb. Cortex* 22, 2894–2903. doi: 10.1093/cercor/bhr364
- Pezzulo, G., Donnarumma, F., and Dindo, H. (2013). Human sensorimotor communication: a theory of signaling in online social interactions. *PLoS ONE* 8:e79876. doi: 10.1371/journal.pone.0079876
- Poljac, E., van Schie, H. T., and Bekkering, H. (2009). Understanding the flexibility of action–perception coupling. *Psychol. Res.* 73, 578–586. doi: 10.1007/s00426-009-0238-y
- Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154. doi: 10.1080/713752551
- Radke, S., de Lange, F., Ullsperger, M., and de Bruijn, E. (2011). Mistakes that affect others: an fMRI study on processing of own errors in a social context. *Exp. Brain Res.* 211, 405–413. doi: 10.1007/s00221-011-2677-0
- Ramenzoni, V. C., Riley, M. A., Shockley, K., and Davis, T. (2008). Carrying the height of the world on your ankles: encumbering observers reduces estimates of how high an actor can jump. *Q. J. Exp. Psychol.* 61, 1487–1495. doi: 10.1080/17470210802100073
- Repp, B. H., and Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *Q. J. Exp. Psychol.* 57, 499–521. doi: 10.1080/02724980343000369
- Richter, L. M. (1995). Are early adult–infant interactions universal? A South African view. *Southern Afr. J. Child Adolesc. Psychiatry* 7, 2–18.
- Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature* 11, 264–274. doi: 10.1038/nrn2805
- Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., and Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Exp. Brain Res.* 226, 473–486. doi: 10.1007/s00221-013-3459-7
- Sartori, L., Becchio, C., and Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition* 119, 242–252. doi: 10.1016/j.cognition.2011.01.014
- Sartori, L., and Betti, S. (2015). Complementary actions. *Front. Psychol.* 6:557. doi: 10.3389/fpsyg.2015.00557
- Schelling, T. C. (1960). *The Strategy of Conflict*. Cambridge: Harvard University Press.
- Schmidt, R. C., and Richardson, M. J. (2008). “Coordination: neural, behavioral and social dynamics,” in *Dynamics of Interpersonal Coordination*, eds A. Fuchs and V. K. Jirsa (Berlin: Springer-Verlag), 281–308.
- Schmitz, G., and Effenberg, A. O. (2012). “Perceptual effects of auditory information about own and other movements,” in *Proceedings of the 18th Annual Conference on Auditory Display*, Atlanta, GA, 89–94.
- Schmitz, G., Mohammadi, B., Hammer, A., Heldmann, M., Samii, A., Münte, T. F., et al. (2013). Observation of sonified movements engages a basal ganglia frontocortical network. *BMC Neurosci.* 14:32.
- Scorolli, C., Miatton, M., Wheaton, L. A., and Borghi, A. M. (2014). I give you a cup, I get a cup: a kinematic study on social intention. *Neuropsychologia* 57, 196–204. doi: 10.1016/j.neuropsychologia.2014.03.006
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76. doi: 10.1016/j.tics.2005.12.009
- Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others’ actions: just like one’s own? *Cognition* 88, B11–B21. doi: 10.1016/S0010-0277(03)00043-X
- Sebanz, N., Knoblich, G., and Prinz, W. (2005). How two share a task: corepresenting stimulus–response mappings. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 1234–1246.
- Skewes, J., Skewes, L., Michael, J., and Konvalinka, I. (2015). Synchronised and complementary coordination mechanisms in an asymmetric joint aiming task. *Exp. Brain Res.* 233, 551–565. doi: 10.1007/s00221-014-4135-2
- Smith, S. J. (2014). A pedagogy of vital contact. *J. Dance Somat. Pract.* 6, 233–246. doi: 10.1386/jdsp.6.2.233_1
- Sowden, S., and Catmur, C. (2015). The Role of the Right Temporoparietal Junction in the Control of Imitation. *Cereb. Cortex* 25, 1107–1113. doi: 10.1093/cercor/bht306
- Springer, A., Brandstädter, S., Liepelt, R., Birngruber, T., Giese, M., Mechsner, F., et al. (2011). Motor execution affects action prediction. *Brain Cogn.* 76, 26–36. doi: 10.1016/j.bandc.2011.03.007
- Stoffregen, T. A., Gorday, K. M., Sheng, Y. Y., and Flynn, S. B. (1999). Perceiving affordances for another person’s actions. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 120–136. doi: 10.1037/0096-1523.25.1.120
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–735. doi: 10.1017/S0140525X05000129
- Tomeo, E., Cesari, P., Aglioti, S. M., and Urgesi, C. (2012). Fooling the kickers but not the goalkeepers: behavioral and neurophysiological correlates of fake action detection in soccer. *Cereb. Cortex* 23, 2765–2778. doi: 10.1093/cercor/bhs279
- Ubaldo, S., Barchiesi, G., and Cattaneo, L. (2015). Bottom-up and top-down visuomotor responses to action observation. *Cereb. Cortex* 25, 1032–1041. doi: 10.1093/cercor/bht295
- van der Wel, R. P. R. D., Knoblich, G., and Sebanz, N. (2011). Let the force be with us: dyads exploit haptic coupling for coordination. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1420–1431. doi: 10.1037/a0022337
- van Schie, H. T., Mars, R. B., Coles, M. G. H., and Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nat. Neurosci.* 7, 549–554. doi: 10.1038/nn1239
- van Schie, H. T., van Waterschoot, B. M., and Bekkering, H. (2008). Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1493–1500. doi: 10.1037/a0011750
- Vesper, C., Butterfill, S., Knoblich, G., and Sebanz, N. (2010). A minimal architecture for joint action. *Neural Netw.* 23, 998–1003. doi: 10.1016/j.neunet.2010.06.002
- Vesper, C., Knoblich, G., and Sebanz, N. (2014). Our actions in my mind: motor imagery of joint action. *Neuropsychologia* 55, 115–121. doi: 10.1016/j.neuropsychologia.2013.05.024
- Vesper, C., and Richardson, M. (2014). Strategic communication and behavioral coupling in asymmetric joint action. *Exp. Brain Res.* 232, 2945–2956. doi: 10.1007/s00221-014-3982-1
- Vesper, C., van der Wel, R., Knoblich, G., and Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Exp. Brain Res.* 211, 517–530. doi: 10.1007/s00221-011-2706-z
- Vesper, C., van der Wel, R., Knoblich, G., and Sebanz, N. (2013). Are you ready to jump? Predictive mechanisms in interpersonal coordination. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 48–61.
- Wahn, B., and König, P. (2015). Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. *Front. Psychol.* 6:1084. doi: 10.3389/fpsyg.2015.01084
- Wahn, B., and König, P. (2016). Attentional resource allocation in visuotactile processing depends on the task, but optimal visuotactile integration does not depend on attentional resources. *Front. Integr. Neurosci.* 10:13. doi: 10.3389/fnint.2016.00013
- Wahn, B., Schwandt, J., Krüger, M., Crafa, D., Nunnendorf, V., and König, P. (2015). Multisensory teamwork: using a tactile or an auditory display to exchange gaze information improves performance in joint visual search. *Ergonomics* 59, 1–15. doi: 10.1080/00140139.2015.1099742
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460
- Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and interaction. *Philos. Trans. R. Soc. Lond. B* 358, 593–602. doi: 10.1098/rstb.2002.1238

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

Copyright © 2017 Vesper, Abramova, Bütepage, Ciardo, Crossey, Effenberg, Hristova, Karlinsky, McEllin, Nijssen, Schmitz and Wahn. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.