Topics in Cognitive Science 1 (2009) 255–259

ISSN: 1756-8757 print / 1756-8765 online
DOI: 10.1111/j.1756-8765.2009.01017.x

Joint Action: Current Perspectives

Bruno Galantucci,[a,*] Natalie Sebanz[b,*]

[a]*Yeshiva University and Haskins Laboratories*
[b]*Radboud University Nijmegen*

Received 4 February 2009; received in revised form 9 February 2009; accepted 10 February 2009

Abstract

In recent years researchers have begun to investigate how the perceptual, motor and cognitive activities of two or more individuals become organized into coordinated action. In the first part of this introduction we identify three common threads among the ten papers of this special issue that exemplify this new line of research. First, all of the papers are grounded in the experimental study of online interactions between two or more individuals. Second, albeit at different levels of analysis, the contributions focus on the mechanisms supporting joint action. Third, many of the papers investigate empirically the pre-requisites for the highly sophisticated forms of joint action that are typical of humans. In the second part of the introduction, we summarize each of the papers, highlighting more specific connections among them.

Keywords: Joint action; Social cognition; Distributed cognition; Interpersonal coordination

Although the study of joint action is an established field in language research (Clark, 1996), it is not until recent years that other disciplines have begun to investigate how the perceptual, motor, and cognitive activities of two or more individuals become organized into coherently coordinated action. This new line of research is rapidly gaining momentum, and progress has recently been made by researchers studying perception–action links (Fowler, Richardson, Marsh, & Shockley, 2008; Knoblich & Sebanz, 2006; Newman-Norlund, van Schie, van Zuijlen, & Bekkering, 2007; Richardson, Dale, & Kirkham, 2007), social cognition in animals and infants (Tomasello, Carpenter, Call, Behne, & Moll, 2005), as well as human communication (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008; Galantucci, 2005; Keysar, 2007; Pickering & Garrod, 2004). However, crosstalk among the scientists who are contributing to the development of this research has only just begun (Sebanz, Bekkering, & Knoblich, 2006). This section of *topiCS* provides a forum for the emerging crosstalk to grow.

*The authors contributed equally to the introduction and to the editing of the special issue and are listed in alphabetical order.

The section comprises 10 papers that provide a sample of current empirical research into joint action, ranging from perception–action couplings to higher-level cognitive processes such as language and theory of mind. Several of the papers can be traced back to the second Joint Action Meeting (JAM2), held at Rutgers University, Newark, in 2007. This meeting provided an opportunity for a diverse group of researchers to exchange their views on joint action and made many of us realize how much we had in common, despite differences in theoretical approach. Beyond their common focus, the papers in the special issue are interconnected in several other ways.

At a most general level, all of the papers in the issue take a social perspective on cognition. That is, they all dispense with the long-held assumption in cognitive science that perception, action, and cognition can be fully understood by investigating single individuals. Some authors continue along the lines proposed earlier by distributed cognition approaches (Hutchins, 1995) by considering individuals acting together as the primary unit of analysis. Other authors focus on cognitive and neural processes within the boundaries of individual minds, but they acknowledge social interaction as a major driving force in shaping these processes. The benefits that may be gained by dispensing with the assumption that cognition can be understood by focusing on single individuals (cf. Roepstorff & Frith, 2004) can be appreciated by considering a recent finding from developmental psychology. Ever since Piaget (1954) first reported perseverative search errors in infants (sometime also referred to as A-not-B errors), explanations for such errors focused on the infant, neglecting the communication context within which the infant performed the task. Manipulating the social context, Topal, Gergely, Miklosi, Erdohegyi, and Csibra (2008) discovered that infants' error rates were substantially reduced when the experimenter did not communicate with the infant, suggesting that subtle communicative cues can play a central role in creating interpretative biases in infants. In a similar vein, the contributions in this volume suggest that we can make progress in understanding cognition by considering the immediate social context within which it occurs. To be sure, the presence of social influences on cognition has been long recognized (Vygotsky, 1978) and studied (Levine & Resnick, 1993; Smith & Semin, 2004). However, little research has directly studied the online interaction between two or more individuals. The papers presented here are all grounded in the experimental study of such online interactions.

The papers in the issue are also connected because, albeit at different levels of analysis, they all investigate the *mechanisms* supporting joint action. In particular, to which extent joint action relies exclusively on low-level mechanisms such as direct perception–action links and to which extent it requires higher-level cognitive processes including memory and theory of mind is a central question in several contributions. This question is related to the ongoing debate about the role of embodiment in cognition (Clark, 1997). How far can we get in explaining joint action with sensorimotor processes alone? To what extent is joint action grounded in sensorimotor functions? How do higher- and lower-level processes work together to support joint action? As we shall see, the answers provided by the authors in this issue differ substantially, covering a large part of the range of possible answers. However, the authors seem to mostly agree on the fact that relying on perception–action links or higher-level cognition alone is not going to provide the answer. Rather, the challenge is to

specify to which extent and in which circumstances different kinds of processes work together to enable our ability to act jointly in flexible ways.

Furthermore, a number of papers in the issue are connected because they speculate about the *prerequisites* for the highly sophisticated forms of joint action that are typical of humans. Some contributions address this issue explicitly (Call, Carpenter, Galantucci), others touch on it more implicitly, by virtue of isolating particular components of joint action in tightly controlled experimental settings (e.g., Bekkering et al.; Marsh, Richardson, & Schmidt; Shockley, Richardson, & Dale; Sebanz & Knoblich). As can be expected, the perspectives provided by the authors in the issue differ substantially, ranging from the claim that human joint action might require little, if any, special prerequisite (Marsh, Richardson, & Schmidt; Shockley, Richardson, & Dale) to the claim that what might be special about humans is the desire to share psychological states (Call; Carpenter). However, all seem to agree on the fact that speculations about prerequisites can and should be grounded in experimental research, even when this implies investigating rough approximations of the phenomena the authors wish to study. To give some examples, Galantucci approximates the multimodal forms of joint action that are typical of human communication with interactions which occur in a purely visual virtual world. Marsh and colleagues, as well as Shockley and colleagues, approximate human social behavior with the behavior of physical systems. Call and Carpenter approximate joint actions involving two adult humans with activities that require limited levels of enculturation, such as the joint actions that are possible with infants and nonhuman primates. These approximations offer important opportunities for developing new theories about joint action and new experimental paradigms to test them.

In the remaining part of this introduction, we provide a brief synopsis of the papers in the issue.

The first three papers originate from the well-established line of research into joint action that focuses on human communication. The papers by Shintel and Keysar and by Brennan and Hanna are connected through a question that has received much attention in the last few years: Does effective communication require processing information related to the mental contents of the people engaged in the conversation? Shintel and Keysar argue that the answer is something like ''most often it does not, and when it does, it does so relatively late in time'' and point to the fact that effective communication can be supported by many cues that do not require access to mental contents. Brennan and Hanna argue that the answer to the question is rather ''often it does, and when it does, it does so relatively early in time'' and point to the fact that processing information about other people's mental contents can be fairly effortless.

Garrod and Pickering propose a model of communication that implies a similar answer as that provided by Shintel and Keysar. However, the major focus of their paper is on the lower-level mechanisms that must be in place for effective communication to occur. This theme is also at the center of the following contribution by Shockley, Richardson, and Dale. Garrod and Pickering start from high-level constraints on communicative goals and ask which low-level mechanisms support them. Their answer is that the core low-level mechanism necessary for effective communication is predictive emulation, which, in turn, implies covert imitation. Shockley and colleagues start instead from low-level constraints on

perception–action couplings and ask which higher-level organizations can emerge from them. Their answer is that human interactions self-organize via the same dynamical principles that are behind physical instances of self-organization such as coordination in human movements.

The following three contributions further expand this discussion by exploring how direct links between perception and action (Marsh, Richardson, & Schmidt) or shared mental representations for one's own and others' actions (Bekkering et al.; Sebanz & Knoblich) contribute to predominantly nonverbal forms of joint action, such as lifting objects together. In line with Shockley et al., Marsh, Richardson, and Schmidt argue that many forms of social interaction and joint action are best understood as emerging through dynamical principles that hold not just within but also across individuals. In this view, connections between humans arise through synchronization of action patterns, rather than, for instance, mental state attribution. While keeping the focus on action, the contribution by Bekkering, de Bruijn, Cuijpers, Newman-Norlund, van Schie, and Meulenbrok addresses the nature of individual cognitive and brain processes in the service of joint action. The authors argue that action simulation based on one's own action repertoire, as well as action monitoring, and action selection are critical component processes of joint action that have identifiable brain bases. The following contribution by Sebanz and Knoblich highlights the necessity to predict others' actions as a core component of joint action. The authors discuss how spatial, temporal, and goal-related predictions about others' actions can be made during joint action by relying on one's own motor system.

The next two papers focus on the origins of human joint action. From a phylogenetic point of view, Call argues that human joint action requires not only an understanding of other creatures' mental processes—which a number of primates exhibit—but also a motivation to share emotions and experiences with others—which seems to be restricted to humans. From an ontogenetic point of view, Carpenter argues that the capability to engage in sophisticated forms of joint action is a fundamental human ability, and demonstrates that it is already fairly developed by the time infants reach their first year of life. Finally, the paper by Galantucci focuses on how the capabilities illustrated by Call and Carpenter can lead, after a history of dyadic interactions, to cultural products that closely resemble human language. In doing so, Galantucci highlights the circumstances under which basic perception–action processes convert into effective communicative behaviors as well as the circumstances under which the conversion does not occur.

Even though the collection of articles of this topic covers a broad range of theoretical perspectives and a large variety of methodological approaches, interested readers should be warned that our coverage omits relevant perspectives provided by recent work in robotics and philosophy, among other domains. Many of the issues raised here can also be found in these fields. For instance, work in robotics is concerned with specifying the prerequisites for joint action (e.g., Breazeal & Scassellati, 2002; Erlhagen, Mukovskiy, Chersi, & Bicho, 2007; Steels, 2003). Likewise, philosophers are developing new perspectives on the question of how joint intentions are formed, and what they entail (e.g., Pacherie & Dokic, 2006; Tollefsen, 2005). This is closely related to the question of the role of higher-level cognition in joint action. At present, we cannot do more than point out these connections, but we hope

that the beginning crosstalk with researchers in these and other disciplines will pick up further, perhaps during future Joint Action Meetings.

References

Breazeal, C., & Scassellati, B. (2002). Robots that imitate humans. *Trends in Cognitive Sciences*, *6*(11), 481–487.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, *106*(3), 1465–1477.

Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press.

Erlhagen, W., Mukovskiy, A., Chersi, F., & Bicho, E. (2007). On the development of intention understanding for joint action tasks. In *Proceedings of the 6th IEEE international conference on development and learning (ICDL 07)* (pp. 140–145). New York: Imperial College London.

Fowler, C. A., Richardson, M. J., Marsh, K. L., & Shockley, K. (2008). Language use, coordination, and the emergence of cooperative action. In A. Fuchs & V. Jirsa (Eds.), *Coordination: Neural, behavioral and social dynamics*. New York: Springer.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: The MIT Press.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, *4*(1), 71–84.

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*, 99–104.

Levine, J. M., & Resnick, L. B. (1993). Social foundations of cognition. *Annual Review of Psychology*, *44*, 585–612.

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*, *10*, 817–818.

Pacherie, E., & Dokic, J. (2006). From mirror neurons to joint actions. *Cognitive Systems Research*, *7*(2–3), 101–112.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–226.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*, 407–413.

Roepstorff, A., & Frith, C. D. (2004). What's at the top in the top-down control of action? Script-sharing and 'top-top' control of action in cognitive experiments. *Psychological Research*, *68*(2–3), 189–198.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76.

Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, *36*, 53–117.

Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, *7*(7), 308–312.

Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences*, *35*(1), 75–97.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–735.

Topal, J., Gergely, G., Miklosi, A., Erdohegyi, A., & Csibra, G. (2008). Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science*, *321*(5897), 1831–1834.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.