velopmental impairments in autism emphasize that for individuals with autism there appears to be a reduced salience of social signals and social rewards (Dawson et al. 2002b; Mundy et al. 2003). This might reflect either a reduced *capacity* to recognize, identify, think, and respond to social signals; alternatively, it might explain how over development individuals with autism have a reduced *motivation* to respond to social signals they fail to find rewarding. As Tomasello et al. highlight, the inevitable consequence of such impairments is an inability to enter fully into human culture.

Toward a construction-based account of shared intentions in social cognition

Peter F. Dominey

Institut des Sciences Cognitives, CNRS UMR 5015, 69675 Bron Cedex, France. **dominey@isc.cnrs.fr**
http://www.isc.cnrs.fr/dom/dommenu-en.htm

Abstract: This commentary analyzes the target article to determine whether shared-intention development could be implemented and tested in robotic systems. The analysis indicates that such an implementation should be feasible and will likely rely on a construction-based approach similar to that employed in the construction grammar framework.

The theoretical framework presented by Tomasello et al. is of particular interest for researchers in the domain of cognitive robotics, whose goals include the development of artificial robotic systems that are endowed with the social cognition capabilities described in the target article. I analyze the article to begin to formalize the internal representations and operations on those representations (e.g. transformations, combinations) in order to determine how they might provide the basis for implementation of an artificial system that could perform like young infants in the development and use of dialogic cognitive representations. Interestingly, one of the outcomes of this analysis is the suggestion that the notion of *construction* as a structured mapping in the construction grammar framework for language can be generalized or extended to the domain of social cognition.

It is assumed that the child comes to the task with a perceptual capability that provides a set of basic representations and operations on those representations. Thus, the child can (1) perceive physical states of objects, (2) perceive physical actions performed by agents that can change the states of these objects, and (3) distinguish between self and other in this context. Likewise the child can (4) perceive the emotional responses of others. In addition, we will assume (5) some form of statistical learning mechanism (e.g., see Saffran et al. 1996) that forms representations that bind together co-occurring and successive events and that can represent variables that can be instantiated in these representations (Dominey & Ramus 2000; Marcus et al. 1999).

Given this machinery, we can consider the ontogenetic development of a form of action construction in which, after repeatedly performing an action (e.g., picking up a block and throwing it), the child will begin to form a representation of the form [initial state, action, and final state] instantiated as [block in hand, throw block, and block on floor]. Although the block may start and end in different locations, and the throwing trajectory may be different, the invariant sequence will be repeated enough times to allow statistical learning to take place, linking the initial state, action, and final state into a coherent cognitive object: an action construction. Interestingly, an analogy with grammatical constructions and the transition in language from fixed holoconstruction to single-argument or verb-island constructions appears here. The throw-block construction is like a fixed holophrase in language: it has a highly limited and specific scope of application. As the child comes to throw different objects, the throw-block construction can become slightly more abstract as the throw(X) construction [in hand(X), throw(X), on floor(X)], analogous to the verb-island constructions in language in a progressive development of more abstract constructions. This introduction of variables and their manipulation could rely on the learning mechanism demonstrated in infants by Marcus et al. (1999) and simulated by Dominey and Ramus (2000) and Dominey et al. (1998).

The resulting relation between initial state, action, and final state enables the development of the intentional construction based on the action construction. In the intentional construction, the final state is the goal, and the initial state and action constitute the intention associated with realizing the goal. These intentional constructions will contribute to the "knowledge, skills" component of Figures 1 and 2 of the target article, and when a goal arises, the construction inventory can be searched to find the appropriate construction and extract the corresponding intention.

We can now consider how this framework might extend to the formation of dialogic cognitive representations. Consider a scenario in which a child is seated on the floor and is reaching for, but failing to reach, a toy block. An adult arrives, sees that the child wants the block but cannot reach it, and so pushes the block to within the child's reach. The child then successfully reaches for and grasps the block, and throws it out of its own reach, and the scenario repeats in a "throw the block" game between the child and the adult enough times so that statistical learning can take place.

In this scenario, the child has a goal of grasping and throwing the block. It has already learned an intentional construction [initial state, action, and final goal-state] of the form [within reach(X), grasp(X), and in hand(X)], and so knows [by chaining the throw(X) and grasp(X) constructions from the construction inventory] that in order to throw the block it must first grasp it. The child attempts to grasp the block but fails because it is out of reach. This creates a subgoal (getting the block within reach) for which there is currently no appropriate action. However, when the adult moves the toy within reach, it becomes graspable, and the intentional construction plan (grasping and throwing the block) can then be executed. Sufficient repetition of the "game" will enable learning mechanisms to bind states and actions together. Interestingly, however, the intentional representation will include actions by both participants. In particular, the subgoal of getting the toy back into reach after it has been thrown is assigned to the adult. More generally, intentional constructions that form part of an intentional plan can be assigned to different participants in a distributed manner, with each participant maintaining his/her representation of the complete distributed plan: the joint intention. Thus, we see mechanistically how joint intentions will be formed in both participants through their participation and observation of the unrolling of actions in a goal-directed manner, with the goal and the execution of the actions shared by the two participants.

The resulting joint intentional representation is a precursor of more articulated dialogic cognitive representations, but still provides a basis for quite interesting interaction. For example, based on the notion of *help* as intervention in an intentional plan by a collaborator (e.g., the adult in the foregoing scenario), the child can learn to request help and to help others under appropriate conditions. Likewise, the notion of *game* naturally develops from this type of representation. This provides both a powerful analytical tool for human behavior and an important step toward the specification of social cognition mechanisms for cognitive robots (Dominey & Boucher 2005).

The current preliminary analysis suggests that the characterization of the mechanisms for shared intentions in the target article is of sufficient detail to enable their implementation in a formalized manner that can eventually be used and tested in the context of artificial systems that interact with humans and/or other systems. In addition, it appears that the functional/computational mechanisms supporting these capabilities can be encompassed within the structure-mapping framework that has been of particular use in the construction grammar approach to natural language (see Dominey 2004; Dominey & Boucher 2005). Indeed, this is not surprising, because language and social cognition are highly interrelated.

https://doi.org/10.1017/S0140525X05270127 Published online by Cambridge University Press