

Infants' attribution of goals and referential intentions to a robot: Cue-based or knowledge-based?

Diane Poulin-Dubois
Concordia University
diane.pouлиндubois@concordia.ca

Adults and even human infants readily interpret the actions of others as goal-directed, but the developmental origins of this important precursor of a theory of mind is hotly debated. According to one view, infants acquire an understanding of actions as goal-directed through gradual experience with particular actions of human agents (Poulin-Dubois, 1999, Tomasello, 1999; Woodward, 1998). An alternative interpretation for the understanding of goal-directed actions holds that infants' reasoning about actions as goal-directed is based on innate specialized systems that are sensitive to abstract behavioral cues and attribute goals to a wide range of entities (Csibra & Gergely, 1998). Such cues include morphological features, biological motion, self-propulsion, and contingent interactions. The first view predicts that young infants should attribute goals to humans but not other agents; the second view predicts that young infants should attribute goals to both human and nonhuman agents. Our main hypothesis is that self-propulsion and morphological features are sufficient for infants to establish joint attention but that mental states will not be attributed to inanimate agents on the basis of these cues. We will present the results of two experiments in which infants had to monitor the goal of a humanoid, self-propelled robot. In the first experiment, 10-month-olds were tested using the infant-controlled habituation paradigm previously designed to test goal attribution to human agents (Woodward, 2003). After being habituated to a robot looking at one of two objects, the location of the two objects was switched and infants saw the robot look at the other object (new goal) or at the familiar object (old goal). (see Figure 1). The results of this experiment indicated that infants did not detect change in the non-human agent's goal, as their looking time on the new goal trials ($M=10.5s$) was similar to their looking time on the old goal trials ($M=11.4s$), $t(26)=.38$, $p>.05$. These results suggest that although very young infants attribute goals to both self-propelled human and nonhuman agents, by 10 months of age the self-propulsion cue is no longer sufficient for goal attribution, even when the agent possesses salient human morphological features. In the second experiment, we tested the hypothesis that 18-month-old infants will not attribute referential intentions to a self-propelled, human-looking robot. Infants aged 18-months were exposed to two new words in either the discrepant or coordinated condition used by Baldwin (1993). In the coordinated condition, the speaker presented the label while looking at a novel toy on which the child was also focused. By contrast, in the discrepant condition, she presented the label while looking at a toy just as the child was focused on a

different toy. In the Nonhuman speaker condition, a programmable robot replaced the human speaker and a second experimenter was present. In both the discrepant and coordinated conditions, a novel word was uttered four times. Infants were subsequently asked comprehension questions by the human experimenter (i.e. "Where is the Dax?") on four test trials per word. The results indicated that infants ($N=26$) followed the robot's gaze during labeling in both conditions. As expected, when tested for comprehension of familiar words, infants selected the correct toy at above chance levels in both the coordinated ($M = 66.99\%$, $SD=28.43$, $t(25)=3.05$, $p<0.01$) and discrepant condition ($M = 65.71\%$, $SD=28.99$, $t(25)=2.76$, $p<0.05$). However, infants did not learn the novel words in either case (Coordinated, $M = 52.24\%$, $SD=32.62$, $t(25)=0.351$, $p>0.05$; Discrepant, $M = 50.64\%$, $SD=35.51$, $t(25)=0.092$, $p>0.05$), suggesting that the 18-month-olds did not attribute referential intentions to the non-human agent. Taken together, the present findings suggest that although goal-directed behaviors are attributed on the basis of behavioral cues such as self-propulsion in the first few months of life, referential intentions and goals are restricted to human agents by the end of the first year.



Figure 1. Experimental set-up for goal attribution task.

References

- Csibra, G. and Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, 1, 255-259.
- Poulin-Dubois, D. (1999). Infants' Distinction between Animate and Inanimate Objects: The Origins of Naive Psychology. In P. Rochat (Ed), *Early Social Cognition*. (pp. 257-280). Hillsdale, N.J.: Erlbaum.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1-34.