To appear in: M.R. Banaji & S.A. Gelman, *Navigating the Social World: What Infants, Chidren, and Other Species Can Teach Us.* Oxford University Press.

Teleological understanding of actions

Gergely Csibra and György Gergely

Central European University, Budapest

All organisms pursue goals that promote their survival and reproduction. For many species, the small repertoire of basic drive-related goals and behavioural means to achieve them are genetically specified and are triggered by appropriate releasing stimuli. Other species, most notably, but not exclusively, mammals and birds, have much more flexible ways to accomplish their goals, tailoring their behaviour to the local environment and adjusting it according to their past experience with similar situations. Humans also display adaptability and learning capacities in their goal-directed actions, but they are flexible in another way as well: they regularly pursue goals that hold only a remote connection to their survival and reproductive potential (Csibra & Gergely, 2006; Gergely, 2010). These include social goals (e.g., increasing one's reputation), culturally defined desires (e.g., enjoying heavy metal music), idiosyncratic aspirations (e.g., climbing Mount Everest), and in many cultures even the most basic survival needs, like consumption of food, are fulfilled by chains of goal-directed sequences composed of actions (e.g., sowing seeds, calling a restaurant to reserve a table) that are far detached from their biologically useful outcome.

Beyond pursuing goals, animals may also benefit from observing other individuals' behaviours because they may provide information (1) about what the observed agents would do next and (2) about the environment. Interpreting behaviours as goal-directed actions is

especially valuable as it carries direct information about likely future events (the expected outcome) and its context (e.g., potential food source). What kinds of information are available in a goal-directed action that would help observers to interpret it as a means directed at a certain end (Csibra & Gergely, 2007)? First, as in many domains, statistical information about the co-occurrence of observed behaviours and their outcomes could assist the formation of bidirectional action-effect associations. Such information may contribute to gradual learning about goal-directed actions, but does not help when an observer is confronted with novel instrumental behaviours. The second source of information for interpreting actions as goal-directed comes from the observer's own motor (or more generally, instrumental) competence. One can rely on the internal mechanisms of motor planning, which chain actions to bring about desired outcomes, to simulate observed actions and to link them to their likely goals. While such simulatory mechanisms play important roles in action prediction (Csibra, 2007), their use is restricted to actions that are within the observer's own motor competence, and are unavailable for interpretation of the behaviour of unfamiliar agents or novel actions of familiar actors.

We have proposed that a third source of information that, beyond temporal association and motor mediation, may link actions to goal states is the *efficiency* of the observed action to the observed or inferred goal in the given environment (Gergely & Csibra, 2003). Briefly, an observed behaviour is interpreted as an action directed to a particular end state if it is judged to be the most efficient means available to the agent for achieving this goal in the given environment. When such an interpretation is established, it creates a *teleological representation* of the action, which is held together by the *principle of efficiency*. The paradigmatic situation in which the functioning of teleological interpretation can be tested is when one observes a behaviour (an agents jumps into the air while moving in a certain direction) leading to an end state (the agent stops next to another object). If, and only if, the behaviour (jumping) is

justified by environmental factors (by the presence of a barrier over which the jumping occurs) will this behaviour be interpreted as a means action to achieve the end-state as the goal of the action (to get in contact with the other object) (Fig. 1). We and others have published extensive evidence that infants from at least 6 months of age form this kind of teleological representations of actions (e.g., Gergely et al., 1995; Csibra, 2008).

Instead of listing evidence and arguments that support the proposal that such a representational system operates early in human development, below we attempt to clarify commonly raised issues about this theory in a question-and-answer format.

***

Would it not be a simpler interpretation of these findings that infants expect agents to always follow the 'shortest pathway'?

It would, but it would be misleading for two reasons. First, such an interpretation hides, rather than reveals, that infants represent the observed actions as goal-directed. Whether or not a certain trajectory is 'the shortest pathway' can only be judged by reference to its end point ('shortest pathway' requires a 'to' argument). Thus, if infants form this expectation, it implies that infants interpret the observed movement as goal-directed. Second, 'shortest pathway' is not the only factor that infants can apply to judge the efficiency of actions. They think, for example, that action sequences that contain fewer steps are more efficient than those with more steps (Southgate et al., 2008), and that the shorter pathway is not necessarily better than a longer detour if it requires the agent to effortfully squeeze through a narrow aperture (Csibra & Gergely, 1998b; Gergely, 2003).

What is the advantage of forming such abstract representations of actions?

Teleological action interpretation enables at least three kinds of inferences to be made. First, it allows action predictions in situations in which the environment changes but the goal does not. In this case, one can predict that the agent would adjust its action to the modified environment, choosing the more efficient action alternative to the goal. In fact, this is exactly how goal attribution is tested in infant paradigms (e.g., Gergely et al., 1995; Csibra et al., 1999, 2003). Second, teleological action representations support the prediction and attribution of goals to ongoing actions even before such outcomes have been realized. Answering the question, 'Toward what end state is the observed behaviour an efficient action in this environment?' enables the observer to infer the goal of unfinished, partly hidden, or even failed actions. Infants can definitely make such inferences at 10 to 12 months of age (Brandone & Wellman, 2009; Csibra et al., 2003), even on the basis of the observation of a single action (Southgate & Csibra, 2009). Third, the assumption of efficiency can provide a basis for inferring situational constraints of an observed goal-directed action. For example, one-year-old infants infer the presence of an occluded barrier when an agent approaches its goal by a jumping action (Csibra et al., 2003). Thus, teleological action interpretation can be a productive process, which is especially important when a child observes novel instrumental actions.

To what kinds of agents do infants attribute goals via teleological interpretation?

Common sense dictates that the interpretation of goal-directedness should be reserved to behaviours of agents, like animals, who do pursue goals. However, feature-based identification of agency does not seem to constrain the application of teleological interpretation in infants. They readily attribute goals to people (Kamewari et al., 2005; Sodian et al., 2004), hands

(Biro et al., in press; Brandone & Wellman, 2009; Phillips & Wellman, 2005; Southgate et al., 2008; Woodward & Sommerville, 2000), puppets (Sodian et al., 2004), robots (Kamewari et al., 2005), boxes (Csibra, 2008), balls (Southgate & Csibra, 2009), and computer-animated two-dimensional shapes (Biro et al., 2007; Gergely et al., 1995; Csibra et al., 1999, 2003; Wagner & Carey, 2003). It may be true that teleological action interpretation is easier to elicit for familiar than for unfamiliar agents (Kamewari et al., 2005), but it is not denied of inanimate boxes even among the youngest age groups (Csibra, 2008). Another proposal suggested that goal attribution might be triggered only by actions of self-propelled agents (Premack, 1990; Leslie, 1994). We were unable to confirm this hypothesis in a study in which we found that infants interpreted the actions of an agent as goal-directed even if they had no evidence of how it started to move and neither were there other disambiguating cues available (such as biomechanical movement) to indicate animacy (Csibra et al., 1999). Indeed, the only agency cue that seems to be reliably linked to goal attribution is some kind of evidence of 'freedom' or 'choice' in terms of variability of the agent's behaviour (Csibra, 2008; see also Biro & Leslie, 2007; Luo, 2011).

Is teleological goal attribution a kind of mental attribution?

It does not have to be. An important insight of this proposal is that the computations that are necessary to set up teleological representations of actions are performed not on mental states but on real (present) and hypothetical (future) states of affairs. These states of affairs correspond to the *contents* of mental states (beliefs, desires, and intentions) that are involved in mentalistic action explanations (Csibra & Gergely, 1998a; Gergely & Csibra, 2003). We think that even adults do not normally appeal to mental states when they explain and predict others' actions, only when they have good reason to assume that the content of those states differs from real states of affairs (e.g., when the agent acts on false beliefs). When we take the

'intentional stance' (Dennett, 1987) and interpret an action by attributing causally efficacious mental states to the agent, the assumption of efficiency is not sufficient; we have to assume that the agent acts *rationally*, i.e., efficiently in the world described by her beliefs. Thus, the *principle of ration action* applies calculations of efficiency either on real states of affairs or on contents of mental states (that represent real, hypothetical, or counterfactual state of affairs). However, when we find evidence of teleological action understanding in infants, it does not tell us whether it is based on mental attribution or not.

So, does the teleological nature of early action understanding suggest that infants are unable to attribute representational mental states?

No. In fact, evidence suggests that even 7-month-old infants attribute representational mental states based on observed perceptual access (Kovács et al., 2010), and that they use this ability for the interpretation of goal-directed actions in the second year of life (Onishi and Baillargeon, 2005; Surian et al., 2007; Southgate et al., 2007). Although infants under one year of age can attribute belief states and evaluate the efficiency of goal-directed actions, it is not yet clear whether they can utilize the former ability in the service of the latter (but see Luo and Baillargeon, 2010). In fact, one way in which experience with goal-directed actions can contribute to the development of action understanding is by providing opportunities to learn to recognize the conditions under which taking into considerations representational mental states of the actor is necessary for correct action understanding. Note also that teleological action interpretation is not the only mechanism of action understanding: communicative-referential actions invite different types of inferences and representations (Csibra, 2003; Gergely, 2010).

Do infants have sufficient knowledge of agents to assess the efficiency of their actions?

In order to evaluate efficiency, observers will have to estimate and compare the effort (the cost) that an agent does or would invest in action alternatives. Such calculations require detailed knowledge of biomechanical factors that determine the motion capabilities and energy expenditure of agents. However, in the absence of such knowledge, one can appeal to heuristics that approximate the results of these calculations on the basis of knowledge in other domains that is certainly available to young infants. For example, the length of pathways can be assessed by geometrical calculations, taking also into account some physical factors (like the impenetrability of solid objects). Similarly, the fewer steps an action sequence takes, the less effort it might require, and so infants' numerical competence can also contribute to efficiency evaluation. As infants learn the biomechanical constraints of familiar agents, like humans, by observation and by experiencing their own movement capabilities, they gradually take into account this knowledge in interpreting observed actions. Until this knowledge base is well established, however, they may make the mistake of expecting human agents to perform an impossible action if it seems the most efficient by some heuristics borrowed from other domains (Southgate et al., 2008).

Is teleological interpretation the only mechanism of goal attribution available to infants?

No. As we mentioned in the beginning of this chapter, there are also other ways to link actions to goals. We hypothesize that infants always attempt to find teleological explanations (i.e., explanations in terms of goal states) for observed actions (a phenomenon that we termed 'teleological obsession', see Csibra & Gergely, 2007). If no reliable information of means adjustment is present in the action, they can use alternative information, like the agent's

apparent choice between two objects (Woodward, 1998). Recent studies have found that a goal that is attributed to an agent on the basis of efficiency of the means action is used by infants to predict which of two objects the agent would choose in a new situation (Biro et al., in press; Hernik & Southgate, submitted). This transfer of goal attribution suggests that whatever information infants use to infer the goal of the action, it feeds into a unitary goal concept (Biro et al., in press). However, when efficiency information is pitted against goal selection, infants tend to rely more on the former than on the latter (Verschoor & Biro, submitted).

Is teleological action understanding human-specific?

Since non-human animals, for example primates, attribute goals, we find it unlikely that they would not use such a simple mechanisms of action evaluation, and evidence supports that they do so (Rochat et al., 2008). However, we think that teleological action understanding is not as productive in other animals as in humans; not because their teleological action representation would not support such productive inferences, but because they lack the 'teleological obsession' with which humans, including human infants, approach actions. Thus, when no other cue indicates a potentially valuable goal (e.g., food) for an observed action, non-human animals do not necessarily start to engage in reasoning about what goal the action is directed to, while we suggest that human infants do so. The reason for this difference is that human infants make also use of teleological understanding of actions for social learning of instrumental actions and artefact functions (Csibra & Gergely, 2007).

Is teleological reasoning applied outside the domain of action perception?

Yes, it is. Humans tend to look for teleological explanations in many domains, most notably to interpret artefacts, social institutions, biological mechanisms, and even non-biological

phenomena (Kelemen, 1999, 2004). In all these cases the underlying intuition is that the explanation must satisfy some kind of optimality criterion: a tool should be better to achieve its desired effect than the available alternatives, biological mechanisms should be adaptive, etc. Whether or not these explanations draw on the same underlying cognitive mechanisms is not known. However, it is plausible to assume that action interpretation and artefact understanding are closely linked in development (Csibra & Gergely, 2007). Infants seek to find functions for artefacts just like they search for action goals (Futó et al., 2010; Träuble & Pauen, 2007), and in many cases the answers to these queries coincide: the function of a tool is the goal which can efficiently be accomplished by its use.

***

Teleological understanding provides a powerful mechanism of learning because it specifies formal criteria of organization of information into a specific representational system. As such, it allows the learning of completely novel action-goal relations from the simplest level of single motor acts to complex actions linking spatially and temporally distant events. It also forms one of the basic representational systems from which our intuitive mentalistic psychology is built, without which human social cognition would be quite different.

References

Biro, S., Verschoor, S.A., & Coenen, L. (in press). Evidence for a unitary goal concept in 12–month-old infants. *Developmental Science.*

Biro, S., Csibra, G., & Gergely, G. (2007). The role of behavioral cues in understanding goal-directed actions in infancy. *Progress in Brain Research, 164*, 303-322.

Biro, S. & Leslie, A.M. (2007). Infants' perception of goal-directed actions: Development

through cue-based bootstrapping. *Developmental Science, 10*, 379-398.

Brandone, A.C. & Wellman, H.M. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science, 20*, 85-91.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society, London B*, 358, 447-458.

Csibra, G. (2007). Action mirroring and action interpretation: An alternative account. In: P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition. Attention and Performance XXII* (pp. 435-459). Oxford: Oxford University Press.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*, 705-717.

Csibra, G., Bíró, S., Koós, S., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*, 111-133.

Csibra, G., & Gergely, G. (1998a). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*, 255-259.

Csibra, G., & Gergely, G. (1998b). Beyond least effort: The principle of rationality in teleological interpretation of action in 1-year-olds. Poster presented at the Swansong Conference of the Medical Research Council Cognitive Development Unit, London.

Csibra, G. & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI* (pp. 249-274). Oxford: Oxford University Press.

Csibra, G. & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica, 124*, 60-78.

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition*, *72*, 237-267.

Dennett, D. (1987). *The Intentional Stance*. MIT Press, Cambridge.

Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition, 117*, 1-8.

Gergely, G. (2003). What should a robot learn from an infant? Mechanisms of action interpretation and observational learning in infancy. *Connection Science, 13*, 191-209.

Gergely, G. (2010). Kinds of agents: The origins of understanding instrumental and communicative agency. In: U. Goshwami, (Ed.). *Blackwell Handbook of Childhood Cognitive Development. 2nd Edition* (pp. 76-105). Oxford: Blackwell Publishers.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The one-year-old's naïve theory of rational action. *Trends in Cognitive Sciences, 7*, 287-292.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165-193.

Hernik, M. & Southgate, V. (submitted). Nine-month-olds attribute goals to single-target events only if the action is efficiently related to the goal.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development, 20*, 303-320.

Kelemen, D. (1999). Functions, goals and intentions: Children's teleological reasoning about objects. *Trends in Cognitive Sciences, 12,* 461-468.

Kelemen, D. (2004). Are children "intuitive theists"?: Reasoning about purpose and design in nature. *Psychological Science, 15,* 295-301.

Kovács, Á.M., Téglás, E., & Endress, A.D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*, 1830-1834.

Leslie, A.M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity, In: Hirschfeld, L.A., Gelman, S.A. (Eds.), *Mapping the Mind: Domain Specificity in Cognition and*

Culture (pp. 119-148). Cambridge University Press.

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science, 14*, 453-460.

Luo, Y., & Baillargeon, R. (2010). Toward a mentalistic account of early psychological reasoning. *Current Directions in Psychological Science, 19*, 301-307.

Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255-258.

Phillips. A.T. & Wellman, H.M. (2005). Infants' understanding of object-directed actions. *Cognition, 98*, 137-155.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, *36*, 1-16.

Rochat, M.J., Serra, E., Fadiga, L., & Gallese, V. (2008). The evolution of social cognition: Goal familiarity shapes monkeys' action understanding, *Current Biology, 18, 227-232.*

Sodian, B., Schoeppner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents? *Infant Behavior and Development, 27*, 31-41.

Southgate, V. & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology, 45*, 1794-1798.

Southgate, V., Johnson, M.H., & Csibra, G. (2008). Infants attribute goals to even biologically impossible actions. *Cognition, 107*, 1059-1069.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by two-year-olds. *Psychological Science, 18*, 587-592.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18*, 580-586.

Träuble, B. & Pauen, S. (2007). The role of functional information for infant categorization. *Cognition, 105*, 362-379.

Verschoor, S.A. & Biro, S. (in press). The primacy of means selection information over

outcome selection information in infants' goal attribution. *Cognitive Science*.

Wagner, L. & Carey, S. (2005). 12-month-old infants represent probable endings of motion events. *Infancy*, *7*, 73–83.

Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1-34.

Woodward, A.L. & Sommerville, J. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, 11, 73-77.

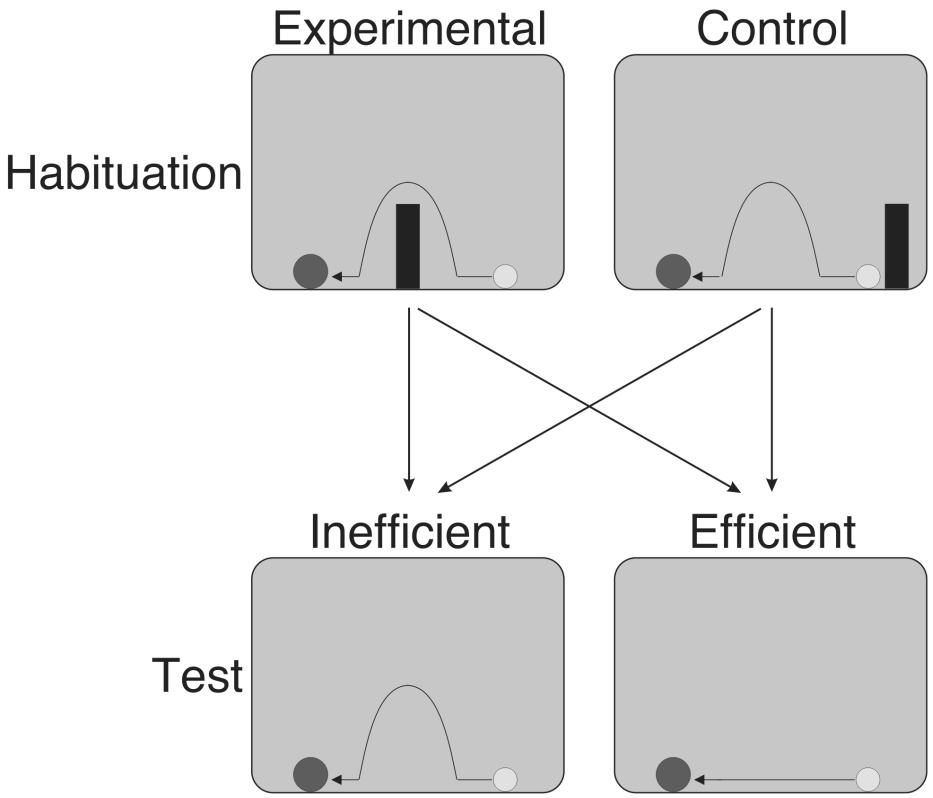

Figure 1. Illustration of the stimuli that were used to demonstrate the link between goal attribution and efficiency evaluation in infants (after Gergely et al., 1995). Infants were habituated to an event in which an agent (the small circle) approached another one (the larger circle) by a jumping action, which was either justified (Experimental condition, top left) or not (Control condition, top right) by the presence of an obstacle in between the agents. Their interpretation of these events were probed in the test trials (bottom row), in which the black box was removed from the screen. The looking times to these test events suggested that infants in the Experimental condition expected the agent to take the shortest pathway (Efficient approach, bottom right) instead of repeating the jumping action (Inefficient approach, bottom left) in the absence of the obstacle. In contrast, infants in the Control condition did not develop any specific expectation for the agent's action in the new situation.