

## More theory and evolution

Radu J. Bogdan

Department of Philosophy, Tulane University, New Orleans, LA 70118,  
bogdan@mailhost.tcs.tulane.edu

**Abstract:** Heyes's skepticism about theory of mind (ToM) in nonhuman primates exploits the idea of a strong and unified theory of mind in humans based on an unanalyzed category of mental state. It also exploits narrow debates about crucial observations and experiments while neglecting wider evolutionary trends. I argue against both exploitations.

Cecilia Heyes' thought-provoking article is a reminder that a convincing argument for a theory-of-mind competence in nonhuman primates remains hard to elaborate and defend. Lack of robust naturalist data and the rarity of decisive experiments are not the only reason, although it is the reason perhaps best exploited by Heyes' skepticism. Equally frustrating may be some lack of conceptual clarity and the reticence to theorize beyond the narrow boundaries of data and experiments. It is these latter shortcomings, also exploited by Heyes, that I want to address briefly in what follows.

**1. Why mental state?** I am puzzled by the notion of mental state placed at the heart of most analyses of the ToM competence. What could it mean? And why treat it as a premise rather than an outcome of inquiry? Philosophers have long disagreed about what mental states are, cognitive scientists don't really care, ordinary language is rather opaque but resolutely flexible and pragmatic about it,

and yet from early on many if not most ToM theorists have made the mastery of a naive category of mental state a test of having the ToM competence. So does Heyes, using the test as a prop for her skepticism. I think this is misleading will explain why in a moment. Right now I want to make a general methodological point.

It is a well known metascientific axiom that theoretical concepts in science, such as gravitation or gene, are usually defined implicitly in a given theory in terms of basic laws, causal or functional implications, and various assumptions. It is a theory of the ToM competence which ought to determine the nature of the categories that structure and run the competence. It cannot be a prior and pretheoretical decision or expectation of what the competence should be. This applies also to the human ToM: we shouldn't regard it as a genuine ToM because it fits pretheoretical expectations about mental state concepts but because its theoretical analysis suggests so. In short, the nature of the ToM categories and skills ought to follow from, not premise, the theoretical inquiry. I am afraid that a good part of the ToM field today suffers from a methodological inversion. This inversion is exploited by skeptics like Heyes. Two further and misleading assumptions also help such skepticism.

**2. Unhelpful premises.** Like many workers in this field, Heyes buys two premises about ToM which are misleadingly modeled on the adult human competence. One premise is that of a strong ToM, based on mental state categories, such as belief, desire, knowledge, which pick up types of mental conditions with causal powers and often clear-cut verbal and behavioral manifestations. The other premise is that of a unified ToM that operates by smoothly integrating such concepts. I think these premises generate the wrong set-up for the argument for or against ToM in nonhuman primates. Not having the space to elaborate this diagnosis (but see Bogdan 1997 and 1999) I restrict

myself to a few pointers.

The first premise first. A ToM competence need not be mentalist (i.e., based on mental state concepts that capture internal conditions with causal powers) in order to be made of specialized, functionally dedicated, and domain-specific skills. This is why an argument for (or against) ToM must first reveal (or rule out) such skills, not their mentalist character. If it turns out that nonmentalistic skills are actually involved in the apes' recognition of emotion, of information access (seeing) and in gaze following, then there is a viable alternative to the exclusive (and probably wrong) disjunction between a mentalist ToM and associative learning (see also Gordon 1998).

The apes' inability to recognize knowledge or attention in gaze or seeing need not therefore disprove the existence of domain-specific and specialized skills but rather clarify their functional profile and limitations. Which brings me to the second premise, about a unified ToM. There is no good reason to believe that ToM is one tightly knit and homogenous competence or that it emerged wholesale at some discrete point in primate phylogeny or human ontogeny. After 20 years of intense research, there are optimistic grounds to believe that the primate ToM is made of a rich battery of skills that had evolved in fits and starts and in a variety of social, interpersonal and cultural ambiances, building upon or converging with or inserting themselves in still other skills, and that only later in human childhood, language-based and thoroughly enculturated upgrades of such skills acquired a mentalist profile and job. But these later acquisitions provide no grounds to downgrade its precursors to mere learning or chancy guesses.

**3. Transcending the proximate.** Yet neither of these points against Heyes' premises would have much bite if we fail to take a wider theoretical and evolutionary view of the matter, something that Heyes' paper does not do. Let me begin with a theoretical point that addresses Heyes' skepticism about ToM in apes.

Most organisms evolve specialized organs and skills under assumptions about the ecology in which they operate. (To cite major instances, Vogel 1988 analyzes this phenomenon in biology in general and Marr 1982 specifically in vision.) The body design of fish and birds evolved under assumptions about the properties of the environments they travel through (water, air); on the artificial side, engineers would design ships and planes under rather similar assumptions. Assumptions are not an explicit part of the design of an organ or skill or artifact (so they cannot be determined by just looking at design) but are part of a wider ecology-organ/skill/artifact package, which is naturally selected or deliberately envisaged as a whole. The same is true of cognitive organs and skills. The visual system works under assumptions about light bouncing off surfaces, boundaries revealing shapes and volumes, and the like.

As part of cognition, the ToM skills are no exception. Assumptions about ecologies, natural and social, are an indelible complement to their design and operation. Assumptions also lead to a sort of division of labor between a skill's reach and the ecology's completion of it: a skill usually exploits ecological regularities or landmarks to do what it was designed to (Bogdan 1994). Thus, it could be that the sociopolitical ecology of apes calls for specialized ToM skills to track gaze and its direction (given their variable manifestations and utilities and the urgency of their manipulation) but not its targets (given that what one gazes at is a basic goal shared by everybody or can be determined contextually from various clues). It could also be that the job of a ToM skill in apes or human children consists precisely in patterning the right ecological, bodily and behavioral clues to guide a response behavior. That would still make the skill domain-specific and functionally dedicated but without any mentalist import. In general, if ecology-sensitive assumptions and divisions of labor are not factored into the analysis of a ToM skill, it may be hard to devise experiments that reveal its nature, particularly

when assumptions or divisions of labor or both are violated, as they may well be in laboratory contexts.

Yet an evolutionary perspective, absent from Heyes' paper, would do more than reveal assumptions and divisions of labor. It could also suggest, perhaps better than observation and experiment, whether there were selective pressures and opportunities for a specialized ToM competence. Such a suggestion would not settle the matter but would most likely constrain and inform the theoretical expectations. An analysis of the genetic proximity between apes and humans could further narrow down the estimates. As important and complementary would be an inquiry into the maturational schedule of the presumed ToM skills. What convinced many researchers that humans develop a specialized ToM competence was the tight ontogenetic scheduling of some of its key skills. Such a schedule suggests genetic expression and the latter in turn suggests some evolutionary pedigree. Wouldn't then be advisable to see whether there is such an ontogenetic scheduling in ape infancy and childhood as well, particularly for skills, such as gaze recognition and gaze following, which many researchers regard as domain-specialized and probably inherited by humans?

Finally, a look at the evolution of ToM skills in primates may also shed light on the interpersonal, cultural and linguistic pressures and opportunities that eventually led to mentalist ToM skills in humans. Whatever the form of these skills (modular, naively theoretical, inferential), the fact that they had to factor in mental conditions with causal powers must be explained (instead of premised) and so must the very possibility of this factoring. It is rather unlikely that such a possibility sprang into existence out of nowhere, without domain-specific and dedicated precursors that reach back into primate phylogeny. It should not be any surprise (should it?) that these precursor skills don't look human and don't operate as they do in humans.

## References

- Bogdan, R. J. (1994) Grounds for cognition. Erlbaum.
- Bogdan, R. J. (1997) Interpreting minds. MIT Press.
- Bogdan, R. J. (1999) Minding Minds. MIT Press.
- Gordon, R. M. (1998) The prior question. Behavioral and Brain Sciences 21: 120-121.
- Marr, D. (1982) Vision. W.H. Freeman.
- Vogel, S. (1988) Life's devices. Oxford University Press.

The inquiry may reveal that, and explain why, older humans do end up with categories of desire, belief and intention that capture some mental (in the head) conditions and their causal role in mentation and action. Yet the explication of these mentalist categories is still the job of a theoretical inquiry, as is its job to determine where they come from and why. Answers to these questions may establish antecedents and rudiments of such categories

subject-world pattern = <disposition (to behave) + memory data + relation type to the world + impact on interpreter>