

Why did the attribution of propositional attitudes evolve?

Hisashi Nakao

Andrews's *Do Apes Read Minds? Toward a New Folk Psychology* is very clear and interesting, and highly recommendable not only for philosophers of mind, psychology, and biology but also for developmental, social, evolutionary, and comparative psychologists. In this book, she argues that the standard folk psychology (SFP) is misguided because our actual folk psychology is more various and not fully described in terms of attribution of propositional attitudes. Moreover, while the SFP argues that folk psychological prediction and explanation require attribution of propositional attitudes, she argues that they do not necessarily.

Although I agree with the basic line of her argument, especially with her criticisms on the SFP, as even many great books often do so, this book also contains some ambiguities and problems. The aim of this short paper is to point out such problems especially in her evolutionary story of the capacity for attribution of propositional beliefs and desires.

Next, I also raise a question with Andrews and Huss's paper "Anthropomorphism, anthropectomy, and the null hypothesis in animal cognition research". In this paper, they assume that for the selective skeptics, the null hypothesis is that "animals do not have human-like cognitive systems, social relations, or normative properties". However, I am skeptical about this assumption. I point out the possibility that such skeptic null hypothesis is *not* the null hypothesis for the selective skeptics.

The main claims of *Do Apes Read Minds?*

This section summarizes Andrews's main claims that I agree with. After that, I point out some problems of her book in the next section.

As I have already briefly mentioned in the introduction, the main opponent of this book is the SFP, which has been endorsed by Alvin Goldman, Stephen Stich,

Shaun Nichols, Robert Gordon, and so on. First, the SFP claims that in order to predict and explain behaviors, we should be able to attribute propositional beliefs and desires. For example, when Poppy does something wrong to me and she looks at me with downcast eyes, I might predict her next behavior by attributing the following propositional attitudes to her: “She believes that I am angry with her, and thinks she ought to apologize to me, so she will say ‘sorry’ soon”. As she clearly shows in Ch. 2 and 3, many contemporary textbooks on philosophy of mind actually assume that folk psychology, i.e., predicting and explaining behaviors, is the attribution of propositional attitudes. Second, relatedly, according to the SFP, explanation and prediction are supported by the same mechanism, that is, the capacity for attributing propositional attitudes. So if we can predict behaviors by attributing propositional attitude, it follows that we should explain behaviors in the same way.

However, she argues that the SFP is misguided: when we predict and explain behaviors, we do not necessarily depend on the attribution of propositional beliefs, and prediction and explanation are supported by different kinds of mechanisms. In her book, this claim is supported both by philosophical and empirical arguments though I focus on the latter for the sake of space here. For example, developmental psychology has shown that 4-years-old children can attribute propositional beliefs and pass the traditional Sally-Anne test although younger infants cannot. However, recently some experiments suggest that much younger infants predict behaviors even though they cannot attribute propositional beliefs and desires (e.g. Onishi and Baillargeon 2005; Repacholi and Gopnik 1997). She distinguishes between *knowing how* (i.e., procedural knowledge) and *knowing that* (i.e., knowledge as propositional attitudes), and argues that it is plausible that knowing how is sufficient and knowing that is not necessary for prediction of behaviors. Moreover, much experimental evidence shows that we have many different biases for prediction without mental attribution such as biases for prediction from the self, stereotypes, traits and so on (Andrews 2012, Ch. 5). So prediction does not require attributing propositional attitudes and is supported by many different kinds of cognitive mechanisms or biases.

Infants under four years old, who cannot attribute propositional beliefs and desires, also actively engage in explanations. Andrews (2012, 146) writes “children are already

exploring the world as two-month-old infants, as demonstrated by their tendency to give tongues protrusions to both social and nonsocial stimuli, and at four to six months with their reaching behavior (Chen et al. 2006)". Chouinard (2007) also shows that infants under two years old actively and frequently ask questions related with explanation such as "What is that?" and "Why does she do this?". Moreover, explanations are very various: normative, linguistic, and causal structures need different types of explanations. So it is plausible that they need different kinds of cognitive mechanisms. Thus contrary to SFP, we do not necessarily need to attribute propositional beliefs and desires to explain behaviors, and which plausibly is supported by different kinds of cognitive mechanisms.

Her criticisms on the SFP and alternative theories or descriptions of folk psychology could be supported by other literature and have further interesting implications. For instance, given that as many developmental psychologists have shown, infants' learning is supported by many different biases (e.g., preference of prestige people Chudek et al. 2012; preference of older adults, Harris and Corriveau 2011; preference of pointers, Palmquist et al. 2012), her claims look very plausible and consistent with the previous researches. Also if we have many different cognitive biases for learning, predicting, and explaining, it should support the massive modularity hypothesis advocated by some evolutionary psychologists (e.g., Barrett and Kurzban 2006; Tooby and Cosmides 1992). So this book is a definitely valuable contribution to philosophy and psychology.

Why the capacity for the attribution of propositional beliefs and desires evolve?

As I have already mentioned, however, some problems also can be found in this book. This section focuses on and points out such problems of her evolutionary story on the capacity for the attribution of propositional beliefs and desires.

Previously, the social intelligence hypothesis has been the most plausible story for the evolution of the capacity for the attribution of propositional beliefs and desires, i.e., theory of mind. According to a version of the social intelligence hypothesis, (1) theory

of mind is useful for better predictions. So organisms with theory of mind are better predictors than organisms without theory of mind. (2) In fierce competitions (especially of primates) where individuals struggle to find mating partners, or get or retain the first rank in a social hierarchy, deceptions are very common. So they should make better predictions to detect such deceptions, and make better deceptions to deceive predictions by other individuals. (3) Thus theory of mind, which allows individuals to predict and deceive better, should have evolved in the fierce competitions.

Andrews (2012, 217-218) proposes an alternative hypothesis for the evolution of theory of mind. First, she points out that theory of mind is not necessary for predictions by saying “[o]ur ancestors who lacked a theory of mind did not need it to predict the behavior of friends and foes who themselves lacked a theory of mind, nor did they need one to deceive them”. I totally agree. As I have already argued, much developmental psychology literature supports this view.

It should be noted, however, that even though we do not need theory of mind to predict behaviors, it is still possible that theory of mind can be useful for *better* predictions. However, Andrews (2012, 219, *italics added*) rejects this claim, writing that:

[A] theory of mind offers an advantage *only when* predicting the behavior of someone else with a theory of mind; when predicting the behaviors of others who lack a theory of mind, attributing beliefs will normally offer *no additional predictive power*.

If she is right and theory of mind has “no additional power” for predictions behaviors by animals who lack theory of mind, it follows that theory of mind is useless for better predictions of behaviors by such animals. So she can reasonably reject the social intelligence hypothesis as the evolutionary story of theory of mind.

Unfortunately, she does not give any clear reasons for this claim. Her examples mentioned for this claim show not that theory of mind is useless for better predictions but just that predictions behaviors do not need theory of mind. For instance, she argues that some monkeys can understand the causal relationship between calls made when they find foods and other individuals coming together around the discoverer, and some individuals can stop to make the call and occupy the foods even if they do not have

theory of mind. I agree with this example if it shows that predicting behaviors does not require theory of mind, but do not agree if it intends to show that theory of mind is useless for better predictions. Thus it is unclear why the second stronger claim is supported and she fails to reject the first claim of the social intelligence hypothesis. The social intelligence hypothesis is still plausible as the evolutionary story of theory of mind.

For the sake of arguments, let me accept that theory of mind did not evolve for better predictions of behaviors by animals who lack theory of mind. Then why did it evolve? Andrews (2012, 220-221) argues as follows:

I suggest that our ancestors would have benefited from having a theory of mind to predict behavior in *anomalous* situations...To explain this [anomalous] behavior, an observer has no recourse but to consider the actor's reasons.

Although I agree with her claim and some reasons should be needed to explain anomalous behaviors, I remain unconvinced with why we need theory of mind to consider actor's reasons. For instance, infants under four years old, who lack theory of mind, actually face anomalous situations every day, but it is plausible that they can explain the anomalous situations even without theory of mind because if they cannot, it follows that every time they face anomalous situations, they have no good explanations and feel affective tensions. Some empirical evidence suggests, however, that younger infants often seek explanations for anomalous situations, as Andrews (2012; e.g., Chouinard 2007) also admits. In these cases, it seems that younger infants can get good explanations even though they do not develop theory of mind. Actually, it is possible that knowing how is sufficient for younger infants or some non-human primates to have good reasons and explanations for anomalous situations, say, through trial-and-error learning, imitation, and a kind of simulation without theory of mind. Andrews does not consider such possibility and her argument seems hasty. At least, she needs to explain whether infants under four years old have good explanations without theory of mind or not, or how they do so if they can.

Worse, at least for her, if I am right and we do not necessarily depend on theory of

mind for explaining anomalous behaviors or situations, we should reject her evolutionary hypothesis of theory of mind according to her criteria (Andres 2012, 219, *italics added*): While it is true that evolution is not a tidy process, we should avoid postulating the development of a unique cognitive process to make better prediction of behavior when the current mechanisms work just fine.

Let me move to the final remark. Andrews suggests that two different anomalous situations are important for the evolution of theory of mind. The first one is norm violations and the second is technological innovations. Andrews (2012, 224) develops an argument for the first as follows:

...rather than developing a theory of mind to predict behavior, humans may have developed a theory of mind to explain norm violations...to perceive a behavior as anomalous, one needs to have an understanding of normal behavior, and to understand behavior as normal is to have at least an implicit understanding of the relevant social norms.

So if we grasp some social norms, we can understand whether some behaviors are norm violations or not, and then such norm violations should require explanations.

I am skeptical about this argument in two regards. First, we do not often need explanations for norm violations but just need to punish, protest, or turn off the violators when we witness norm violations. For instance, Hamlin et al. (2006) shows that even 6-months-old infants autonomously dislike objects violating norms. Thus it is unclear whether norm violations often require explanations or not. Second, even if you do not grasp social norms, you can perceive some of my behaviors as anomalous just when you know my ordinary behaviors. For instance, every time I go to your office with a cup of coffee, and you perceive this as my usual or common behavior, you can perceive my going to your office without a cup of coffee as anomalous. So because of these two points, I do not understand why we need to refer to social norms to explain the evolution of theory of mind.

Let me summarize my questions and criticisms on Andrews' book. First, because Andrews fails to consider the possibility that infants under four years old (who do not develop theory of mind) can have good explanations, and that knowing how is sufficient

for explaining anomalous behaviors, I remain unconvinced with why theory of mind is needed to explain anomalous behaviors or situations including norm violations. Second, her alternative story especially on norm violations is also unconvincing because it is not clear whether we need explanations for them. Worse, since she does not give any clear reasons why theory of mind is useless for better predictions of behaviors by animals who lack theory of mind, she fails to reject the social intelligence hypothesis.

Andrews and Huss's "Assumptions in animal cognition research"

Now let me turn to the Andres and Huss's paper. In this paper, first, they distinguish between the categorical skeptics and selective skeptics. Although the categorical skeptics argue that the whole or entire research program of comparative animal research fails, the arguments by the selective skeptics are more modest. According to the authors, they do not deny the possibility of the research program of comparative animal research, while they argue that we should employ the following hypothesis (the skeptic null hypothesis) as the null hypothesis, i.e., "For selective skeptics, the null hypothesis is that animals do not have human-like cognitive systems, social relations, or normative properties". After such characterization, the authors argue that the selective skeptics confront a kind of dilemma by using the distinction between Type-I and -II errors.

I think their arguments are interesting but the problem is their assumption: My question is whether the skeptic null hypothesis is really the null hypothesis for the selective skeptics or not. It seems that anthropomorphism and anthropechtomy are biases not for the research methods or the null hypothesis rather for interpretations of the results, and that the selective skeptics do not argue that we should consider the optimistic hypothesis (i.e., non-human animals do have human-like psychological properties) only after the skeptic null hypothesis is first rejected. So more charitable interpretations of what the selective skeptics argue is that the optimistic researchers should be cautious of their interpretations of the experimental results: they also think of and test the possibility that the animals do not have human-like psychological properties. Actually, the selective skeptics can use the optimistic hypothesis to construct the research design as follows:

Expectation: If the animals do have a human-like psychological property P (optimistic hypothesis), we should expect that the animals would behave like B in an experimental situation E.

Result: The animals fail to behave like B in E.

Conclusion: So we can reasonably reject the optimistic hypothesis that the animals have P.

So if I am right, I can conclude that Andrews and Huss's assumption is misguided and they fail to depend on the rest of the arguments to criticize the selective skeptics.

Conclusion

Although I agree with Andrews in that the SFP is misguided, I am skeptical about her evolutionary story of theory of mind. She fails to show any convincing reasons to reject the social intelligence hypothesis. Also Andrews and Huss's paper develops an interesting argument though it seems that their assumption is misguided and that they need to restructure the argument.

References

- Barrett, H. C. and Kurzban, R. 2006. Modularity in cognition: Framing the debate. *Psychological Review* 113: 628-647.
- Chen, X., Reid, V. M., and Stiano, T. 2006. Oral exploration and reaching toward social and non-social objects in two-, four-, and six-month-old infants. *European Journal of Developmental Psychology* 3(1): 1-12.
- Chouinard, M. M. 2007. Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development* 72(1): 1-112.
- Chudek, M., Heller, S., Biro, S., and Henrich, J. 2012. Prestige-biased cultural learning: Bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, 33(1): 46-56.
- Hamlin, J. K., Wynn, K., and Bloom, P. 2007. Social evaluation by preverbal infants. *Nature* 450: 557-559.
- Harris, P. L. and Corriveau, K. H. 2011. Young children's selective trust in informants. *Philosophical Transactions of the Royal Society B*, 366: 1179-1187.

- Onishi, K. H. and Baillargeon, R. 2005. Do 15-month-old infants understand false beliefs? *Science* 308: 255–258.
- Palmquist, C. M., Burns, H. E., & Jaswal, V. K. 2012. Pointing disrupts preschoolers' ability to discriminate between knowledgeable and ignorant informants. *Cognitive Development*, 27: 54-63.
- Repacholi, B. M. and Gopnik, A. 1997. Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology* 33(1): 12–21.
- Tooby, J. and Cosmides, L. 1992. The psychological foundations of culture. In J. Barkow, L. Cosmides, and J. Tooby (eds.) *The adapted mind: Evolutionary psychology and the generation of culture*, New York: Oxford University Press, pp. 19-137.