Do Apes Attribute Beliefs to Predict Behavior? A Mengzian Social Intelligence Hypothesis

Kristin Andrews

Abstract: I defend a Mengzian version of the Social Intelligence Hypothesis, according to which humans think about one another’s beliefs and desires—and reasons for action—in order to solve our social living problems through cooperation, rather than through competition and deception, as the more familiar Machiavellian version has it. Given this framework, and a corresponding view about the function of belief attribution, I argue that while apes need not attribute propositional attitudes to pass the “false belief task,” we should not conclude that apes may be behaviorists. Rather, the Mengzian Social Intelligence Hypothesis perspective offers another interpretation of ape behavior, intermediate between behaviorist and propositional attitude schemas. I argue that the false belief task can be solved by individuals who have an agency schema which takes others to be minded beings who have goals, emotions, and perceptions, but who fail to consider propositional attitudes or reasons for behavior. I then argue that a true test of belief attribution in great apes would be one that shows they seek explanations in terms of reasons for behavior. However, no such test yet exists.
I. Debunking Beliefs about False Belief

False beliefs drive our comedic and tragic narratives. Romeo thinks Juliet is actually dead, so kills himself. Countess Olivia falls in love with Viola, thinking she is a man.

We love these stories, as demonstrated by Shakespeare’s continued popularity hundreds of years after the stories’ first telling. Do we love them merely because they are entertaining, or is there a deeper reason that ties into our evolutionary history? Some say the latter; according to leading theories in evolutionary psychology, ascribing false beliefs is part of a lynchpin in the evolution of human intelligence, and an ability unique to humans.

Telling stories isn’t the key, according to these views. Rather, the stories are a pleasant side effect of the primary function of belief attribution, namely predicting and manipulating behavior. This view is expressed by some supporters of the Social Intelligence Hypothesis, who suggest that the driving force in the evolution of cognition is social complexity, because the biggest problems in the lives of social creatures are social problems, not ecological ones. The Machiavellian version of this hypothesis stems from a picture of our evolutionary social context as one resembling Hobbes’s account of humans before the social contract, or Xunzi’s picture of human nature without artifice. Humans are portrayed as selfish and manipulative, and in order to thrive in such an environment, we had to learn to predict, deceive, and manipulate behavior. We did this, so the story goes, by considering the mental causes of behavior. In particular, the Machiavellian Social Intelligence Hypothesis suggests that humans learned to manipulate behavior by coming to think that others have beliefs and desires that together cause behavior, that these beliefs can be true or false, and that by instilling false beliefs in others we can cause them to act as we want them to act. Acts of deception—making people think falsely—are seen as the key to human cognitive development.

It is uncontroversial to claim that human social complexity differs significantly from the complexities we see in other animals, and the Machiavellian Social Intelligence Hypothesis offers an explanation for the difference —humans alone came up with the social technology of belief attribution. This hypothesis has been consistent with forty years of research on nonhuman animals. While psychologists accept that young children attribute beliefs, until recently, attempts to find evidence of belief attribution in the nonhuman great apes—chimpanzees, orangutans, bonobos and gorillas—were unsuccessful. All the while, we continued to find what appeared to be converging evidence that the ability to attribute beliefs is at the core of human cognition, with claims that even human infants implicitly attribute beliefs to predict behavior. Until recently, the tentative conclusion has been that humans are the only species who think about beliefs, who tell stories about beliefs, and who intervene into others’ beliefs to manipulate their behavior.

But this conclusion seems to be in conflict with more recent research. Primatologists recently discovered that great apes, too, can track false beliefs in the same way that infants do, and they have concluded that apes also understand false beliefs, at least implicitly. So, do great apes and infants both attribute beliefs, and is belief attribution an ability shared among ape species? Or, was there a misstep somewhere in this line of reasoning? The finding that great apes track what we call false belief behavior makes this a good time to re-evaluate the Machiavellian Social Intelligence Hypothesis, its picture of human nature as selfish and competitive, and its corresponding commitment to prediction and manipulation as the primary function of belief attribution. The re-evaluation suggests another picture of human cognitive development, more Mengzian than Machiavellian—a picture that empha-
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

sizes the power of storytelling, cooperation, and community concern, and that downplays the role of belief attribution for predicting behavior.

To engage in this re-evaluation, we can begin by recalling the history of belief-desire psychology and the Social Intelligence Hypothesis, and see how the question of whether chimpanzees understand others’ minds became a question about whether chimpanzees have what researchers call a “theory of mind”—with little attention to whether or not adult humans have the kind of theory assumed by the question. We will see how the false belief task became a litmus test for theory of mind in apes and children. With this background in hand, I will argue that passing the false belief task, something infants and apes are now known to do, does not require a theory of mind. However, I will also argue that it requires more than behavior-reading, which is the familiar alternative account of ape social cognition. Instead, I present a middle ground hypothesis, and defend the claim that chimpanzees are mentalists who take others to be agents, but that there is currently no evidence that they are belief-desire psychologists who consider others’ reasons for action. This alternative hypothesis challenges the Machiavellian Social Intelligence Hypothesis as an account of human social cognition. Rather than using belief-desire psychology in order to predict others’ behavior, I argue that humans use belief-desire psychology to explain others’ behavior. This claim, that the function of belief-desire psychology is for explaining behavior rather than predicting behavior, suggests new ways of studying chimpanzee social cognition, and it also supports a Mengzian Social Intelligence Hypothesis.

II. BENEFICIARIES OF BELIEF-DESIRE PSYCHOLOGY

The idea that deliberate human action is caused by beliefs and desires is almost as old as Western philosophy. In De Anima Aristotle tells us, “These two at all events appear to be sources of movement: appetite and mind.” Appetite, what today we would call desire, is the fundamental mover of all animals, human and nonhuman, but in order to direct us to action that will in fact fulfill our desire, we need mind to guide us—or in contemporary terms, we need belief. Aristotle focuses on how we come to act when we have competing desires, and he appeals to the mind as the means for deciding between them. Desire alone won’t lead me to act, because while I may want to run a marathon, I may also want to be slothful, and I need mind to guide me.

Some 2000 years later, David Hume ratchets up the role for belief and desire in his causal account of action. On the view he proposes, all our intentional actions are caused jointly by reasons (belief) and passions (desire), because one without the other can’t lead to any action at all. I may believe that the bakery is open now, but if I don’t want to do anything in the bakery, my belief isn’t going to cause me to walk there. Likewise, if I desire to eat bread, but don’t have any beliefs about how to acquire bread, I won’t be moved to act. In either case there is a missing causal link needed for action.

The Machiavellian Social Intelligence Hypothesis is a natural development of Humean belief-desire psychology. The reasoning goes like this: Human action is caused by beliefs and desire, so if I want to manipulate others’ behavior, then I can manipulate their belief, or I can manipulate their desire. Realizing the relationship between beliefs and desires and developing the skill of creating false beliefs or discrepant desires, a Machiavellian can reap benefits that in turn will motivate competitors to likewise develop this skill, and the looping effect leads to the spread of belief-desire psychology through human society.
The theory of mind research program is one branch of research that stems from the Machiavellian Social Intelligence Hypothesis, and is aimed at examining the development and evolution of the ability to predict behavior based on the ascription of beliefs and desires. Given the assumption that typical mature humans regularly engage in this behavior, then a natural question is to ask when humans gain this ability, and whether humans are alone in having this ability.

David Premack and Guy Woodruff initiated the research into great ape belief attribution, and introduced the term “theory of mind.” They originally defined the term as follows:

An individual has a theory of mind if he imputes mental states to himself and others. A system of inferences of this kind is properly viewed as a theory because such states are not directly observable, and the system can be used to make predictions about the behavior of others.

Premack and Woodruff’s definition is wedded to the Aristotelian and Humean views insofar as it assumes that in order to predict intentional behavior, one has to understand the cause of that behavior. By adopting belief-desire psychology, Premack and Woodruff accepted that humans do indeed need to think about the cause of behavior in order to predict behavior. This is how Humean belief-desire psychology led to the development of the next stage of our story—the so-called false belief task, variations on which have been essential to the investigation of the development and evolution of belief reasoning.

III. The False Belief Task

In 1978, Premack and Woodruff asked whether Sarah, a fourteen-year-old chimpanzee, had a theory of mind. They showed Sarah films of a human trying to achieve a goal, such as warming up a cold room with a heater or opening a locked door. To solve the problem, Sarah chose a photograph representing what the human needed: a match to light the heater, a key to open the door. Premack and Woodruff concluded that Sarah understood the actor’s beliefs, and that she was able to choose the right object by thinking about what the actor intended and believed. However, this interpretation was largely rejected by commentators and later by the authors themselves, after considering alternative explanations grounded in associative reasoning. Sarah, who lived in a world of locked doors and heaters, could have simply connected matches with heaters, and keys with locks. In commentary, Daniel Dennett, Jonathan Bennett, and Gilbert Harman each suggested that a real test would be to ask whether Sarah understands that others can have and act on false beliefs. The idea is that when someone has a false belief, simple association can’t be used to solve a prediction problem, because there is no association between the way things are in actual fact. Instead, the attributor needs to think about how the world appears to the target, and use that appearance rather than the reality to draw a conclusion about the target’s likely behavior.

While we didn’t see another published chimpanzee theory of mind study for twenty years, this idea led to an explosion of research on human children, examining both typical and atypical human development, based on what became known as the moved object false belief task. The original version goes like this: children watch a puppet show in which Maxi hides a piece of chocolate in one cupboard and then leaves the room. While Maxi is out, his mother finds the chocolate and moves it to another cupboard. Then Maxi comes back to get his chocolate. Maxi says, “I’m going to look for my chocolate now.” At this
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

point the show stops, and children are asked to point to where Maxi is going to look for his chocolate—in the cupboard where he left it, or the cupboard where it actually is. None of the 3 or 4 year-old children in the original study were able to correctly predict that Maxi would seek out his chocolate where he left it. Rather, they thought Maxi would go where the chocolate really was. This result was interpreted as demonstrating that young children are unable to attribute false belief, and that the older children who pass the test can. A meta-study finds that Western children tend to pass this task beginning at between 4-5 years of age. However, it is a matter of ongoing discussion as to why young children don’t pass it.

The quest then became to develop a nonverbal version of the test that could be used with human infants and with nonhuman great apes. The first comparative false belief moved object task examined chimpanzees, orangutans, and human children. A Communicator, but not the research participants, could see in which box a Hider placed a food item, and would mark the box that contained the treat by placing a token on top. The participants all learned to trust the Communicator’s token, choosing the marked box that had the food in it. In the false belief test, the Hider hid the food, the Communicator left the room, and then the Hider switched the food to a different box, in view of all the subjects (yet of course they still didn’t know in which box the food had been hidden, or where it was now). When the Communicator came back into the room, he marked the original location of the food. Those who can reason about belief should avoid the token in this case, according to the theoretical framework, since the Communicator doesn’t know where the food actually is. Since chimpanzees and orangutans persisted in picking the marked box, but five-year-old children were able to make the switch in the new context, the researchers concluded again that the ability to attribute beliefs might be unique to humans.

Despite variations on this task over the years, great apes continued to fail them, and human children continued to pass them—and pass them at younger and younger ages. Over time, many researchers cautiously concluded that chimpanzees don’t understand false belief. In their review of 30 years of research on chimpanzee mindreading, Call and Tomasello wrote, “chimpanzees probably do not understand others in terms of a fully human-like belief–desire psychology in which they appreciate that others have mental representations of the world that drive their actions even when those do not correspond to reality. And so in a more narrow definition of theory of mind as an understanding of false beliefs, the answer to Premack and Woodruff’s question might be no, they do not.”

On the infant research front, things developed very differently. In the first published study of infant mindreading, researchers measured how long infants looked at a person acting according to her false belief about the location of an object compared with how long they looked at a person acting inconsistently with her false belief. Infant of fifteen months acted as though they expected the person to act according to a false belief—but they also acted as though the agent was ignorant. In order to gain evidence that infants are indeed able to attribute false belief, a number of other experiments were run by various researchers. Prominent among those was a study by Victoria Southgate and colleagues using eyetracking technology to examine twenty-five-month-old infants’ anticipatory looking behavior. Infants were familiarized to an agent who wanted to obtain a ball hidden by a puppet. The false belief conditions involved the puppet moving the ball between two boxes, leaving it in the right-hand box, and exiting the scene. Then the agent turned away, the puppet returned, opened the right-hand box and took the ball with him out of the scene. When the agent turned back, and the infant was cued that the agent was about to seek the
ball, infants looked at the right-handed box, even though the ball was no longer inside, and even though the ball had spent the same amount of time in both boxes. The authors conclude, “The data presented in this article strongly suggest that 25-month-old infants correctly attribute a false belief to another person and anticipate that person’s behavior in accord with this false belief.”

Southgate’s study revolutionized the investigation of great ape theory of mind. Krupeyne and colleagues modelled their research on Southgate’s and used materials that were more likely to catch the apes’ attention to elicit the same behavior as that seen in the infants. For example, apes watched a movie showing a human who was attacked by someone in a gorilla suit (King Kong). The ape participants first saw King Kong attack the human, and then run into one of two haystacks to hide. The human grabbed a stick and hit the haystack where King Kong was hiding. In the false belief conditions, the human had to leave the scene to get the stick, at which point King Kong changed position and left the scene. When the human came back, the chimpanzee subjects looked at the haystack in which King Kong hid before the human left, anticipating that the human would beat that haystack with the stick (See Figure 1).

Following on the heels of this study was another paper claiming that great apes attribute false belief behavior in an active helping task. Again, the task was modeled on a test first given to human children, in this case eighteen-month-old infants. In the original task, children too young to pass the verbal task were able to help a human open a box, and they were able to determine which box he wanted to open even when he had a false belief. The experimental set-up involved a blue box and a yellow box and an object that the researcher was playing with. In the false belief condition, the researcher placed the object in the blue box and locked the box, giving the key to the child, and then leaving the room. While the researcher was gone, another adult unlocked the box and moved the object to the yellow box, locking both boxes after him. When the researcher came back into the room, he tried

Figure 1. False belief 2 condition from Krupeyne et al.’s experiment 1. The chimpanzee subject watches as the human sees King Kong hide in the right haystack, and then goes inside to get a stick, closing the door behind him. While the human isn’t watching, King Kong moves from the right haystack to the left, and then leaves the scene. Then the door opens, the human comes out with the stick raised over his head, and the chimpanzee subjects look more at the right most haystack, where the human last saw King Kong. These looks are interpreted as a prediction that the human will hit the rightmost haystack, and an attribution of a false belief that King Kong is hiding in the right haystack. Figure © 2016 Amy Noseworthy. Reprinted by permission.
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

to open the blue box. The children with the key tended to open the yellow box for him, revealing the desired object. Importantly, in the true belief condition, which involved the same movements except that the other adult didn’t move the object, when the researcher came back into the room and tried to open the (empty) yellow box, the children helped him open that box. The children seemed to be able to determine the researcher’s goal in the two conditions, and this was interpreted as evidence that they know the content of the researcher’s belief about the location of the object in each condition. Using that belief attribution, and seeing the researcher’s action, children are able to work out the missing part of the causal frame, namely what the researcher desires.

When the same experimental set-up was run with chimpanzees, orangutans, and bonobos, researchers found the same kind of difference between the true and false belief conditions (but interestingly, the apes were worse than the children in the true belief condition). In these so-called implicit false belief tasks, great apes, like humans younger than four, are able to track a person’s behavior and goals when they have a false belief. The ape researchers, like the infant researchers, take success at this task to be evidence that the subjects understand false belief:

[O]ur results, in concert with existing data, suggest that apes solved the task by ascribing a false belief to the actor, challenging the view that the ability to attribute reality-incongruent mental states is specific to humans.20

Great apes thus may possess at least some basic understanding that an agent’s actions are based on her beliefs about reality. Hence, such understanding might not be the exclusive province of the human species.21

Almost forty years after Premack and Woodruff first asked their question, researchers were finally able to elicit false belief tracking behavior in the nonhuman great apes. But the question that remains is how do they—and we—do it? What is needed to track false belief behavior? The Machiavellian Social Intelligence Hypothesis’s ready-to-hand answer, drawn from Humean belief-desire psychology, is that we attribute beliefs to track the sort of reality-discrepant behavior we see in the moved object false belief experiments. And this is the explanation that the researchers themselves endorse.

Before we can evaluate, much less endorse, the claim that humans and other apes need to think about false belief in order to solve these kinds of social problems, we need to know what belief attribution gets us. So long as belief is an unanalyzed notion, interpreting these behaviors in terms of an ability to attribute beliefs will be uninformative at best. It’s time to start examining the assumptions about the role of beliefs that underlie the Machiavellian Social Intelligence Hypothesis.

IV. Belief, Behavior, and Perception

With all this discussion of false belief, very little has been said about the nature of a belief itself, other than that it can be true or false, and that the ability to represent both is an essential part of having a belief. But a belief is more than something that can be false. When we say that Romeo believes that Juliet is dead, we attribute to Romeo a commitment that takes as true the meaning of the sentence “Juliet is dead.” While Juliet looks dead, she also looks like she has taken a sleeping potion, and we don’t say of Romeo that he believes “Juliet is dead or she is asleep due to a potion,” because that belief, along with his desire not to live
without Juliet, would not cause him to kill himself. The disjunctive belief would cause him, we should hope, to examine the other possibility.

In contemporary philosophy, a belief is understood as a type of propositional attitude.  
A proposition, or the meaning of a sentence, can be the object of a variety of attitudes. We can hope that Juliet is dead, worry that Juliet is dead, desire that Juliet is dead, or believe that Juliet is dead. Propositional attitudes have the peculiar logical property of referential opacity. A proposition on its own will not change its truth value when we substitute different names for the same referent. If Cesario weighs 150 lbs, then Viola does, too. That inference is justified by the fact that Cesario and Viola are one and the same entity. However, if we embed the proposition in an attitude, the same substitution can change the truth value of the sentence. Countess Olivia believes that Cesario is worthy of her love, but she does not believe that Viola is worthy of her love, even though Cesario is Viola.

As a referentially opaque propositional attitude that can be true or false, and along with desire causes behavior, a belief is something that requires quite a bit of cognitive sophistication. Given the high cognitive demands, some critics of the false belief task claim that the apes can solve these sorts of problems by using a theory of behavior rather than a theory of mind. The idea behind this challenge, dubbed the logical problem, is that chimpanzees who have observed chimpanzee behavioral patterns are able to read behavior; they can anticipate what another chimpanzee will do without knowing why the chimpanzee acted, and even without knowing that apes are agents.

Cecelia Heyes thinks that the logical problem arises for the King Kong study as well. Heyes proposes what she calls a sub-mentalizing hypothesis: the apes may be focusing on observable features of the situation, such as the appearance and disappearance not of the human, but of the colorblock of the persons’ shirt, and the movement of the colorblock of King Kong when the colorblock of the person’s shirt is in the scene compared to when it is out of the scene. The return of the colorblock of the person’s shirt could cause the subjects to recall the location of King Kong when the colorblock of the shirt was last there, hence giving the impression that the apes appear to understand false belief. According to the sub-mentalizing hypothesis, the ape subjects do not see the movement of the agents as a story with narrative structure, and they do not see the human and the King Kong as agents who act for reasons.

In all versions of behavior-reading hypotheses, the proposal is that apes can know what others will do without knowing why they do what they do. The why is Skinner’s discarded intervening variable between observable actions. According to Humean belief-desire psychology, those internal mental states supply an answer to that why-question because they complete the causal chain. The critics rightly point out that for typical behavior, knowing hidden inner causes isn’t needed to predict behavior. Dominants seek out food that they are oriented towards. Bared teeth facial expressions are followed by violence. Where the behavior-reader can make an inference directly from these observable features to future behavior, to attribute belief one must first make an inference to a hidden mental state, take it to be a cause, and then draw a conclusion from that causal relationship in order to predict behavior.

While the critics are right to point out that quotidian behaviors are transparent in this way, such critics need not also reject the idea that great apes know something of others as agents. There is a middle ground between belief attribution and sub-mentalizing/behavior reading. We can also know, without entertaining a propositional attitude, what someone feels, wants, and, importantly, sees. We know that Romeo sees Juliet, and that he also sees
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

her as dead. We know that Countess Olivia sees Viola, but that she sees her as Cesario. In developmental psychology, knowing what someone else sees is called level-1 perspective taking, and knowing that someone sees something as having some property is level-2 perspective taking. By three years of age, children are already engaged in level-2 perspective taking, and can tell that an object behind a colored filter looks different to people taking different perspectives.

The researchers who introduced the King Kong task realize this possibility when they admit their findings are consistent with a competing hypothesis:

We acknowledge that all change-of-location false-belief tasks are, in principle, open to an abstract behavior rule-based explanation —namely, that apes could solve the task by relying on a rule that agents search for things where they last saw them.

And the authors of the active helping task suggest that they are providing converging evidence to bolster the finding of the King Kong study:

In the Krupenye et al. study, apes could have passed the test simply by predicting that the actor would go to the last place he saw the object; thus, converging evidence is needed from tests that use different methods.

Studies only provide converging evidence if they are different enough from one another to count as independent. Since both the King Kong and the active helping studies can be explained in terms of the intermediate, mentalizing hypothesis that apes know what others can and cannot see, and because there is independent evidence that apes are sensitive to what others see, these studies don’t provide much in the way of evidence that other animals think about belief.

In the King Kong study, the ape subjects might predict that the human will go to the left haystack because that’s where the actor last saw King Kong, and actors seek out things where they last saw them. Indeed, this same hypothesis can account for passing the verbal Maxi task. This is not a behavior reading type of rule, since it appeals to a mental concept, namely seeing. But this alternative explanation doesn’t invoke level-2 perspective taking, requiring only the simpler level-1 perspective taking of knowing what others can and cannot see. Similarly, in the active helping task, subjects understand the target’s goal by understanding that actors seek out things where they last saw them, so in the false belief condition, the ape thinks the researcher’s goal is to acquire the object because the researcher is looking where he last saw it.

The seeing hypothesis is a mentalistic hypothesis, but it is a nonpropositional mentalistic hypothesis that deals in states of affairs of the world. The seeing hypotheses converges well with other things we know about how apes treat other apes. We know that chimpanzees perceive a difference between intentional behavior and accidental bodily movement. We know that they are more impatient with humans who are unwilling to provide food than they are with humans who are unable to give them food. We also know that chimpanzees can identify goals and spontaneously help humans achieve their goals, picking up dropped objects and passing them back to a human companion.

We also know quite a bit about chimpanzee sensitivity to seeing. Early field researchers reported that adult chimpanzees monitor gaze, though infant chimpanzees don’t attend to gaze when making requests from their mothers until they are about ten and a half months. In addition, low-ranking chimpanzees act differently when they are out of the dominant’s sight. When a low-ranking chimpanzee is invisible, he will take advantage of
his situation by mating with a preferred partner, or by eating food that would not be available to him if the dominant were present. Chimpanzee food competition studies reinforce these observations from the wild.

From field studies we have evidence that great apes provide social support for one another; they arguably teach their young; they communicate using gestures; they have care practices that protect vulnerable members of the community; and they protect their territory by attacking outgroup members. None of this is best explained by claiming that apes don’t understand others to be agents with goals and emotional states. However, none of this is best explained by saying that apes take others to be Humean belief-desire creatures, either.

We have seen three hypotheses about great ape social understanding. One is that apes are sub-mentalizing behavior-readers who do not see others as intentional agents, but are simply able to infer future behavior from past behavior, just as they are able to anticipate when trees will fruit given past experience with fruiting trees. The second hypothesis is my intermediate non-propositional mentalizing hypothesis, that great apes are able to predict behavior by understanding others as agents, who, among other things, can see. The third hypothesis is the one proclaimed by ape researchers involved in the current studies, that great apes have a concept of belief, and are able to ascribe propositional attitudes such as false beliefs to others in order to predict behavior.

Of all these hypothesis, the non-propositional mentalizing hypothesis is the best supported. The sub-mentalizing hypothesis is at odds with the naturalistic and experimental data we have about ape social cognition, and supporters of this hypothesis would have to offer alternative explanations for chimpanzee affiliative relationships, emotion recognition, and the ability to discriminate intentional from accidental behavior. The belief-attribute hypothesis entails many more things about chimpanzees than the non-propositional mentalizing hypothesis does. For example, it entails that chimpanzees have level-two perspective taking, which is something we don’t at the time have converging evidence for. It entails that chimpanzees have a concept of true and false, that they are able to think about referential opacity, and that they understand that how the world appears to a target is part of the cause of the target’s behavior. In contrast, the non-propositional mentalizing hypothesis is compatible with what we know about chimpanzee ability to track what others see. Furthermore, as we will soon see, it does not rely on suspect Humean belief-desire psychology.

If apes and human children can predict behavior without thinking about beliefs, the Machiavellian Social Intelligence Hypothesis—at least the part of it that takes belief attribution to have the function of predicting behavior via the kind of causal reasoning assumed by Humean belief-desire psychology—needs to be rejected. We can retain the idea that social complexity drives cognitive complexity, but reject the key role given to belief attribution in order to manipulate others in the story of our evolutionary history.

V. Why Humans Attribute Beliefs

Humans talk about people’s beliefs, so we know that humans can attribute beliefs. But, as we have seen, humans are skilled at predicting behavior without thinking about beliefs. Can humans deceive without thinking about beliefs, too? Following Richard Byrne and Andrew Whiten, we can identify two types of deception: tactical deception, or “acts from the normal repertoire of the agent, deployed such that another individual is likely to mis-
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

interpret what the acts signify, to the advantage of the agent. and true deception, which also requires that the deceiver intends to modify the target’s mental state to create a false belief.

Tactical deception goes a long way towards providing the benefits associated with manipulating behaviors. I don’t need to be able think about changing other people’s minds if I know how to change their behavior, and a clever behaviorist can recognize patterns and intervene in those patterns in order to effect outcomes. Take the example of finding food but not giving the species-typical food alarm call. An animal who doesn’t understand belief can figure out that suppressing the call results in keeping the food for oneself. How could they learn this? First, the animal could note that whenever they find food, they make a particular call, and when they make that call other members of the community come running. Through causal reasoning, the animal could learn not to call when they want to eat all the food. In a competitive social environment, this causal discovery offers an advantage, but it is not an advantage that requires the ability to think about beliefs.

The emphasis on false belief in developmental and comparative psychology stems from a commitment to the importance of false belief in contemporary human adult life. On my theory of Pluralistic Folk Psychology, this commitment is misplaced, as belief attribution is only one among many cognitive strategies humans use to predict behavior, and it is a strategy that does not offer much in the way of accuracy. Other methods can be used to predict what people are going to do, such as expecting that others will do what we tend to do and appealing to stereotypes, norms, and social roles to predict that people will do what they should do. When we know the people we are predicting, we can also use strategies such as inductive generalizations over past behavior or trait attributions. We can attribute propositional attitudes, of course, but then again we can also think about others’ non-propositional mental states such as emotions, moods, and intentions.

I’ve argued that attributing beliefs to others is not one of the more important tools in our predictive toolset, because this method isn’t likely to be accurate, and we’re better at anticipating behavior than we should be if we used this method. For one, determining the correct and relevant belief is difficult because the appropriate belief/desire sets are underdetermined by theory and observable behavior. I might turn off the light because the light hurts my eyes, or because it is time to turn off the lights, or because I don’t want anyone to know that I’m home. How is an observer to decide between these, and the many other beliefs that combined with a desire could lead me to turn off the lights? The underdetermination problem leads to computational intractability when trying to decide which attitudes a person holds.

Another problem arises because any particular belief/desire set is going to serve as a hypothesis about hidden causes of behavior. Once we adopt a hypothesis, confirmation bias kicks in, leading us to seek evidence in favor of our hypothesis and ignore evidence that undermines it. In one study demonstrating this bias, college students were asked to predict whether they would act in a friendly or unfriendly way toward another student. Participants in the experimental group were also asked to consider their reasons for action in order to make the prediction, and those in the control group weren’t given this instruction. Researchers found that when people thought about their reasons for actions, they ended up with a higher degree of confidence in their prediction, but their predictions were more likely to be wrong. If we can generalize from this finding, thinking about beliefs and desires actually decreases our ability to accurately predict behavior.
Finally, there is evidence from response time experiments suggesting that we don’t automatically attribute beliefs to others, but do so only deliberately, with conscious attention.\(^47\) It takes longer for adult humans to answer a probe question about someone’s belief in a “false belief” scenario than it does for them to answer a probe question about the location of the moved object.

Given these problems with the idea that belief attribution is a quotidian and accurate means for predicting behavior, the Machiavellian Social Intelligence hypothesis loses quite a bit of its appeal. Early hominins didn’t need to think about beliefs in order to predict behavior, or to manipulate or deceive. If our ancestors didn’t need to attribute beliefs in order to become better predictors of behavior, yet we do today attribute beliefs in our stories and narratives, what led to this practice? Is there any evolutionary advantage to telling stories about the causes of people’s behavior?

**VI. Normative Folk Psychology and the Mengzian Social Intelligence Hypothesis**

One positive answer might come from the idea that these stories we tell about beliefs and desires reflect reality, and that this folk psychology maps on to scientific psychology. This view is consistent with Humean belief-desire psychology, and it is a popular one. For example, Jerry Fodor would give an answer like this, given his commitment that scientific psychology will vindicate folk psychology by finding causally effective and semantically evaluable physically instantiated states that map onto the beliefs and desires we talk about in our stories.\(^48\) On this view, humans who first started attributing beliefs and desires discovered something true about human psychology.

Another positive answer is agnostic about whether our belief talk originally identified real cognitive states, and this is the answer that I endorse. If we examine when and why we talk about beliefs, it appears that we do so most typically when called on to justify our behavior, or when justifying someone else’s behavior.\(^49\) Given this picture, belief attribution serves a fundamentally normative role, not a fundamentally causal one. The stories we tell about people’s actions as caused by beliefs and desires can be valuable as stories even if they do not get things right.\(^50\)

If this assessment of the way we talk about beliefs reflects how and when we think about beliefs, then we can consider an alternative to the Machiavellian Social Intelligence Hypothesis—we can call it the Mengzian Social Intelligence Hypothesis, after the ancient Chinese philosopher Mengzi who argued that human nature is good, and that human society helps children exercise their pro-social natural tendencies to sustain strong communities. A few years before anyone started talking about Machiavellianism, Allison Jolly put forth her own version of a social intelligence hypothesis.\(^51\) Based on her experience observing the lemurs of Madagascar, where she discovered the little primates organized in female dominant societies, complete with alloparenting, Jolly proposed that the complexities associated with sustaining a cohesive social life and social learning drove the evolution of sophisticated primate cognition.

Jolly’s Mengzian Social Intelligence Hypothesis is consistent with the human practice of telling stories and offering explanations in terms of beliefs in order to explain, and better yet, to justify behavior, as a means of solidifying communities when individuals act outside of the group norms. Consider again the proposal that predicting behavior is the
primary role for belief attribution. In a world without those concepts, the innovator who first started thinking about beliefs to predict behavior would only need to do so when all the other means of predicting behavior failed. This would make the situation in which belief attribution must be used quite unusual—a situation in which a person acts counter to her personality, her social role, and her past behavior. However, to understand this unusual situation we would need to first interpret the situation by offering an explanation of the unusual behavior, before being able to predict what the individual would do next. Someone fails to eat the food in front of her, or vigorously rubs sticks together, and the social partner doesn’t understand why she’s doing what she’s doing, or what she will do next. The social partner needs to explain the behavior before predicting what the individual will do. And that explanation could be in terms of the individuals’ beliefs and desires—reasons for action. Predicting behavior by attributing propositional attitudes relies on a prior ability to explain behavior in terms of the beliefs and desires one should have in order to be a responsible group member. This argument suggests that the Machiavellian hypothesis gets things backwards: conceiving of beliefs and desires could not have been an adaptation for making better predictions of behavior, because before this capacity could provide any additional predictive power, the belief and desire concepts would have to be part of the cognitive repertoire already.

My account of the role of belief attribution presents it as a type of cultural technology that offers two kinds of benefits to the community: promoting group cohesion and supporting the spread of beneficial innovations. Attributing belief allows those who deviate from group norms to remain in-group members, by allowing others to rationalize their actions. When someone acts against the norm, other group members feel discomfort with the unexpected and unusual behavior. By explaining the behavior, they are able to reduce their discomfort and continue living in peace with the individual. Importantly, it doesn’t matter whether these sorts of explanations are true or not. Their function isn’t to predict any other behavior, but to reduce cognitive dissonance. The explanation fills in a missing part of the story, and closes the narrative structure. Our fiction is rife with cases of false justifications used to repair relationships. Mr. Chater forgives Septimus for sleeping with his wife after Septimus falsely tells him that Mrs. Chater proposed the affair in return for a positive review of Mr. Chater’s poetry book. Harry decides that Snape was acting nastily because he wanted to help Dumbeldore’s cause, and not because Snape hated Harry’s father. The truth doesn’t matter, so long as the relationship persists.

Attributing beliefs to those who deviate from typical behavior can also lead to the adoption of beneficial innovations, such as novel tool sets, hunting or fishing techniques, or the control of fire and cooking. With the ability to tell stories about why people are acting, our hominin ancestors could come to understand such anomalous behaviors, see value in them, and adopt them as new community norms. In these cases, it is beneficial for other members of the community to strive for an accurate explanation of the innovators’ behavior. For example, if one saw another person knapping obsidian for the first time, explaining the novel behavior in terms of the knapper’s actual reasons for acting—say, to make a sharp tool for cutting meat—would help in the uptake and spread of the knapping behavior.

But the ability to tell stories about the causes of others’ action can also work when the explanations are false, and when those reasons are not the actor’s own. If we feel better about an individual because we think we understand why they acted as they did, then we will be less likely to ostracize the individual. And the tolerance for different behaviors on
its own can drive technological evolution, if the ends of the novel actions are transparent. For example, the group member who first took hard won meat and put it in destructive fire might have been ostracized by less tolerant group members, but if a norm violation like this was rationalized—whether the rationalization were true or not—the tolerant group member who tasted the cooked meat could independently realize that cooked meat tastes better, and in this way the novel behavior could spread through the community.

This alternative hypothesis for the role of belief attribution is consistent with Fodor’s commitment to the vindication of folk psychology, but it is also consistent with instrumentalist views that portray our belief and desire talk as an interpretive gambit rather than a causal system, as Daniel Dennett would have it. Whether beliefs are understood as functional states, brain processes, dispositions, or instrumental ascriptions, the power that comes from our practice of attributing beliefs is what makes possible the cohesion of large groups of people structured to promote cumulative cultural development.

VII. DO APES NEED TO ATTRIBUTE BELIEFS?

The attempts to answer the question of whether other apes attribute beliefs have been based on false assumptions about the role of belief attribution in human society. If we want to know whether apes attribute beliefs, we should first start by examining whether they need to attribute beliefs. That is, we can ask a series of related questions: Can apes construct explanations? Do apes benefit from innovative behaviors that spread through their society? Do they have group norms, and do they benefit from tolerating violations of group norms? Can apes tell stories with narrative structure?

We have a kind of answer for some of these questions already. Apes seem to understand the causal structure of physical systems, suggesting they may be able to explain—at least to themselves—physical functioning, and can figure out how to access puzzle boxes and process food. Apes also introduce and acquire behavioral innovations. Different ape communities have different practices—chimpanzees crack nuts with rocks and anvils at many different sites across Africa, but in the Lopé Reserve in Gabon, chimpanzees don’t bother eating nuts, not due to some ecological reason, but, it seems, because they don’t know how to crack nuts. Rehabilitant orangutans raised by humans and released onto an island habitat began to swim and fish, behaviors that have never been observed in wild orangutans, and which these rehabilitants likely learned by observing the humans who cared for them.

While apes can engage in causal reasoning, they might not be able to construct causal explanations. For evidence of that we’d have to first examine whether apes engage in explanation seeking—do they look for explanations? Looking for explanations is a response to a wondering why, so this question will also involve examining ape affective states. We can state this as another kind of question—do apes experience cognitive dissonance, and, like humans, attempt to restructure what they know so as to reduce the negative affective state?

Seeking explanations can be understood in terms of the following pattern: (a) the expression of a curiosity state (surprise, interest, horror) directed at an unexpected event, (b) engaging in exploratory behavior associated with the state of affairs, and finally (c) resolving the curiosity state and returning to an emotional baseline. Given this definition, we might be able to find evidence that great apes seek explanations. We do observe apes engaging in exploratory and innovative behaviors. Some exploratory behavior is directed
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

at other agents. For example, in the imitation recognition studies, subjects demonstrate that they know they are being imitated. In one study, a chimpanzee named Cassie would systematically alter his behavior in response to the imitator’s movements, and would do so only while being imitated.\(^5\) When children act this way, developmental psychologists interpret their behavior as the children testing the hypothesis that they are being imitated.\(^5\) If Cassie was testing the hypothesis that he was being imitated, he was also engaged in explanation seeking behavior—he would have started in a curiosity state that motivated him to generate the hypothesis, and test it.

Better evidence of great ape explanation seeking might come from examining the topography of the behavioral pattern as a whole: a facial expression or bodily posture indicating a curiosity state; directed exploratory behavior; then finally a resolution of the curiosity state, indicated by a satisfaction facial expression or bodily posture. Using research on emotions and facial expressions in chimpanzees, including a chimpanzee version of Paul Ekman’s FACS (Facial Action Coding System), we can classify the emotions expressed by great apes.\(^6\)

While it may be difficult enough a task to examine whether chimpanzees seek explanations, the questions about norms and narratives are even more difficult to answer. Chimpanzees might have social norms in their societies, such as a norm against infanticide, if norms are defined as group behavioral regularities, the violation of which provokes negative reactions by bystanders.\(^6\) But if norms are consistent with the existence of anti-normative behavior and a lack of negative bystander reaction, they are going to be much more difficult to observe. Some human norms, like the norms of marriage in contemporary industrialized society, are routinely violated, but so long as something that takes the role of a justifying explanation is provided, even an “explanation” such as irreconcilable differences, we tend to accept the anti-normative behavior. The same can be said of contra-normative behaviors such as being late to an appointment, truth telling, and even cases of public nudity and, sometimes, killing. Such is the power of justification.

Great apes clearly can’t justify others’ behavior the same way we can. They can’t tell stories. But can they think stories? The King Kong experiment can be interpreted as a test of story comprehension—the apes anticipate the human will look in the rightmost haystack because that’s where he should look for King Kong. It’s not unlike Premack and Woodruff’s original test, in which Sarah the chimpanzee was asked what comes next in a familiar narrative. Cold rooms require heating, and to start the gas heater the person needs a match; locked cages need unlocking, and to unlock the cage the person needs a key. These are simple stories, indeed, but experiments that explicitly set out to test story comprehension can help to shed light on whether humans alone can follow, or enjoy, more complicated stories. Anecdotal reports of captive great apes enjoying soap operas and experimental evidence that chimpanzees yawn contagiously in response to animations of yawning chimpanzees offer some suggestive reason for thinking that apes might respond to fictional stories and characters.\(^6\)

Furthermore, in our study of pantomime communication in orangutans, Anne Russon and I report on one pantomime incident that is suggestive of narrative structure in the context of production. Kikan, an infant orangutan, was observed to re-enact part of a past event. Here’s the background. A researcher who is well versed with orangutan behavior and forest medicines, Agnes Ferisa, was observing a group of rehabilitant orangutans, as was her practice. She noticed that one of the orangutans, Kikan, was mouthing her foot. Kikan,
who in the past had not been particularly friendly toward Agnes, allowed her to approach and examine the foot. Agnes saw that Kikan had a sliver in the sole of her foot, so she used her pencil to remove it. To close the wound, Agnes broke a leaf stem off a ficus tree, and used the latex to seal the hole. A week later, Kikan approached Agnes, pulling on Agnes’s leg a few times before getting her attention. When Agnes turned to look at Kikan, the orangutan lifted up her foot, then picked a leaf and poked its stem at the sole of her (now healed) foot, just as Agnes had done while doctoring her. Russon and I argue that Kikan’s pantomime reveals narrative thinking in the orangutan.63

VIII. Refocusing the Lens

Apes may or may not have developed the cognitive technology of belief attribution. We won’t know that until we start asking the right sorts of questions. What we should see, however, is that the questions we have been asking have started with some problematic assumptions about human nature, and our evolutionary social environment—our puzzle pieces were flawed to begin with. When we adopt a lens through which to see the world, we frame our future movements. If that lens is faulty, it can lead us far from where we want to be. If we want to know what makes humans different from other animals, we need to begin with a clear picture of what kind of animals humans are. At least sometimes, we are animals who think we are more special than we really are.

The Mengzian Social Intelligence Hypothesis offers a different lens through which to see human social evolution. Our evolutionary history of life in cohesive groups with social norms--groups that tolerate just enough contra-normative behavior to permit the introduction of technological innovations and maintenance of large communities--may have led us to create stories about violations. These stories gave us the power to feel like we understand others, even when we do not.

Notes

2. Byrne and Whiten, Machiavellian Intelligence.
3. Scott and Baillargeon, “Early False-Belief Understanding.”
4. Buttelmann et al., “Great Apes Distinguish True from False Beliefs in an Interactive Helping Task”; Krupenye et al., “Great Apes Anticipate that Other Individuals Will Act According to False Beliefs.”
5. Andrews, Do Apes Read Minds?
7. Premack and Woodruff, “Does the Chimpanzee Have a Theory of Mind?”
8. Ibid., 515.
10. Wimmer and Perner, “Beliefs about Beliefs.”
12. Call and Tomasello, “A Nonverbal False Belief Task.”
Andrews: Do Apes Attribute Beliefs to Predict Behavior?

14. Onishi and Baillargeon, “Do 15-Month-Old Infants Understand False Beliefs?”
16. Ibid., 590.
17. Krupenye et al., “Great Apes Anticipate.”
18. Buttelmann et al., “Great Apes Distinguish.”
22. E.g. Schwitzgebel, “Belief.”
25. Flavell, “The Development of Inferences about Others.”
27. Krupenye et al., “Great Apes Anticipate.”
29. Call et al., “‘Unwilling’ versus ‘Unable.’”
32. Whiten and Byrne, “Tactical Deception.”
33. Hare et al., “Chimpanzees Know”; Hare et al., “Do Chimpanzees Know”; Karg et al., “Differing Views.”
34. Wittig et al., “Social Support.”
35. Musgrave et al., “Tool Transfers.”
38. Mitani and Watts, “Boundary Patrols.”
41. Andrews, Do Apes Read Minds?
44. Andrews, Do Apes Read Minds?
45. Apperly, Mindreaders; Zawidzki, Mindshaping.
46. Wilson and LaFleur, “Knowing What You’ll Do.”
47. Apperly et al., “Is Belief Reasoning Automatic?,” 841; Back and Apperly, “Two Sources of Evidence.”
References


Andrews: Do Apes Attribute Beliefs to Predict Behavior?


Meltzoff, A. N. “Foundations for Developing a Concept of Self: The Role of Imitation in Relating Self to Other and the Value of Social Mirroring, Social Modelling, and Self
Andrews: Do Apes Attribute Beliefs to Predict Behavior?


