Anthropomorphism, Parsimony, and Common Ancestry

ELLIOTT SOBER

Abstract: I consider three theses that are friendly to anthropomorphism. Each makes a claim about what can be inferred about the mental life of chimpanzees from the fact that humans and chimpanzees both have behavioral trait B and humans produce this behavior by having mental trait M. The first thesis asserts that this fact makes it *probable* that chimpanzees have M. The second says that this fact provides *strong evidence* that chimpanzees have M. The third claims that the fact is *evidence* that chimpanzees have M. The third thesis follows from a plausible Reichenbachian model of how a common ancestor is probabilistically related to its descendants. The first two theses do not, and they have no general evolutionary justification.

Morgan's (1894) canon has exerted a large influence on the study of mind and behavior, both human and nonhuman. The canon says that higher mental faculties should not be postulated to explain a behavior if the behavior can be explained by postulating lower mental faculties alone. Morgan did not suggest that this principle is a priori correct; rather, he tried to harness evolutionary theory to justify it. Morgan's evolutionary argument does not work (Sober, 1998), so the question arises of whether some other justification can be provided.

My project here is not to inquire further into the justification of Morgan's canon, but to consider an evolutionary argument for something like the canon's mirror image. Morgan argued that evolutionary considerations should make us wary of anthropomorphism; the opposite line of thought is to argue that evolutionary theory lends support to anthropomorphic inferences. This is the kind of position that De Waal (1991) presents. Here are two of his formulations:

> By far the simplest assumption regarding the social behavior of the chimpanzee, for example, is that if this species' behavior resembles that of ourselves then the underlying psychological and mental processes *must* be similar too. To propose otherwise requires that we assume the evolution of divergent processes for the production of similar behavior (p. 298).

> The most parsimonious assumption concerning nonhuman primates is that if their behavior resembles human behavior the psychological and mental processes involved are *probably* similar too (p. 316).

My thanks to David Baum, Michael A. Bell, Branden Fitelson, Matt Kopec, Daniel Schneider, Larry Shapiro, Karen Strier, David S. Wilson, and an Editor of this journal for useful discussion.

Address for correspondence: Philosophy Department, 5185 Helen C. White Hall, University of Wisconsin, Madison, WI 53706, USA.
Email: ersober@wisc.edu

Mind & Language, Vol. 27, No. 3 June 2012, pp. 229–238.

I italicize 'must' and 'probably' to mark a difference to which I will return. In these passages as well as in others, De Waal isn't noting what parsimony considerations entail in order to distance himself from a dictum he disdains; on the contrary, his point is to endorse the recommendations that parsimony considerations provide (De Waal 1999, 2009).

Karin-D'Arcy (2005, pp. 185-86) criticizes De Waal's argument on the ground that it appeals to introspection and also on the ground that it bases its inference on too small a sample. The first criticism can be addressed by dropping the mention of introspection. Instead of starting with the premise that we know *by introspection* that human beings have mental characteristic M, the argument can simply start with the more modest statement that human beings have mental characteristic M. The argument then is merely one of extrapolation—from us to another species. As such, the form of argument applies with equal force to nonmental traits. For example, if human beings transport oxygen to their tissues by using the hemoglobin molecule, the argument will conclude that the most parsimonious hypothesis is that chimpanzees use the same proximate mechanism if they too transport oxygen. Karin-D'Arcy's second criticism of De Waal is less easily countered. How strong an inference can one draw about the characteristics of chimpanzees based just on the observation of a single, albeit closely related, species? Surely the conclusion that chimpanzees *must* be like us is too strong. And perhaps the conclusion that they *probably* are like us is too strong as well.

Even if De Waal's formulations go wrong, there remains a more modest anthropomorphic thesis with which to reckon. This is the claim that if human beings and a closely related species (e.g. chimpanzees) both exhibit behavior B, and if human beings produce B by occupying mental state M, then this is *evidence* that M is also the proximate mechanism that chimpanzees deploy in producing B. It is useful to understand this thesis by using the Bayesian conception of confirmation:

> Given background assumptions A, observation O confirms hypothesis H if and only if $\Pr_A(H|O) > \Pr_A(H)$.

The 'A' subscript on the probability function means that probabilities are assigned on the assumption that A is true. For Bayesians, confirmation means probability-raising. Applied to the case at hand, the thesis is that:

(*) $\Pr_A$(chimpanzees have M | humans have M) > $\Pr_A$(chimpanzees have M).

The background assumption A in (*) is that humans and chimpanzees both have the behavioral trait B. The idea I wish to explore is that the fact that humans and chimpanzees have a common ancestor provides a justification for this inequality. Notice that proposition (*) does not entail that $\Pr_A$(chimpanzees have M|humans have M) > ½, which is why this evidential thesis about anthropomorphism is more modest than the second of De Waal's two formulations, quoted above.

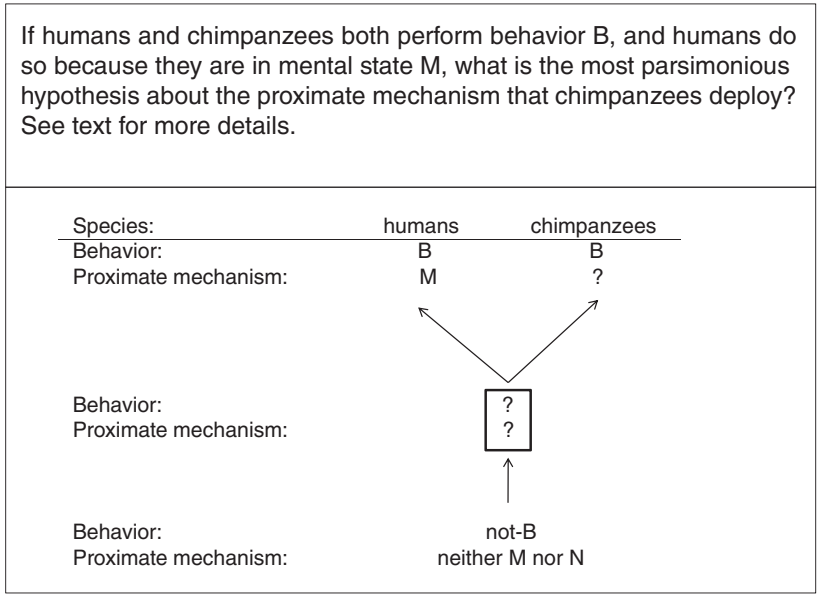

Figure 1 *Cladistic parsimony solution*

Before discussing whether (*) is true, I want to describe how parsimony leads to the kind of conclusion that De Waal endorses. The concept of parsimony that De Waal has in mind is *cladistic* parsimony, according to which the complexity of an evolutionary hypothesis is measured by counting how many changes in character state the hypothesis requires. The problem that De Waal is addressing, and the solution that cladistic parsimony provides, are represented in the accompanying figure (Figure 1). We observe that humans and chimpanzees perform behavior ***B***, and we know that humans do so by using the proximate mechanism ***M***. We assume that humans and chimpanzees have a common ancestor. The box in the figure depicts their most recent common ancestor (MRCA); below that, a more ancient common ancestor is represented, which is assumed to lack the behavioral trait ***B***. What inference should be drawn concerning the proximate mechanism that chimpanzees deploy? Suppose, to make the problem simple, that there are exactly two proximate mechanisms (***M*** and ***N***) that could produce the behavior ***B*** and that each proximate mechanism suffices for ***B*** to occur.

Since humans and chimpanzees perform behavior ***B***, the most parsimonious hypothesis about the MRCA is that it performed ***B*** as well. This assignment is most parsimonious because it requires only a single change in character state—from *not B* in the ancient ancestor to ***B*** in the MRCA. Assigning *not B* to the MRCA is less parsimonious, since it requires that at least two changes must have occurred—one in the lineage leading from the MRCA to humans, the other in the lineage leading from the MRCA to chimpanzees. The parsimonious assignment of ***B*** to MRCA

means that the MRCA used either *M* or *N*. But which of these two proximate mechanisms should be assigned to the MRCA, and which to chimpanzees? The most parsimonious hypothesis is that both use *M*. Cladistic parsimony endorses anthropomorphism in this simple inference problem.[1] Notice that the logic of this inference applies to any pair of related species that share a phenotype where the proximate mechanism that causes this phenotype is known for one species but unknown for the other. The cladistic argument has nothing essentially to do with *us* or with the fact that the proximate mechanism is *psychological*.[2]

Cladistic parsimony now is controversial in evolutionary biology (Sober, 1988, 2008, 2011). Many evolutionary biologists appeal to parsimony when they make phylogenetic inferences, but many others prefer to use statistical inference procedures that explicitly assume a particular model of the evolutionary process. Skeptics about parsimony may be inclined to reject De Waal's parsimony argument in favor of anthropomorphism[3] and to draw the same negative verdict about proposition (*). This, I suggest, would be to conflate the bathwater and the baby. Proposition (*) can be defended on grounds that have nothing to do with parsimony.

The argument I have in mind for (*) is due to Hans Reichenbach (1956), who wasn't thinking about anthropomorphism or evolution when he made it. His more general and abstract project was to describe a probability model in which two effects trace back to a common cause. Reichenbach's model applies to the problem at hand if we think of human beings and chimpanzees as effects and their most

[1] How might bonobos (which are just as closely related to humans as chimpanzees are) and gorillas be brought into the picture? If bonobos have the behavioral trait *B*, this strengthens the parsimony inference that the MRCA of humans and chimpanzees has *B*. If bonobos lack *B* and gorillas do too, this raises the question of whether the presence of *B* in humans and chimpanzees is a homoplasy; if it is, parsimony would not favor attributing *M* to chimpanzees.

[2] The literature on theory of mind contains another kind of parsimony argument, one that does not discuss the evolutionary process, but rather has to do just with how an input/output device works now. This line of argument traces back to Morgan's Canon and aims to impugn anthropomorphism. Povinelli and Vonk (2003), for example, argue that a model that postulates behavior reading only is more parsimonious than a model that postulates behavior reading and mind reading both. Sober (2001) argues, to the contrary, that introducing an additional intervening variable can sometimes increase a model's parsimoniousness.

[3] Povinelli and Vonk (2003, p. 157) make the following argument: 'The first possibility is that the chimpanzee's mind seems similar to ours precisely because it is similar. Biological parsimony would seem to support such an assumption: chimpanzees and humans arose from a common ancestor about six million years ago. Alas, invoking biological parsimony will not help. After all, humans and chimpanzees are different in several other important ways, but this in no way denies their evolutionary relatedness. By way of analogy, the fact that some bats echolocate, but their closest living relatives do not, hardly constitutes a crisis for evolutionary theory.' Povinelli and Vonk's argument succeeds in showing that the existence of a common ancestor 6 million years ago does not *entail* that humans and chimps must be psychologically similar, but it does not undermine more modest theses that are friendly to anthropomorphism. Bolhuis and Wynne (2009, p. 833) assert that 'evolutionary analyses ... because they are analyses of history, cannot uncover how an animal achieves a particular feat', but they do not explain why parsimony arguments about proximate mechanisms are not to be trusted.

recent common ancestor as the common cause. Here is Reichenbach's general idea: Suppose two effects E_1 and E_2 and their common cause C each occupy one of two states (coded as 0 and 1). I will write '$E_1 = 0$' to represent the proposition that E_1 is in state 0, and similarly for the other events and states. Reichenbach's model of how these events are related to each other involves three assumptions:

(i) $\Pr(C{=}i)$, $\Pr(E_1{=}j|C{=}k)$, and $\Pr(E_2{=}j|C{=}k)$ have values between 0 and 1 noninclusive, for all i, j, k.

(ii) $\Pr(E_1{=}i|C{=}j){=}\Pr(E_1{=}i|C{=}j\ \&\ E_2{=}k)$, for all i, j, k.

(iii) $\Pr(E_1{=}i|C{=}i){>}\Pr(E_1{=}i|C{=}j)$ and $\Pr(E_2{=}i|C{=}i){>}\Pr(E_2{=}i|C{=}j)$, for all $i{\neq}j$.

In other words: (i) all probabilities are strictly between 0 and 1, (ii) the common cause screens-off the two effects from each other (i.e. the probability of one effect's being in a given state, given the state of the common cause, is the same regardless of what state the other effect happens to occupy),[4] and (iii) there is a positive correlation between the common cause and each of its effects. Reichenbach showed that these three assumptions entail that the two effects will be positively correlated:

$$\Pr(E_1{=}i\&E_2{=}i){>}\Pr(E_1{=}i)\Pr(E_2{=}i), \text{ for each } i.$$

This correlation entails that:

$$\Pr(E_2{=}i|E_1{=}i){>}\Pr(E_2{=}i), \text{ for each } i.$$

If E_2 and E_1 are chimpanzees and humans, C is their most recent common ancestor, character state i is the mental state M, and we consider a probability function that embeds the assumption that humans and chimpanzees both have behavioral trait B, proposition (*) follows.

Reichenbach's assumptions apply to many common causes and their effects.[5] In addition, these assumptions are routine in the probabilistic models of the evolutionary process that biologists use. Reichenbach's model is neutral on whether evolution is driven by drift or by selection, and also on whether rates of evolution are constant or shifting. Common ancestry, understood in terms of Reichenbach's model, justifies proposition (*).

As noted above, proposition (*) says nothing about what the absolute value is of $\Pr_A$(chimpanzees have M| humans have M). Even so, (*) entails that estimating

[4] This is equivalent to each state of the common cause's rendering the two effects conditionally independent of each other.

[5] It is not an assumption of Reichenbach's model that the common cause C is the *only* cause of E_1, nor is it assumed that C is the only cause of E_2; E_1 and E_2 may each have their separate causes and Reichenbach's argument still goes through (Sober and Barrett, 1992, p. 10).

the value of $\Pr_A$(chimpanzees have M) would place a lower bound on the value of the conditional probability of interest. I suspect, however, that nothing much can be done with this idea. For example, consider the fact that the frequency of behavior B in the species that now populate the earth is something we can estimate from observations. If B is a necessary condition for M, the frequency of M can't be greater than the frequency of B. Does this fact about frequency suffice to justify the conclusion that $\Pr_A$(chimpanzees have M) is less than the frequency of B? Even if it does, the result is useless. An *upper* bound on $\Pr_A$(chimpanzees have M) does not allow (*) to place a *lower* bound on $\Pr_A$(chimpanzees have M| humans have M).

Another question may be more tractable. Proposition (*) says that observing that humans have M is evidence that chimpanzees do too. But how strong is the evidence thus provided?[6] Here we need a measure of the degree to which observation O confirms hypothesis H. Several measures of Bayesian *degree of confirmation* have been proposed; see Fitelson (1999) for a review of the pluses and minuses of each. The measure I want to consider here is the likelihood ratio:

$$\mathrm{DOC}(O, H) = \frac{\Pr(O|H)}{\Pr(O|not\,H)},$$

so the next question to investigate is what makes the quantity

$$\frac{\Pr(\text{humans have } M|\textit{chimpanzees have } M)}{\Pr(\text{humans have } M|\textit{chimpanzees lack } M)}$$

large and what makes it small.[7] Reichenbach's model entails that this ratio is larger than one; the question is how much larger than one it is.

First some notation: I will discuss three taxa, humans (H), chimpanzees (C), and their most recent common ancestor (A). I will write 'C1' to represent the proposition that chimpanzees have mental trait M and 'C0' to represent the proposition that they do not. Similarly for the other taxa and their possible character states.[8] The above likelihood ratio can be expanded as follows:

$$\frac{\Pr(H1|C1)}{\Pr(H1|C0)} = \frac{\Pr(H1|A1)\Pr(A1|C1) + \Pr(H1|A0)\Pr(A0|C1)}{\Pr(H1|A1)\Pr(A1|C0) + \Pr(H1|A0)\Pr(A0|C0)}$$

[6] Notice that the cladistic argument for anthropomorphism, *per se*, does not address this quantitative question. Parsimony just provides an ordering of how well supported various hypotheses are.

[7] I drop the subscript A from the probability function in stating the question I want to pursue because I don't want to assume that the most recent common ancestor of Humans and Chimpanzees has the behavioral trait B.

[8] So 'A0' means that the ancestor A either has N or has neither M nor N (in which case it lacks the behavior B). Unlike the formulation of proposition (*), the present analysis does not assume that the taxa have B.

The right-hand side of this equation has the form

$$(^{**}) \qquad \frac{ax + b(1-x)}{ay + b(1-y)}.$$

Notice that the numerator and denominator in (**) each take a weighted average of a and b. As for a and b themselves, it is standard to assume that $a > b$; this just means that a descendant (H) and its ancestor (A) are positively correlated, which is a given in all Markov models of evolution (Sober, 2008, pp. 215-6). If a isn't much bigger than b, then there is no way for (**) to be much greater than unity. For example, if $a = 0.7$ and $b = 0.6$, the largest value (**) could have is $0.7/0.6 = 1.2$.[9]

The values of a and b— namely of Pr(H1|A1) and of Pr(H1|A0)—depend on the amount of time separating the most recent common ancestor from present day humans *and* on the probability per unit time of the lineage's changing character state. The fact that this lineage has a duration of about 6 million years does not suffice to say anything about the values of a and b, though it is true (for fixed probabilities of change per unit time) that the more time there is between the MRCA and human beings, the closer these two probabilities will be to each other; they start with values of 1 and 0 when the lineage has a duration of zero and then edge towards each other as time marches on. It is interesting that this very basic consideration about the value of the ratio (**) concerns human evolution alone; chimpanzees have nothing to do with it.

The chimpanzee side become relevant if a and b are very different. For example, if $a = \Pr(H1|A1) = 0.9$ and $b = \Pr(H1|A0) = 0.01$, the values of x and y are worth considering. The ratio (**) can be rewritten as:

$$\frac{x(a-b) + b}{y(a-b) + b}.$$

Since a>b, this ratio is made larger by maximizing x and minimizing y. Bayes' Theorem allows each of those terms to be expanded as follows:

$$x = \Pr(A1|C1) = \frac{\Pr(C1|A1)Pr(A1)}{\Pr(C1)}$$

$$y = \Pr(A1|C0) = \frac{\Pr(C0|A1)Pr(A1)}{\Pr(C0)}.$$

Given the common term Pr(A1), the way to make x large and y small is to make Pr(C1|A1) large and Pr(C1) small. What can be said about these two quantities?

[9] Although there is no precise cut-off separating strong evidence from weak, Royall (1997) recommends using a likelihood ratio of at least 8 to define what strong evidence is.

For Pr(C1|A1) to be large, it must be true that chimpanzees would probably have ***M*** if the most recent common ancestor they share with humans had that trait. The question here is not whether the most recent common ancestor in fact had ***M***; the question is whether chimps would probably retain that trait *if* the ancestor already possessed it. Perhaps the answer is *yes*, on the ground that selection would favor the retention of ***M*** once it was in place. What about Pr(C1) being small? If behavior ***B*** is rare among current taxa, then the frequency of ***M*** must be lower still. Perhaps this provides a reason to think that the prior probability Pr(C1) is small. So the question about strength of evidence—that is, of how much larger than one the ratio (**) is—needs to be answered by attending to the values of the four parameters just discussed. Proposition (*) follows from weak Reichenbachian assumptions, and this entails that the ratio (**) is *bigger* than one, but further assumptions are need to say whether (**) is *much bigger* than one.[10]

De Waal's assertion that 6 million years is too short a time for humans and chimpanzees to have different proximate mechanisms for producing a shared behavioral trait ***B*** should raise eyebrows when viewed against the backdrop of the growing number of documented cases of rapid evolution, especially ones in which different proximate mechanisms evolve that produce 'the same' phenotype. Consider, for example, the evolution of lactose tolerance, a trait that some 25% of present day human beings possess. This phenotype evolved since agriculture came on line about 10,000 years ago, and the trait evolved multiple times, with different gene complexes evolving in different human populations (Tishkoff *et al*. 2007). This example does not show that 6 million years suffices for humans and chimps to evolve different mental mechanisms (***M*** and ***N***) for producing behavior ***B***. But it does show that the claim that '6 million years is not enough' requires a detailed argument that focuses on the specifics of the traits and the species involved. Cladistic parsimony does not furnish that argument;[11] neither does the fact that 6 million years is a mere flicker in the 3.8 billion years since life on earth began.

Biologists are now exploring data and theory concerning how two species can have the same phenotype though the underlying proximate mechanisms differ. True and Haag (2001) call the process that can produce this outcome 'developmental systems drift', but they do not mean to restrict the process to neutral evolution. For example, two species might exhibit behavior B2 as a pleiotropic effect of different gene complexes. Suppose that in the first species gene G1 now causes B1 and B2

[10] The distinction between the truth of the statement '***O*** confirms ***H***' and an assessment of the degree to which ***O*** confirms ***H*** also needs to be drawn in connection with the first step of the cladistic analysis of the problem depicted in the figure. Weak assumptions entail that if humans and chimps have the behavioral trait ***B***, that this is evidence that MRCA has ***B*** as well. These assumptions do not suffice to answer the question of how strong the evidence is that MRCA has ***B*** (Sober ,2011, pp. 155-57).

[11] Notice that the cladistic argument represented in the Figure is not affected by how long ago the MRCA was. This shows that the cladistic argument is insensitive to one of the factors that influences strength of evidence.

whereas in the second G2 now causes B2 and B3. Selection for B1 in the first species or for B3 in the second can be why the proximate mechanisms diverged. Selection can lead to this 'unparsimonious' solution even when B2 is present in both species *and* in their most recent common ancestor. True and Haag (p. 110) say that 'at the short end of the time scale, there is a great deal of evidence of DSD between recently diverged species.'

I conclude that there is no general evolutionary reason for thinking, when both humans and chimpanzees produce behavior ***B***, that the fact that humans have mental state ***M*** is *strong* evidence that chimpanzees have ***M*** as well. But drop the word 'strong' and a weaker thesis reappears: the fact that humans have ***M*** is *evidence* that chimpanzees do too. This modest anthropomorphism is the kernel that remains after the chaff has been discarded.

Department of Philosophy
University of Wisconsin, Madison

References

Bolhuis, J. and Wynne, C. 2009: Can evolution explain how minds work? *Nature,* 458, 832–3.

De Waal, F. B. M. 1991: Complementary methods and convergent evidence in the studyof primate social cognition. *Behaviour,* 118, 297–320.

De Waal, F. B. M. 1999: Anthropomorphism and anthropodenial—consistency in our thinking about humans and other animals. *Philosophical Topics, 27*, 255–80.

De Waal, F. B. M. 2009: Darwin's last laugh. *Nature,* 460, 175.

Fitelson, B. 1999: The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science,* 66, S362–S378.

Karin-D'Arcy, M. R. 2005: The modern role of Morgan's canon in comparative psychology. *International Journal of Comparative Psychology*, 18, 179–201.

Morgan, C. L. 1894: *Introduction to Comparative Psychology*. London: Walter Scott.

Povinelli, D. and Vonk, J. 2003: Chimpanzee minds: suspiciously human? *Trends in Cognitive Science*, 7, 157–60.

Reichenbach, H. 1956: *The Direction of Time.* Berkeley, CA: University of California Press.

Royall, R. 1997: *Statistical Evidence—A Likelihood Paradigm*. Boca Raton, FL: Chapman and Hall.

Sober, E. 1988: *Reconstructing the Past—Parsimony, Evolution, and Inference.* Cambridge, MA: MIT Press.

Sober, E. 1998: Morgan's canon. In D. Dellarosa Cummins and C. Allen (eds), *Evolution of Mind*. Oxford: Oxford University Press.

Sober, E. 2001: The principle of conservatism in cognitive ethology. In D. Walsh (ed.), *Naturalism, Evolution, and Mind*. Cambridge: Cambridge University Press.

Sober, E. 2008: *Evidence and Evolution—The Logic Behind the Science*. Cambridge: Cambridge University Press.

Sober, E. 2011: *Did Darwin Write the Origin Backwards?* Amherst, NY: Prometheus Books.

Sober, E. and Barrett, M. 1992:Conjunctive forks and temporally asymmetric inference. *Australasian Journal of Philosophy,* 70, 1-23.

Tishkoff, S., Reed, F., Ranciaro, A., Voight, B., Babbitt, C., Silverman, J., Powell1, K., Mortensen, H., Hirbo, J. Osman, M., Ibrahim, M., Omar, S., Lema, G., Nyambo, T., Ghori, J., Bumpstead, S., Pritchard, J.,Wray, G. and Deloukas, P. 2007: Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics,* 39, 31–40.

True, J. and Haag, E. 2001: Developmental system drift and flexibility in evolutionary trajectories. *Evolution and Development*, 3, 109–19.