

Anthropomorphism, anthropectomy, and the null hypothesis

Kristin Andrews · Brian Huss

Received: 30 July 2013 / Accepted: 11 March 2014
© Springer Science+Business Media Dordrecht 2014

Abstract We examine the claim that the methodology of psychology leads to a bias in animal cognition research against attributing “anthropomorphic” properties to animals (Sober in *Thinking with animals: new perspectives on anthropomorphism*. Columbia University Press, New York, pp 85–99, 2005; de Waal in *Philos Top* 27:225–280, 1999). This charge is examined in light of a debate on the role of folk psychology between primatologists who emphasize similarities between humans and other apes, and those who emphasize differences. We argue that while in practice there is sometimes bias, either in the formulation of the null hypothesis or in the preference of Type-II errors over Type-I errors, the bias is not the result of proper use of the Neyman and Pearson hypothesis testing method. Psychologists’ preference for false negatives over false positives cannot justify a preference for avoiding anthropomorphic errors over anthropectic (Gk. anthropos—human; *ektomia*—to cut out) errors.

Keywords Animal cognition · Mindreading/theory of mind · Ape cognition · Null hypothesis · Anthropomorphism · Hypothesis testing

Introduction

In the context of animal cognitive research, “anthropomorphism” can be defined as the attribution of human psychological, social, or normative properties to non-human animals. However, the term is often defined as an error—a misattribution of a human property to a nonhuman animal. For example, the author of the foremost textbook on animal cognition, Sara Shettleworth, defines anthropomorphism as “the attribution of human qualities to other animals, usually with the implication it is

K. Andrews (✉) · B. Huss
Department of Philosophy, York University, Toronto, M3J 1P3, Canada
e-mail: andrewsk@yorku.ca

done without sound justification” (Shettleworth 2010a, 477). We will focus on the second definition in order to investigate errors that might be made in controversial areas of animal research including mindreading, episodic memory, and error monitoring. We think there is a problem with the epistemology associated with the methods of determining whether animals have these properties that are uncontroversially attributed to adult humans.

One way psychologists justify their claims about what an animal can do is to use Neyman and Pearson’s (1928) hypothesis testing methods: a null hypothesis is devised—a hypothesis that reflects what is expected to be the norm, and against which the researcher is looking for a statistically significant discrepancy. Data is collected, analyzed, and the results are reported and interpreted. For experiments designed to investigate whether an animal has a particular psychological, social, or normative property had by humans, the null hypothesis is typically that the animal does not have the property in question. When a purportedly human property is attributed to an animal without prior methodologically sound investigation of this sort, that attribution is considered anthropomorphic. In the current debate about the role of anthropomorphism in animal cognition research, some scholars have raised the concern that the standard psychological methods result in a bias against attributing properties to animals when those properties are seen as somehow *special* human (Sober 2005; de Waal 1999)—special because they are psychological, social, or normative properties that have been identified as potential markers for human uniqueness. We disagree, arguing here that on no interpretation do these methods lead to such a bias.

There has been much recent discussion about the charge of anthropomorphism (e.g. essays in Mitchell et al. 1997; Datson and Mitman 2005). We can categorize those skeptics who are particularly worried about anthropomorphism into two types: categorical skeptics who think that animal cognition research cannot be good science, and selective skeptics who think that some kinds of attributions are not justified. Many of the defenses of animal cognition research have addressed categorical skeptics, such as Kennedy (1992), who think that animal cognition research is an unscientific field of research. These skeptics think that animals are not the right sorts of things to apply the concepts to. For categorical skeptics, the charge of anthropomorphism is a pre-empirical one. They think that researchers in animal cognition are making a category mistake by asking whether animals have certain properties (for this critique see Bekoff and Allen 1997; Fisher 1990, 1991; Keeley 2004).

If the charge of anthropomorphism is a pre-empirical one, the justification for it must be philosophical, in the sense that either the concepts appealed to in the charge are defined as uniquely human, or the nature of the concept or topic under investigation, added to some well-established empirical or theoretical claims, entails that some features are unique to humans. While there are philosophical arguments against the existence in animals of some human psychological properties, such as having belief (Davidson 1975, 1982; Stich 1979) or consciousness (Carruthers 2000, 2004), these arguments are quite controversial, and should not be taken to be so well established as to undermine an entire research program (and we expect that Carruthers would be loath to have his work used that way). Responses to the

categorical skeptics amount to the charge that they are begging the question (Fisher 1990; Keeley 2004). We are sympathetic with this analysis.

In this paper, rather than adding to the arguments against the categorical skeptics, we address the selective skeptics, specifically those skeptics who are also animal cognition researchers. Note first that the field of animal cognition research is by no means a unified one. Researchers come from a variety of disciplines, including anthropology, biology, and psychology, and they use different methods for collecting data, including observational field studies, non-invasive field experiments, and a variety of laboratory experiments and observational studies. Researchers study a variety of species across taxa, and develop different communities, research questions, and standards based on their shared interest in a species or other taxa. The animal subjects in these studies also live in a range of different kinds of settings, including natural habitat, sanctuaries in or near their typical habitat, zoos with different social or environmental conditions, human-like research settings, and laboratory cages or tanks. As may be expected with a field as diverse as this one, there is disagreement from within about how to approach the research into animal cognition.

One division among animal cognition researchers that is particularly apparent among ape researchers is between those who emphasize the similarities between humans and nonhuman animals, and those who emphasize the differences. Both camps justify their position by appeal to evolutionary considerations. The ethologist Frans de Waal, for example, argues that when we see similarities in behavior between humans and other apes, we should expect to see similarities in cognitive processes and functions, because the similarities in behavior suggest that the individuals derived from a common ancestor. He writes, “The...cladistic rationale applied to humans and their close relative should lead us to adopt cognitive similarity as the default position, thus making anthropomorphism a virtual nonissue” (de Waal 1999, 259). Other primatologists clearly feel the same way about interpreting ape behavior. In the introduction to his book on chimpanzee culture, the anthropologist William McGrew writes, “Tickle a chimpanzee, and she laughs; startle a chimpanzee, and he grimaces; threaten a chimpanzee, and she lashes out; groom a chimpanzee, and he sprawls relaxed. All of these signals of feelings are recognized readily by the average person. More dramatically, when we see an orphaned ape with her dead mother, her demeanor or ‘body language’ is one that, if seen in a human child, would be interpreted as grief” (McGrew 2004, 8–9).

Other animal cognition researchers express great concern about this way of describing other species. One of the foremost selective critics, the psychologist Daniel Povinelli, has vigorously argued that contemporary animal cognition researchers are too eager to undermine claims of human uniqueness. The worry is both that the science of animal cognition is harmed when the researchers assume similarity, because very real differences will not be discovered, and that current biology is inconsistent with views like de Waal’s that there is cognitive continuity between closely related species (e.g. Povinelli and Bering 2002; Povinelli et al. 2000; Penn et al. 2008). For example, Povinelli and Bering write, “if the dramatic resculpting of the human body and brain that occurred over the past 4 million years or so involved the evolution of some qualitatively new cognitive systems, then this

insistence on focusing on similarities will leave comparative psychologists unable to investigate hallmarks of their own species—or chimpanzees, for that matter. It [seeking to find similarities across species] is an agenda that does justice to no one” (Povinelli and Bering 2002, 116). Similar concerns are shared to some degree by other animal cognition researchers (e.g. Shettleworth 2010a, b; Silk 2002; Blumberg and Wasserman 1995; Wynn 2004, 2007). For selective skeptics, the null hypothesis is that animals do not have human-like cognitive systems, social relations, or normative properties. Further, at least some selective skeptics claim that current research supports such hypotheses (despite the methodological prohibition against affirming the null, which we will discuss below).

We argue that the special worry about anthropomorphism as expressed by selective skeptics such as Povinelli is unwarranted. We do two things. First we challenge the idea that the special human properties can be unproblematically identified, and hence that the null hypothesis can be unproblematically stated. We conclude that in some cases animal cognition researchers are not in a position to decide, before they engage in research, what the null hypothesis should be.

We then examine the concern that the Neyman and Pearson method leads to a bias against attributing some special human properties to animals. The worry is that scientists who choose the skeptical hypothesis as the null bias their results against finding some similarity. If, on the other hand, scientists who emphasize similarity were to choose an optimistic null hypothesis, the results would be biased in the other direction. For example, Elliott Sober identifies a bias in animal cognition research as due to the acceptance of the rule of thumb according to which it is better to fail to reject a false null hypothesis than it is to mistakenly reject a true null hypothesis. This he traces to the use of Morgan’s Canon—an earlier rule of thumb which states “in no case is an animal activity to be interpreted in terms of higher psychological processes, if it can be fairly interpreted in terms of processes which stand lower in the scale of psychological evolution and development” (Morgan 1903, 292). In our investigation we focus on the current rule of thumb, and conclude that preference for some types of errors over others, correctly applied, should not be understood as biasing anthropotomy (Gk. *anthropos*—human; *ektomia*—to cut out) over anthropomorphism.

While the selective skeptics’ worries about anthropomorphism should be set aside, there is a kernel of truth in their concern about the use of some human terms in the animal context. Addressing that worry offers a productive way for animal cognition research to move ahead by clearly defining the terms used, and that requires being aware of the extent to which we understand such terms when they are used in their more typical human context.

Human properties and the null hypothesis

A human property can also be, unproblematically, an animal property. As animals, humans and non-human animals share a number of biological, morphological, and relational properties. That humans and salamanders both have mass is indisputable, and no special investigation is required to justify the claim that this property is

shared. More interestingly, it is widely accepted that humans and animals can both be attributed some psychological properties such as the ability to *fear* (e.g. a predator) or *desire* (e.g. food). Here too no experimental studies are done to defend these conclusions, even if they are based on behavioral observations of animals—perhaps the same kinds of observations we use to justify the notion that other humans experience fear and desire.

On the other hand, other features are thought by some to be special kinds of human traits, including psychological states such as beliefs, personality traits such as confidence or timidity, emotions such as happiness or grief, social organizational properties such as culture or friendship, and moral behaviors such as cooperation or punishment. Why are these properties more problematic than the others? There has long been an intuition that some properties are higher, and others are lower, and that the lower human properties can be unproblematically attributed to animals while the higher ones are all possible candidates for justifying human uniqueness. The idea of a hierarchy with humans at the top and animals toward the bottom can be traced back to antiquity and The Great Chain of Being. It has proven to be an enduring idea; we see such claims by ancients such as Plato and Aristotle, eighteenth century naturalists such as Linneaus and Darwin, and by the end of the 19th century we see it in psychology, with C. Lloyd Morgan and E.L. Thorndike. And however much people attempt to avoid using the terms “higher” and “lower”, they are still often heard today. The reason to stop using the terms is that, despite all attempts, no satisfactory account of what “higher” and “lower” mean has been offered (see, e.g. Allen-Hermanson 2005; Fitzpatrick 2008, 2009; Sober 1998, 2005; de Waal 1999). And those who have drawn the distinction have disagreed about where the line should be drawn. The Great Chain of Being understands it as relative placement on God’s hierarchy of value. Morgan’s interpretation of Darwinianism led him to believe that reasoning in terms of sense experience was an early evolutionary development, and that reasoning conceptually in terms of general principles was evolutionarily later. His Canon is an epistemic principle that advises that if we can explain a behavior in terms of some evolutionarily earlier cognitive capacity, we should. Thorndike uses “lower” to refer to those animals whose behavior can be accounted for in terms of “a bundle of original and acquired connections between situation and response” whereas human behavior is more appropriately described in terms of consciousness and insight (Thorndike 1911, 4). Shettleworth (2010b) suggests that contemporary psychologists generally use “lower” to refer to cognitive capacities such as associative learning or untrained species-specific behaviors (or what used to be called innate behaviors), and “higher” to refer to cognitive process other than associative learning (such as reasoning, planning, or insight). All these accounts have their problems, even the last two, given recent work suggesting that associative learning may be implicated in supposedly more sophisticated cognitive processes (De Wit and Dickinson 2009; Dickinson 2009; Rescorla 1988). It may be that even insight can be understood as a series of associative processes—at least in some cases (Shettleworth 2010a).

If appeal to higher or lower properties cannot help us to divide human properties into the ones that are problematically and unproblematically attributed to animals, how else can such a division be justified? One answer is that the problematic

properties are those that require a degree of interpretation to identify, those that are still more opaque than transparent. Those human properties that currently defy a robust scientific account are also those that are most often cited as problematically anthropomorphic.

The selective skeptics express concern about the use of certain terms that, when applied to humans, have a rich web of connotations. Joan Silk, for example, has raised concerns about using the term “friend” to describe nonsexual relationships between male and female baboons, because it implies that the baboons share social bonds that are in some way similar to human friendships insofar as they serve the same emotional, psychological, and adaptive functions (Silk 2002). There is concern (Penn 2011) about claims that chimpanzees have a concept of death, and that they may even grieve the death of relatives (as reported by Anderson et al. 2010; Biro et al. 2010). And the debate about whether chimpanzees cooperate, or help one another when there is no immediate reward to one’s self is another area in which some researchers can be interpreted as holding that others are making anthropomorphic claims (see e.g. Greenberg et al. 2010; Jensen et al. 2006; Melis et al. 2006; Silk et al. 2005; Vonk et al. 2008; Yamamoto et al. 2009; Warneken and Tomasello 2006; Melis and Tomasello 2013).

In some cases the debate is about the evidence in a particular study, but often such discussions also give rise to a worry that the property being investigated is *special*, and that additional evidence would be needed to conclude that an animal has that particular property. Rather than claiming that these properties are “higher” human properties, some researchers suggest that the special human properties are folk psychological (Penn 2011; Penn and Povinelli 2007; Povinelli and Giambrone 1999; Povinelli and Vonk 2003, 2004). The “insidious role that introspective intuitions and folk psychology play” in comparative cognition research is identified as being at the heart of the anthropomorphic approach to the science (Penn and Povinelli 2007, 732). The worry appears to be that folk psychological concepts are introspective explanations for human behavior that are then attributed to animals in analogous situations. While such explanations may be “simpler for us” to understand (Heyes 1998, 110), such explanations are not the result of good science. They are problematic attributions in the first place because they are based on possibly false folk account of the cause of *human* behavior. Then, to make matters worse, those same properties are attributed to animals.

We think that identifying the special human properties as folk psychological also fails to do the work that the selective skeptics need it to do; the distinction between folk psychological concepts and scientific psychological concepts will not map onto the distinction between anthropomorphic human properties and shared properties. Consider that folk psychology is “(a) a set of attributive, explanatory, and predictive practices, and (b) a set of notions or concepts used in these practices” (Von Eckardt 1994, 300). The practices of folk psychology include things such as predicting, explaining, justifying, evaluating, and coordinating behavior. And the concepts of folk psychology include theoretical mental entities such as beliefs, desires, intentions, memories, emotions, sensations, and other notions such as goals and personality traits (Andrews 2012). If the selective skeptics were to identify folk psychological terms as anthropomorphic, they would have to accept that application

of any folk psychological term to an animal is impermissible. But the skeptics cannot claim that any use of folk psychological language is problematic, because they make great use of many folk psychological concepts in their scientific papers—concepts including beliefs, memories, goals, desires, and emotions such as fear. As well, Povinelli and Vonk accept that chimpanzees have beliefs; they write “Here the lack of analogy with the behaviorism debate becomes apparent: everyone agrees that the chimpanzee’s mind contains mental representations—that is, intervening variables. The question is: are these intervening variables representations of behavioral abstractions and mental states (as theoretical entities), or behavioral abstractions alone? (Povinelli and Vonk 2003, 158). Povinelli and Vonk are fine with reinterpreting the chimpanzees as having beliefs and desires, and as they see it, the question in the chimpanzee mindreading debate is whether chimpanzees also reinterpret chimpanzee behavior in terms of folk psychological intervening variables. Since the selective skeptics help themselves to some folk psychological concepts in order to do science, they cannot consistently dismiss any use of folk psychology as unscientific. The selective skeptic cannot sustain a general worry about the use of folk psychology in animal cognition research, and so the *special* properties cannot be identified as coextensive with folk psychological properties.

Does the methodology of standard animal cognition research promote the selective skeptic’s approach? Are de Waal and Sober correct in thinking that the methods bias some hypotheses over others? We turn to this question now.

Types of errors

In animal cognition research, like psychology more generally, some kinds of errors are thought to be worse than others. Students of psychology are taught early in their training that committing a Type-I error is worse than committing a Type-II error. These errors are identified in terms of the null hypothesis being investigated.

Since our methodological concerns have to do with how Type-II errors are defined and understood in practice, we will not immediately define the two types of error in any detail. For now, let us understand a Type-I error as a false positive and a Type-II error as a false negative. In the context of animal cognition research, a Type-I error (a false positive) involves ascribing a psychological property to an animal when it lacks that property; this is the error of anthropomorphism. A Type-II error (a false negative), on the other hand, involves denying a psychological state to an animal who actually has that mental state. False positives seem to be associated with permissive and sentimental thinking, whereas Type-II errors, while still errors, are thought to demonstrate a kind of hard-nosed conservatism that has long been taken to be a virtue of the serious scientist. Sober (2005) has suggested that the charge of anthropomorphism is often based on this understanding of the difference between the two kinds of error.

At the same time, Sober thinks that this understanding of the errors has resulted in a different kind of bias in animal cognition research—a bias we call anthropectomy. It has been noted that false positives in animal cognition research have a widely recognized name—anthropomorphism. False negatives on the other

hand do not, despite the best intentions of de Waal and others to coin a phrase [de Waal suggests “anthropodenial” (de Waal 1999) while Sheets-Johnstone calls it “reverse anthropomorphism” (Sheets-Johnstone 1992)]. The skeptics think that anthropomorphism is the worse error. Others think the errors are equally problematic. In his insightful discussion of this point, Sober says of both anthropomorphism and anthropectomy that they are:

...maxims of ‘default reasoning’. They say that some hypotheses should be presumed innocent until proven guilty, while others should be regarded as having precisely the opposite status. Perhaps these default principles deserve to be swept from the field and replaced by a much simpler idea – that we should not indulge in anthropomorphism *or* in anthropodenial [anthropectomy] until we can point to observations that discriminate between these two hypotheses. It is desirable that we avoid the type-1 error of mistaken anthropomorphism, but it is also desirable that we avoid the type-2 error of mistaken anthropodenial [anthropectomy] (Sober 2005, 97).

What Sober suggests is that the methodological position of preferring Type-II errors is the position of preferring anthropectomy over anthropomorphism, and it seems the skeptic would agree with that analysis; where Sober and the skeptic disagree is on the benefit of using Neyman and Pearson’s (1928) method of designing a research program around a null hypothesis and the predicted alternative hypothesis in animal cognition. Indeed, elsewhere Sober criticizes the entire Neyman and Pearson approach to testing hypotheses (Sober 2008).

Although we are sympathetic to Sober’s suggestion that we start with an even playing field, we do not think that it is necessary to do away with Neyman and Pearson testing methods, or to revise the methodological rule of thumb of preferring Type-II over Type-I errors. In order to understand why, we need to get clear on what, exactly, the two kinds of errors are supposed to be and how they *should* be thought of.

The most obvious way to commit a scientific error is to make a false claim; we will call this a *fundamental error*. There are two general ways of making a fundamental error. First, one might claim that something is the case when, in fact, it is not. (Such a claim constitutes a false positive—a Type-I error as defined above.) Second, one might deny that something is the case when, in fact, it is. (Such a denial constitutes a false negative—a Type-II error as defined above.) The two kinds of mistake are on a par. That is, the two kinds of mistake are equally errors; they both constitute a failed attempt to describe the world accurately. If our best science is right, then it is just as much a scientific error to deny that light sometimes behaves as a wave as it is to make the positive claim that light sometimes behaves as a drunken sailor. Falsity does not admit of degrees and all falsehoods in the scientific context are fundamental errors.

If this take on scientific error is even roughly correct, then it is curious that in practice one kind of error is sometimes thought to be *more* of an error than another kind. Why, in the case of animal cognition, do scientists sometimes act as if attributing certain properties to animals is somehow riskier than denying that they have those properties? The claim that an animal lacks a capacity in a world where it

in fact has the capacity is just as false as the claim that an animal has a certain capacity in a world where the capacity is not had. So why is the attitude of anti-anthropomorphism more common in psychology than the attitude of anti-anthropectomy?

In order to answer this question, let us begin by making a few fairly fine-grained distinctions between different types of putative errors:

- 1A Claiming that some Fs are Gs when in fact no Fs are Gs
- 1B Not claiming that no Fs are Gs when in fact no Fs are Gs
- 2A Claiming that no Fs are Gs when in fact some Fs are Gs
- 2B Not claiming that some Fs are Gs when in fact some Fs are Gs

The first thing to note is that 1B and 2B are not errors in any strict sense. 1B and 2B describe attitudes that are consistent with agnosticism with respect to the relationship between Fs and Gs. And while it might be an error, in a loose sense of the term, to reserve judgment in certain cases, reserving judgment can never yield falsity. So 1B and 2B do not constitute fundamental errors. Both 1A and 2A, on the other hand, constitute fundamental errors, and equally so.

In psychology, Type-I and Type-II errors are defined in terms of the null hypothesis in such a way that Type-I errors—of which mistaken anthropomorphism is an example—are fundamental errors, whereas Type-II errors are not. If Type-II errors are considered to be errors of anthropectomy, then it makes sense that there is emphasis on anti-anthropomorphism in animal cognition research. Consider these typical definitions:

Type-I Error Rejecting a null hypothesis when it is in fact true

Type-II Error Failing to reject a null hypothesis when it is in fact false.¹

Here, the null hypothesis is taken to be the default situation; it is what is assumed unless and until investigation shows it to be false, and it can never be proven true. Critically, in the research we are focusing on, it is typical to formulate the null hypothesis in terms of the animal lacking a special human property.

A Type-I Error is equivalent to 1A above. It is a fundamental error. But a Type-II Error is equivalent to 2B above, and 2B is *not* a fundamental error. So the explanation for the overemphasis on anti-anthropomorphism is this: Investigators rightly see that Type-I Errors are more serious—constitute greater errors—than do Type-II Errors. And in the case of animal cognition, a Type-I Error is committed when, for example, one falsely claims that members of a taxa have a theory of mind, cooperate, or grieve. Hence, we see the hesitancy to attribute capacities such as theory of mind to non-human animals. Type-II errors are viewed—again, rightly—

¹ See, for example, Garner (2005): “Type-I error is rejecting a null hypothesis that is true. Type-II (or beta) error is failing to reject a null hypothesis that is, in fact, false” (135). But note that this way of defining Type-II errors is not universal. See Fisher (1971), who defines “errors of the second kind” in terms of “accepting the null hypothesis ‘when it is false’” (17, emphasis added). For reasons that should become clear soon, it is of the utmost importance to determine whether Type-II errors should be defined in terms of “failing to reject” or in terms of “accepting”, for these phrases describe two entirely different doxastic states. Our assessment of the problem with much animal cognition research is that it is unclear whether this important distinction is made in actual practice by researchers.

as less serious than Type-I errors, and so investigators are inclined to not reject the null hypothesis—to not reject the claim that animals lack a theory of mind, for example—even though a failure to reject it runs the risk of failing to say something true.

The problem with this general approach is that when investigators go out of their way to avoid Type-I Errors, they not only run the risk of committing Type-II errors, but they also run the risk of committing the much more serious 2A error. Anthropectomy involves a claim about the nonexistence of a property. It is not a position of agnosticism, and so it is a mistake to prefer anthropectomy to anthropomorphism. Of course careful researchers can avoid this slide from Type-II error to 2A errors, but as we discuss in the next section, in practice this is indeed a problem.

To put the same point in a slightly different way, we think it is a mistake—a serious mistake having to do with the application of psychological methodology—to both:

- a. hold that Type-I errors are more serious than Type-II errors
and
- b. view Type-I errors as errors of anthropomorphism and Type-II errors as errors of anthropectomy.

To see why, consider again how we might define the two kinds of error:

Type-I error = rejection of a null hypothesis when it is in fact true.

Understood this way, a Type-I error is indeed a fundamental error to be avoided. So far, so good. The problem for the skeptic can be put in terms of a dilemma concerning the definition of Type-II errors. Either:

(Horn 1) Type-II error = failure to reject a null hypothesis when it is in fact false.

or

(Horn 2) Type-II error = acceptance of a null hypothesis when it is in fact false.

If the selective skeptic accepts (Horn 1) of the dilemma, then he has some explaining to do. First, as defined in (Horn 1), Type-II errors are not fundamental errors, while Type-I errors *do* constitute fundamental errors. Note that *not* committing a Type-II error under this definition is not very epistemically virtuous. A rock commits a Type-II error under this definition, and so do a great many people not involved in any kind of scientific research and who have no opinion at all concerning the field of study: there they sit, failing to reject the investigator's false null hypothesis. The selective skeptic cannot deny that animals have some special human property, if he accepts (Horn 1). That is, the method cannot be used to justify claims of human uniqueness. Rather, the skeptic must adopt agnosticism, since it is here that the methodological prohibition against affirming the null hypothesis comes into play.

This leads to a second concern with (Horn 1), namely that the skeptic needs to explain why we should take anything like “animals do not have psychological properties” to be the null hypothesis. Surely to *begin* the investigation with the skeptical view as the default position is to bias the investigation. In other words, the onus is on the skeptic to explain why the skeptical hypothesis and not the optimistic

hypothesis that animals *do* have psychological properties is the proper null hypothesis.

There are often cases where pragmatic or moral concerns might justify counting a particular hypothesis as the null hypothesis. For example, in drug trials there is good reason to use “drug x is ineffective” as the null hypothesis, given the health risks associated with putting any new drug on the market. In criminal jurisprudence, the hypothesis that the accused is “innocent until proven guilty” serves as a reasonable null hypothesis since we have decided that sending innocent people to prison is morally worse than letting guilty people go free. But when our concerns are purely epistemic, as they presumably are in the case of animal cognition, it’s less clear why either the skeptical or optimistic hypothesis should get preferential treatment from the outset. Unless there is some *prima facie*, pre-empirical reason to think that one of the hypotheses is more plausible, or there is some independent empirical evidence that the skeptical hypothesis is statistically more common, neither should be counted as a null hypothesis. To insist that one must be counted as the null hypothesis is to beg the question against the other hypothesis.

One might object that we have presented a mere caricature of null-hypothesis selection in animal cognition research. Surely, the objection goes, there can be a variety of good grounds for choosing this or that hypothesis as the null hypothesis and the reasons researchers have for choosing this or that hypothesis as the null are to be evaluated on their merits. Haven’t we unfairly painted with too-broad strokes here?

Our response is to emphasize that we are not even trying to raise an objection for the use of null hypotheses or Type-I and Type-II error analysis in general. Rather, the objection is that *how* these tools are *employed* when considering whether or not an animal has some psychological property is often unwarranted. Consider, for example, some possible reasons for counting a skeptical hypothesis as the null:

1. There is prior anecdotal evidence that fails to report evidence of the property in question in the species in question.
2. There are statistically-based reasons to think the species in question does not have the property in question because closely related species seem to lack the property.
3. There are anecdotally-based reasons to think the species in question does not have the property in question because closely related species seem to lack the property.
4. There are independent theoretical grounds for thinking that the species in question lacks the property in question.²

We maintain that all of these are potentially good reasons for treating a skeptical hypothesis as the null hypothesis. The question is just whether these reasons are usually or often had when the skeptical hypothesis is chosen as the null hypothesis. It is perhaps worth repeating the following point: Unless there is some *prima facie*, pre-empirical reason to think that one of the hypotheses is more plausible, or there is some independent empirical evidence that the skeptical hypothesis is statistically

² Thanks to Richard Moore for providing this list of reasons one might take a hypothesis as the null.

more common, neither should be counted as a null hypothesis. In the list above, (1) through (3) fall under the provision that there be some “independent empirical evidence that the skeptical hypothesis is statistically more common”. Item (4) falls under the provision that there be some “prima facie, pre-empirical reason” to think that the skeptical hypothesis is more plausible.

It is worth noting, with respect to (4), that one has to tread carefully. First, in order to use “theoretical grounds” as a rationale for choosing the skeptical hypothesis as the null hypothesis, it is important that researchers ensure that those grounds really are independent of the question at issue. For example, it will hardly do for a researcher to try to justify her use of “the subject chimpanzees will not exhibit a theory of mind” as a null hypothesis on the grounds that that’s what her favorite theory says or directly entails. Here, the proposed null hypothesis and the theoretical grounds on which it is based are “too close”; they in effect say the same thing. So the independence of the theoretical grounds are of the utmost importance. It can be difficult to determine when a theory is independent enough for it to be used as grounds for deciding a null hypothesis, and there might not be any sharp distinction between the independent-enough and the not-independent-enough. But suffice it to say that animal cognition researchers must pay attention to the concern of independence.

Another difficulty in making use of (4) is that because theories are always underdetermined by existing evidence, it is not clear which “theoretical grounds” are the correct grounds to use in determining what the null should be. This difficulty is not just a philosopher’s problem. Especially in the field of animal cognition research, evidence can underdetermine theory in a fairly straightforward, obvious way that is of practical significance to anyone working in the field. For every well-respected and widely-held theory that suggests species X lacks characteristic ϕ , there is another, equally well-respected and widely-held theory that suggests species X has characteristic ϕ . If this characterization is even remotely accurate, then it will be very difficult for a researcher to justify using one set of pre-empirical grounds over another set for choosing her null hypothesis.

In short, if any of (1) through (4) are met, then great. Researchers can then reasonably choose one null over another. Our worry is just that the skeptical hypothesis is often chosen as the null when it is very questionable whether any of (1) through (4) are actually met.

At this point it is worth stressing that our concern in this paper is with the correct methodology for a subset of animal cognition research that deals with hypotheses about human uniqueness, not empirical research in general, or about animal behavior, or even other areas of animal cognition (such as serial memory). We do not want to overstate our case. Even if one’s goals are purely epistemic, there might be reason to choose a particular claim as a null hypothesis because prior evidence suggests it is the norm. If, for example, statistics suggest a particular rate of pregnancy among sexually active women of a certain age range, it would seem quite reasonable to use that rate of pregnancy as the null hypothesis when investigating the effectiveness of a new birth control method. But there is no such statistical norm regarding animals’ possession of special psychological properties, nor is there any unproblematic theoretical ground. Or, at least, no such evidence has been provided

in such a way that would warrant a *general* null hypothesis to the effect that animals lack this or that cognitive capacity.

In a textbook on statistics that seems to be fairly representative, Garner (2005) defines Type-II errors as in (Horn 1) and offers the following “tip” for “setting up a good null hypothesis”:

The null hypothesis usually specifies a dull or disappointing out-come. The null hypothesis means that we cannot announce exciting research results, take action, or establish a new finding. If the null hypothesis is true, there is no relationship between two variables; a supplier has not cheated us; a new curriculum does not improve reading scores – in short, we get a boring result. Therefore, the null hypothesis is a statement we would like to be able to reject (140).

Before saying something about Garner’s advice, we should again emphasize that we are not here arguing against her way of understanding the null hypothesis in general. Our argument is centered on null hypotheses and Type-1 and Type-2 errors as they relate to animal cognition research, and in particular, on questions about whether animals have some special human property. Perhaps Garner’s advice is very good advice in other areas of research. Still, we think this passage is illuminating, but not in the way Garner intends. Whether a result is “dull”, “disappointing”, or “boring” is something we would have thought is in the eye of the beholder. Moreover, it is not at all clear what it means to “like to be able to reject” a hypothesis when one’s concerns are purely epistemic. What does “like” mean here? Most of us would like to be able to reject the hypothesis that there is a secret government plot to allow aliens to harvest our bodies for food. But such a hypothesis is certainly not boring. Perhaps the point is just that the null hypothesis is one we would like to be able to reject if we wanted the world to be a very exciting place. But the fact remains that one’s epistemic concerns should be immune to potential levels of excitement.

When an investigation has to do with animal cognition and the like, it is especially difficult to see how or why Garner’s advice should be followed. It is true that some people would find it amazing if mere beasts shared our capacity to read others’ minds, for example. But others would find it equally amazing if nonhuman animals, especially those who are very closely genetically related to humans, lacked *all* of our cognitive capacities, e.g. our ability to appreciate that others see material objects. So, again, even if the kind of advice Garner offers is on the right track generally, it is very difficult to see how it could be on the right track when it comes to the kind of animal cognition research we are interested in here. For even if we were to try to follow her advice, we would be at a loss in determining whether the skeptical or optimistic hypothesis should be the null hypothesis.

While the selective skeptic could choose an optimistic hypothesis as the null hypothesis, and then risk a fundamental Type I error in concluding that animals lack the property attributed to them, the empirical work required to do so would be much more demanding, given the methodological rule of thumb to avoid Type 1 errors. We conclude that the selective skeptic cannot accept (Horn 1) of the dilemma, because it does not allow her to claim that animals lack a certain property, and

because the evidential requirements for defending a negative existential claim have not been met.

These problems associated with accepting (Horn 1) arise because the definition of Type-II errors in (Horn 1) makes it so that there is an asymmetry between Type-I and Type-II errors. Type-I errors are fundamental errors, but under the (Horn 1) definition, Type-II errors are not. Consequently, if the skeptic accepts (Horn 1), then she has good reason to think that Type-II errors are less serious than Type-I errors, but she has no reason to think that the skeptical hypothesis about psychological properties of animals should count as the null hypothesis.

If the skeptic remedies this situation by accepting the (Horn 2) definition of Type-II errors, then the problem is just that she has no reason whatever to think that Type-I errors are more egregious than Type-II errors. Under the (Horn 2) definition, both types of error are fundamental errors. What reason could there possibly be to prefer the acceptance of one falsehood over the acceptance of another falsehood? Again, if one's goals are primarily epistemic, and not pragmatic, prudential, or moral, then there can be no satisfactory answer to this question (barring the existence of known statistical norms or appropriately distanced theoretical considerations). So, under the (Horn 2) definition, it is perfectly legitimate for a researcher to choose something like "animals do not have psychological properties" as her null hypothesis, but it is not legitimate for her to think that accepting this hypothesis when it is in fact false is worse than rejecting it when it is in fact true.

Now, the same researcher might choose "animals *do* have psychological properties" as her null hypothesis. The lesson just is that once Type-I and Type-II errors are made symmetrical, neither error is worse than the other. And since the errors are defined in terms of the null hypothesis, it follows that there is no direct epistemic reason to choose the skeptical hypothesis as the null hypothesis and no reason to choose the optimistic hypothesis as the null hypothesis. It's a wash. If, for purposes of statistical analysis, it is useful to deem a particular hypothesis the null hypothesis, then fine. But the "deeming" must be of no methodological consequence—it must not bias one hypothesis over the other unless there are pre-experimental reasons for doing so. If the methods of statistical analysis themselves treat one of the hypotheses as having special epistemic status in the absence of previously established statistical norms, then so much the worse for those statistical methods. For, as we have shown, there is no *epistemic* reason to treat either the skeptical or the optimistic hypothesis in animal cognition preferentially, and if (Horn 2) is accepted, then there is no prohibition against affirming the null.

To summarize: If the skeptic takes on (Horn 1) of the dilemma, then she has good reason to think that Type-II errors are less serious than Type-I errors, but she is unable to defend the claim that animals lack a particular property. If, on the other hand, the skeptic accepts (Horn 2), then she is free to deem the skeptical hypothesis the null hypothesis, but then she has no reason to think that Type-II errors are less serious than Type-I errors. Either way, the claim that the risk of anthropectomy is less troubling than the risk of anthropomorphism is unwarranted on either interpretation of errors. The Neyman and Pearson hypothesis testing method itself doesn't bias researchers against accepting that other animals share psychological characteristics with humans.

To illustrate, let's take an example. Consider Michael Tomasello's claim that apes don't share joint attentional frames because they don't point to one another in the wild. The null hypothesis is that apes don't have joint attention, and we can operationalize joint attention as pointing (setting aside for simplicity's sake the fact that other behaviors, such as eye movements, could also count as evidence of joint attention). Tomasello's claim that apes and humans differ in this respect may be seen as an assertion of the skeptical null hypothesis, not an agnostic position, and as such we would situate him on Horn 2 of the dilemma. He would not be able to assert the negation of the null were he to accept Horn 1. Thus, if he is wrong he is committing a fundamental error. Now consider Juan-Carlos Gómez, who, using Treverathan's methods from developmental psychology, claims that apes do engage in joint attention. If we take Gómez as starting with the skeptical null hypothesis, we can interpret Gómez as denying the null and making the positive claim, thus risking a false positive (a fundamental error), but in so doing he is able to accept either horn of the dilemma. Here both Tomasello and Gómez are taking the same risk in making their claims—something we think they are both fully aware of.

Now, if, for example, Tomasello took as the null hypothesis that there is no difference in the ability to share joint attentional frames between chimpanzees and humans, then he would be able to accept Horn 1 of the dilemma *and* deny that chimpanzees have joint attention. Here again, if he were wrong, he'd be making a fundamental error. But consider the evidence needed for denying the null hypothesis that humans and other apes share joint attention. The researcher would have to search everywhere before being able to assert the nonexistence of a trait, in order to make sure that the experimental design or implementation isn't the issue, and that the particular participants are not somehow unmotivated, too young, mentally ill, etc. Empirical justification of a lack of some property requires a thorough search, and the conclusion can only be tentative. Negative existential claims are justified in the realm of logic, not so much in empirical science. To defend the hypothesis that there is a difference between the two species, then, it would be empirically simpler to use the skeptical hypothesis as the null and take Horn 2 of the dilemma. But if that were the case, there would be no reason to prefer Type II errors over Type I.

Conclusion

Like de Waal and Sober, we think that there is an unjustified worry about attributing special human properties to animals. However, we think that on no interpretation of the errors does the Neyman and Pearson hypothesis testing method lead to the conclusion that anthropectomy is preferable to anthropomorphism.

The selective skeptics may reply to our argument that anthropectomy is as bad an error as anthropomorphism by reaffirming their status as skeptics rather than slayers—they may claim that they do not deny the existence of special human properties in animals, and hence are not open to the charge of anthropectomy, but that they simply remain agnostic. However, that response isn't consistent with the sorts of claims the selective skeptics make. Take, for example, this claim: "...whatever "good trick" (Dennett 1996) was responsible for the advent of human

beings' ability to reinterpret the world in a symbolic-relational fashion, it evolved in only one lineage—ours. Nonhuman animals didn't (and still don't) get it" (Penn et al. 2008, 129). This isn't a cherry-picked example, and such negative existential claims abound in selective skeptics' writing. It shouldn't be surprising that such claims are made, when selective skeptics such as Penn and Povinelli identify the division between themselves and other animal cognition researchers as a division between those who emphasize the differences and those who emphasize the similarities between humans and animals. The selective skeptics are unwarranted in making anthropocentric claims based on the standards of the Neyman and Pearson hypothesis testing method, and by making such claims they do a disservice to the science of animal cognition research.

Before closing, we should note that not all selective skeptics make this error. In revisiting the worry about folk psychology in animal cognition research, we can interpret the concern as offering a helpful reminder. While no general prohibition against using folk psychology in animal cognition can be justified in the face of current practice, given the ubiquity of such terms in scientific psychology, the worry about folk psychology does point to the need to have well-defined terms in hypotheses as well as in interpretations of results. We must try to avoid using fuzzy language to describe animal behavior or cognition, especially when the functions and the mechanisms of such behavior or cognitive capacities are not well understood in humans.

This is, in fact, the lesson that Morgan wanted us to learn. Morgan thought that mentality could only be interpreted, in humans and in nonhumans, and that we humans tend to see our own behavior as more clever than it really is. For Morgan, there was no question that there was an animal psychology to be studied, but he cautioned us not to over-intellectualize human cognition. That is the first step in doing good animal cognition: "To interpret animal behavior one must learn also to see one's own mentality at levels of development much lower than one's top-level of reflective self-consciousness. It is not easy, and savors somewhat of paradox." (Morgan 1930, 250).

It is true that when we use the same term to describe baboon friendship and human friendship, and the term has not been operationalized in the same way in its application to humans and nonhumans, its careless use may have unintended implications. Researchers can avoid unintended implications by carefully choosing the terms they use to interpret animal behavior, and by reminding us that some terms, like 'friendship' refer to a range of human relationships that differ from one another in innumerable ways (across age ranges, across cultures, etc.). 'Friendship' may be an umbrella term, or a family resemblance term, and when such a word is used to describe animal relationships, researchers can remind their audience—especially popular audiences—that there is no straightforward analogy between a single human relationship and an animal friendship.

This is especially important when animal cognition researchers speak about their research to the popular press. For example, a scientific study showing evidence that orangutans use iconic or pantomimic gestures to communicate their desires (Russon and Andrews 2011) was reported by many in the popular press (including *National Geographic*) as suggesting that orangutans play charades. Reports about animal

behavior that make them seem more human-like appeal to our biases about animals, and it also appeals to the human imagination. The results are more striking when reported in such a way, but such reports also do a disservice to science.

This leads us to conclude that researches should set aside any worry about *special* human psychological, social, or normative properties, given the difficulty in even identifying what such properties might be. Rather, animal cognition researchers who want to make comparisons across species should carefully identify the property of interest in the comparison species before they begin to ask whether it exists in the target species. Some properties, such as the capacity for theory of mind, are still so poorly understood in the human case that it isn't surprising that looking for them in animals has led to so much controversy. The better defined the question, the better the science. While that is a general principle that extends beyond animal cognition research, it is one that bears repeating in this context.

Acknowledgments Thanks to helpful comments from the audience at the Southern Society for Philosophy of Psychology, Kyoto University, and comments on the draft from Irina Meketa and Richard Moore. We also are very grateful for help with the Greek from David Curry and Daniel Devereux.

References

- Allen-Hermanson S (2005) Morgan's canon revisited. *Philos Sci* 72(4):608–631
- Anderson J, Gillies A, Lock L (2010) *Pan* thanatology. *Curr Biol* 20(8):R349–R351
- Andrews K (2012) Do apes read minds? Toward a new folk psychology. MIT Press, Cambridge
- Bekoff M, Allen C (1997) Cognitive ethology: Slayers, skeptics, and proponents. In: Mitchell RW, Thompson NS, Miles HL (eds) *Anthropomorphism, anecdotes, and animals*. State University of New York Press, Albany, pp 313–334
- Biro D, Humle T, Koops K, Sousa C, Hayashi M, Matsuzawa T (2010) Chimpanzee mothers at Bossou, Guinea carry the mummified remains of their dead infants. *Curr Biol* 20(8):R351–R352
- Blumberg M, Wasserman E (1995) Animal mind and the argument from design. *Am Psychol* 50(3):133–144
- Carruthers P (2000) *Phenomenal consciousness: a naturalistic theory*. Cambridge University Press, Cambridge
- Carruthers P (2004) On being simple minded. *Am Philos Quart* 41:205–220
- Datson L, Mitman G (eds) (2005) *Thinking with animals: new perspectives on anthropomorphism*. Columbia University Press, New York
- Davidson D (1975) Thought and talk. In: Guttenplan S (ed) *Mind and language*. Oxford University Press, Oxford, pp 7–24
- Davidson D (1982) Rational animals. *Dialectica* 36(4):317–327
- Dennett DC (1996) *Kinds of minds: toward an understanding of consciousness*. Basic Books
- de Waal FBM (1999) Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. *Philos Top* 27:225–280
- de Wit D, Dickinson A (2009) Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychol Res* 73(4):463–477
- Dickinson A (2009) What are association formation models. *Learn Behav* 37:21–24
- Fisher R (1971) *The design of experiments*. Hafner Press, New York
- Fisher JA (1990) The myth of anthropomorphism. In: Bekoff M, Jamieson D (eds) *Interpretation and explanation in the study of animal behavior: Vol. 1, Interpretation, Intentionality, and Communication*. Westview Press, Boulder, pp 96–116
- Fisher JA (1991) Disambiguating anthropomorphism: an interdisciplinary review. In: Bateson PPG, Klopfer PH (eds) *Perspectives in ethology*, vol 9. Plenum, New York, pp 49–85
- Fitzpatrick S (2008) Doing away with Morgan's Canon. *Mind Lang* 23(2):224–246

- Fitzpatrick S (2009) The primate mindreading controversy: a case study in simplicity and methodology in animal psychology. In: Lurz R (ed) *The philosophy of animal minds*. Cambridge University Press, New York
- Garner R (2005) *The joy of stats*. Broadview Press, Peterborough
- Greenberg J, Hamann K, Warneken F, Tomasello M (2010) Chimpanzee helping in collaborative and noncollaborative contexts. *Anim Behav* 80:873–880
- Heyes C (1998) Theory of mind in nonhuman primates. *Behav Brain Sci* 21(1):101–134
- Jensen K, Hare B, Call J, Tomasello M (2006) What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proc R Soc B Biol Sci* 273:1013–1021
- Keeley B (2004) Anthropomorphism, primatomorphism, mammalomorphism: understanding cross-species comparisons. *Philos Biol* 19(4):521–540
- Kennedy JS (1992) *The new anthropomorphism*. Cambridge University Press, New York
- McGrew W (2004) *The cultured chimpanzee: reflections on cultural primatology*. Cambridge University Press, Cambridge
- Melis A, Tomasello M (2013) Chimpanzees' (*Pan Troglodytes*) strategic helping in a collaborative task. *Biol Lett* 9:1–4
- Melis AP, Hare B, Tomasello M (2006) Chimpanzees recruit the best collaborators. *Science* 311:1297–1300
- Mitchell R, Thompson N, Lynn Miles H (1997) *Anthropomorphism, anecdotes, and animals*. State University of New York Press, Albany
- Morgan C (1903) *An introduction to comparative psychology*. Walter Scott Pub. Co, London
- Morgan C (1930) *Autobiography of C. Lloyd Morgan*. In: Murchison C (ed) *History of psychology in autobiography*, vol. 2. Clark University Press, Worcester, pp 237–264
- Neyman J, Pearson ES (1928) On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. *Biometrika* 20A(1/2):175
- Penn D (2011) How folk psychology ruined comparative psychology: and how scrub jays can save it. In: Menzel R, Fischer J (eds) *Animal thinking: contemporary issues in comparative cognition*. MIT Press, Cambridge, pp 253–266
- Penn D, Povinelli D (2007) On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philos Trans R Soc B* 362:731–744
- Penn D, Holyoak K, Povinelli D (2008) Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav Brain Sci* 31:109–178
- Povinelli D, Bering J (2002) The mentality of apes revisited. *Curr Dir Psychol Sci* 11(4):115–119
- Povinelli D, Giambrone S (1999) Inferring other minds: failure of the argument by analogy. *Philos Top* 27:167–201
- Povinelli DJ, Vonk J (2003) Chimpanzee minds: Suspiciously human? [opinion]. *Trend Cogn Sci* 7(4):157–160
- Povinelli D, Vonk J (2004) We don't need a microscope to explore the chimpanzee's mind. *Mind Lang* 19(1):1–28
- Povinelli D, Bering J, Giambrone S (2000) Toward a science of other minds: escaping the argument by analogy. *Cogn Sci* 24(3):509–541
- Rescorla RA (1988) Pavlovian conditioning: it's not what you think it is. *J Exp Psychol Anim Behav Process* 43:151–160
- Russon A, Andrews K (2011) Orangutan pantomime: elaborating the message. *Biol Lett* 7(1): 627–30. doi:10.1098/rsbl.2010.0564
- Sheets-Johnstone M (1992) Taking evolution seriously. *Am Philos Quart* 29(4):343–352
- Shettleworth SJ (2010a) Clever animals and killjoy explanations in comparative psychology. *Trend Cogn Sci* 14(11):477–481
- Shettleworth S (2010b) *Cognition, communication, and behavior*, 2nd edn. Oxford, New York
- Silk J (2002) Using the 'F'-word in primatology. *Behaviour* 139(2–3):421–446
- Silk J, Brosnan S, Vonk J, Henrich J, Povinelli D, Richardson A, Lambeth S, Mascaro J, Schapiro S (2005) Chimpanzees are indifferent to the welfare of unrelated group members. *Nature* 437(7063):1357–1359
- Sober E (1998) Morgan's canon. In: Cummins D, Allen C (eds) *The evolution of mind*. Oxford University Press, New York, pp 224–242
- Sober E (2005) Comparative psychology meets evolutionary biology: Morgan's canon and cladistic parsimony. In: Mitman G, Datson L (eds) *Thinking with animals: new perspectives on anthropomorphism*. Columbia University Press, New York, pp 85–99

- Sober E (2008) Evidence and evolution: the logic behind the science. Cambridge University Press, Cambridge
- Stich S (1979) Do animals have beliefs? *Australas J Philos* 57(1):15–28
- Thorndike EL (1911) Animal intelligence. Macmillan, New York
- Von Eckardt B (1994) Folk psychology. In: Guttenplan S (ed) A companion to the philosophy of mind. Blackwell, Cambridge, pp 300–307
- Vonk J, Brosnan SF, Silk JB, Henrich J, Richardson AS, Lambeth S, Schapiro J, Povinelli DJ (2008) Chimpanzees do not take advantage of very low cost opportunities to deliver food to unrelated group members. *Anim Behav* 75:1757–1770
- Warneken F, Tomasello M (2006) Altruistic helping in infants and young chimpanzees. *Science* 311(5765):1301–1303
- Wynn C (2004) The perils of anthropomorphsim. *Nature* 428(606):606
- Wynn C (2007) What are animals? Why anthropomorphism is still not a scientific approach to behavior. *Comp Cogn Behav Rev* 2:125–135
- Yamamoto S, Humle T, Tanaka M (2009) Chimpanzees help each other upon request. *PLoS ONE* 4(10):e7416