Anthropomorphism as Cognitive Bias
Mike Dacey

Philosophers and psychologists have long worried that a human tendency to anthropomorphize leads us to err in our understanding of nonhuman minds. This tendency, which I call *intuitive anthropomorphism*, is a heuristic used by our unconscious folk psychology to understand the behavior of nonhuman animals. I argue that the dominant understanding of intuitive anthropomorphism underestimates its complexity. It does often lead us to err, but not always. And the errors it produces are not only overestimations of nonhuman intelligence. If we want to understand and control intuitive anthropomorphism, we must treat is as a cognitive bias, and look to the empirical evidence. The literature on controlling implicit social biases is particularly helpful. That literature suggests that the most common for intuitive anthropomorphism, Morgan's Canon, should be rejected, while others need supplementation. It also suggests new approaches.

1. Introduction

Humans naturally anthropomorphize. As David Hume put it: "There is an universal tendency among mankind to conceive all beings like themselves . . . We find faces in the moon, armies in the clouds" (Hume 1957, pg. 29). Philosophers and psychologists attempting to understand the minds of nonhuman animals have long worried that this tendency leads us to error. This worry is shared across fields and across theoretical attitudes about anthropomorphism more generally. Kennedy (1992) argues forcefully against any form of anthropomorphism, and sees this tendency, which is "simply *built into us*" (pg. 28, emphasis his), as the reason it is so problematic. Rivas & Burghardt (2002) believe that some forms of anthropomorphism are valuable, but caution against its naïve forms: "Anthropomorphism is like Satan in the bible - it comes in many guises and can catch you unawares!" (pg. 15). I call the human tendency to anthropomorphise *intuitive anthropomorphism*, and it is the specific target of this paper. Existing views get intuitive anthropomorphism wrong, and as a result, fail to control its effect on the sciences of nonhuman minds.

To start, the term 'anthropomorphism' (*simpliciter*) needs clarification. In its strictest sense, anthropomorphism is sometimes defined as a *kind of error*: overestimating the intelligence of animals by attributing to them human-like traits they do not have (perhaps consciousness or belief-desire psychology). A broader sense of the term applies to the belief that a nonhuman animal possesses any of these 'characteristically human' traits, while allowing that that belief may be true or false. A still broader sense treats anthropomorphism as a way of thinking, or a *process* of forming beliefs about non-human minds: coming to understand the minds of nonhuman animals by analogy to our own, while leaving open whatever beliefs we might form. I treat anthropomorphism as a process. Though anthropomorphism-as-an-error is probably most often made explicit, usage of the term tends to slide between these, especially in debates over the role of anthropomorphism in science (more in section 6). I view anthropomorphism as a process because it allows more productive engagement with the crucial problem.[1] I say this for three reasons.

Firstly, this use allows me to remain agnostic about anthropomorphism generally. There is a large and contentious debate about whether there are legitimately scientific anthropomorphic strategies (e.g. Burghardt 1985, Rivas & Burghardt 2002, de Waal 1991, 1999, Mitchell 2005, Wynne 2007). I side-step this debate to focus specifically on intuitive anthropomorphism, which is just one kind of anthropomorphism.

Secondly, it does not prematurely restrict anthropomorphism to certain posits, better capturing the nature of intuitive anthropomorphism. Intuitive anthropomorphism is a heuristic

[1] There are two other reasons I won't argue for at length. First, treating anthropomorphism as a process is anthropocentric in a much less pernicious way than the other definitions. They mark off a class of 'characteristically human' traits, but why think we have special claim to them? But, we are human, so it is natural that we interpret other species from our own perspective. Second, this view shows how one can be concerned about intuitive anthropomorphism without thinking that animals are unintelligent, or that they don't have minds. This is not what is at issue here.

employed by our unconscious folk psychology. Our unconscious interprets behavior of nonhuman animals in the same way it interprets human behavior. It is an empirical question what effects this may have. Like any cognitive heuristic, it likely leads to errors in many cases, but it does not always do so, and we can't know how or when it does until we've tested it. It is a mistake to assume what effects it has at the outset. In fact, intuitive anthropomorphic error is *not* merely a matter of overestimating intelligence by positing a set of specific mental states, so controls that simply aim at correcting this mistake are ineffective. The effects of intuitive anthropomorphism are complex, and the errors it produces share nothing besides their common source. Focusing on any particular kind of posit or error before we understand intuitive anthropomorphism dooms us from the start.

Thirdly, this way of looking at anthropomorphism opens up a new approach to debates about it. Too often these debates proceed as follows: one theorist claims that common explanations of some behavior (read: not their own preferred theory) are either overly anthropomorphic or overly averse to anthropomorphism, while another claims the opposite. These are not really arguments about anthropomorphism itself; they are arguments about *those theories* of animal cognition. Very little in comparative psychology is settled, so it should not be surprising that these debates make little headway. This approach makes sense if anthropomorphism is a kind of error: to identify instances of anthropomorphism in the field, we must first identify errors in the field. If anthropomorphism is a process, there is another approach available at the level of individual psychology. So I do not argue that the field is overly anthropomorphic, I argue that any particular judgement one makes about the psychology of nonhuman animals might be subject to the influence of intuitive anthropomorphism. As such, any particular judgement is potentially subject

to an intuitive anthropomorphic bias.[2] To understand what this means, I look to the empirical literature.

Intuitive anthropomorphic bias is one kind of cognitive bias that results from reasoning heuristics applied in our unconscious (e.g. Tversky & Kahneman 1974). The literature on cognitive biases and unconscious processing generally will help understand intuitive anthropomorphism (section 2). The literature on implicit social biases will help understand how to control it (section 3). This is the largest literature on controlling unconscious biases, and it targets interventions of the right form for the current discussion. Collectively, this literature suggests that existing strategies for controlling intuitive anthropomorphism are ineffective (sections 4-6), and can help develop new controls (section 7).

Figure 1.

[2] This might appear to *imply* that the field is overly anthropomorphic. But I argue that the effects of intuitive anthropomorphism are poorly understood, so it is unclear what it means to say that the field is overly anthropomorphic in the first place.

2. The Intuitive Anthropomorphic Bias

Figure 1 shows Ham the Chimpanzee in 1961, on his way to the capsule that will launch him into space. He is about to become the first hominid to orbit the earth. The look on his face appears to be one of excitement or pride (internet comments on the photo show how common this interpretation is). Perhaps he perceives excitement in the behavior of those around him, or perhaps he is just pleased by the attention. Unfortunately, this is unlikely. In chimpanzees, this expression is known as the 'fear grin.' As much as we can rationalize the thought that Ham is pleased, fear is a more likely reactions to being strapped into a seat and carried to a strange enclosure.

This is a case where intuitive anthropomorphism leads us astray. We see an animal grinning and it *looks to us* like happiness. That is a pretty good heuristic for dealing with other humans, but we go wrong because our unconscious folk psychology applies it the same way in nonhumans. Even knowing that we are wrong, the perceptual Gestalt remains: Ham still looks happy. Note that we are not positing human-like capacities in Ham that he incapable of having. There is no substantive difference in the intelligence required for fear and happiness.[3] Intuitive anthropomorphism does lead us to err, and those errors do not just overestimate intelligence (though sometimes they do). We also should not think that intuitive anthropomorphism *always* leads us to error (anthropomorphism is a way of thinking, not a kind of error), but there is good reason to think it will do so often.

Daily experience with pets and animal cartoons demonstrates how easy it is: we talk to our dogs, and do not blink when cartoon dogs talk to their owners. There are also evolutionary

[3] I am speaking loosely of 'happiness' and 'fear,' but I do not mean the *human* mental states; I mean whatever is the chimpanzee analogue.

reasons to think we *should* anthropomorphize: Intuitive anthropomorphism is the kind of fast and frugal heuristic that can work for evolutionary purposes, even if it is not up to the epistemic standards of science. As some have argued (e.g. Caporeal & Heyes 1997, Gallup, Marion & Eddy 1997), our intuitive folk psychology most likely evolved to inform social interactions with other humans, and then was exapted to handle interactions with nonhuman animals. The intractable problems of interpreting other minds are a prime target for 'good enough' predictive strategies (Dennet 1989 argues in a similar spirit).

There is also considerable empirical evidence that humans make errors because of intuitive anthropomorphism. In one of the classic studies of psychology, Heider & Simmel (1944) showed that participants attributed intentional actions to a collection of two dimensional shapes 'interacting' in a short cartoon. This indicates a tendency to see (quite literally) intentional action when certain cues are present. In this experiment, the cues are irregular movements that seeming respond to one another. This, by itself, is insufficient evidence to really conclude that behavior is intentional, but it seems to be enough for our unconscious folk psychology.

In general, our anthropomorphic unconscious folk psychology seems to trigger on simple or irrelevant cues, suggesting it triggers too often. Other simple cues like hands, eyes, and faces influence the attribution of conscious mental states (e.g. Arico et. al. 2011, Fiala, Arico, & Nichols 2011). Magnifying this, humans are wired to err in the direction of seeing faces that aren't there over missing faces that are (e.g. Liu et. al. 2014). Finally, humans are biased in attributing conscious mental states to animals that move at about the same speed as us (Morewedge, Preston, & Wegner 2007).

Children, of course, anthropomorphise wildly (e.g. Gebhard, Nevers, & Billmann-Mahecha 2003). And adults anthropomorphize when explaining and imagining behavior. Religious

participants asked to describe God or tell stories involving God will anthropomorphize God, even when doing so contradicts their theological beliefs (e.g. these are anthropomorphic errors by the subjects' own lights; Barret & Keil 1996). Participants presented with written descriptions of situations and asked to assess the 'reasonableness' of various mental state attributions ascribe mental states to dogs that are of the same kind as they do a human child (though quantitatively simpler; Rasmussen and Rajecki 1995).

One might wonder whether these effects apply to scientists. While scientists may not make the flagrant errors that children make, or that experimental subjects make in snap judgments, we should expect that they do make some errors. In general, unconscious social biases influence behavior even during careful deliberation by experts. This has been shown in hiring decisions (Bertrand & Mullainathan 2003), including hiring by scientists (Moss-Racusin et al. 2012), medical decision making (Green et. al. 2007), the judicial process (Banks, Eberhardt, & Ross 2006, Mitchell et. al. 2005, Rachlinski et. al. 2009). Expertise and deliberation are not themselves sufficient to stop cognitive biases. In fact, the self-perception that one is objective *increases* social bias (Uhleman & Cohen 2007).

So scientists and philosophers should not expect themselves to be immune. In addition, scientific practice is complex, and intuitive anthropomorphism can arise at every stage of research: model construction, experimental design, data gathering, and model choice. At each stage, it can arise in different ways, and bias at one stage will be compounded in later stages: attention and recall are biased towards stereotype confirming evidence (e.g. confirmation bias supporting implicit bias; Bodenhausen & Wyer, 1985; Darley & Gross, 1983). This, along with the subtlety of cognitive biases generally, mean that the effects of anthropomorphic bias can be complex and, in at least some cases, counterintuitive (recall Ham).

Putting this all together, there is good reason to believe that intuitive anthropomorphism is problematic in a context like comparative psychology. It is triggered by simple stimuli like eyes and hands, and its effects can be subtle and unpredictable. Awareness of the bias, even with deliberation and expertise, is not sufficient to control it. We cannot be sure how and to what degree any specific judgment about animal cognition is (or is not) influenced. So what can we do?

3. Controlling Implicit Social Bias

The literature on controlling implicit biases provides the best current evidence about controlling cognitive biases, and some of the interventions discussed there are especially instructive to the current discussion. So I turn to that literature now. The upshot is that taking steps to ensure that counter-stereotypical information is salient in reasoning is more effective than simply directing participants to avoid or 'correct' for bias.

The Weapon Identification Task is a common test of bias. Participants identify an image as a tool or a weapon, but are very briefly shown an image of a white or black face before. If, for instance, subjects mistakenly say that a tool is a weapon after a black face more often than a white face, that is taken as a characteristic indicator of bias.

Not all interventions that make intuitive sense work. Using this task, Payne, Lambert, & Jacoby (2002) tested the effects of instructions on bias. Participants in the two experimental conditions were first told that research had shown that people possess implicit racial biases that can influence performance on the task, and then told either to "avoid race" or "use race" in making their decision. Controls were not given any of these instructions. In fact, the two experimental groups were *more likely* to make errors consistent with racial bias than the control

group. Thus, the authors conclude, calling attention to race has the effect of *increasing* bias, whether participants are told to use or avoid it. The instruction itself activates racial stereotypes that lead to biased judgments. Similarly, Legault, Gutsell, & Inzlicht (2011) showed that motivating participants to avoid racial bias *increases* bias if they perceive the motivation as coming from an external source, but can reduce bias if it is seen as self generated.

Stewart & Payne (2008) found a more effective approach. They used *implementation intentions*, which are if-then action plans that make it easier for participants to accomplish a goal than general intentions to do so. For instance, the implementation intention "if I leave work, I will stop and exercise at the gym" is more effective than the general intention "I will exercise more." Stewart & Payne asked participants to form one of three implementation intentions. The first: "whenever I see a black face on the screen, I will think the word, *accurate*." The second: "whenever I see a black face on the screen, I will think the word, *quick*." The third: "Whenever I see a black face on the screen, I will think the word, *safe*." Those participants were told: "By thinking the word 'safe,' you are reminding yourself on each trial that you are just as safe interacting with a Black individual as with a White individual" (pg. 1336). Only the 'think safe' condition reduced bias.

Another common intervention is to show participants images of admired or counter-stereotypical members of the stereotyped group. Dasgupta & Greenwald (2001) found this effect using the Implicit Association Test. In one manipulation, Dasgupta & Greenwald (2001) showed participants images of admired black individuals and disliked white individuals before giving them the IAT. They found that doing so reduced bias, whether the images were presented immediately before or 24 hours before administering the IAT. Govan & Williams (2004) achieved a similar result using counter-stereotypical examples in the experiment itself, even

using non-social stereotypes about flowers and insects. Imagining a positive, productive interaction with members of the stereotyped group before performing an IAT can also reduce bias (Turner & Crisp 2010).

So the most consistently effective interventions are those that make counterstereotypical information salient in reasoning, like the 'think safe' intention, or imagined interactions. What doesn't work is demanding accuracy. I now apply the lessons of the discussion so far to existing methods of controlling anthropomorphism.

4. Morgan's Canon

Perhaps the most commonly advocated method of addressing anthropomorphism is an updated version of *Morgan's Canon*, a famous statement by the 19th century comparative psychologist C. Lloyd Morgan (1894). The modern interpretation of Morgan's Canon councils that researchers should adopt the model that describes the *simplest* psychological process that can predict behavior. This practice is widespread, and is still motivated largely by concerns about anthropomorphism (Graham 1998, Manning & Dawkins 1998, Shettleworth 2010, Wilder 1996, Wynne 2007). In order to control anthropomorphism generally, it must control intuitive anthropomorphism specifically.

de Waal (1998), Sober (2005), and Fitzpatrick (2008) have argued that Morgan's Canon leads to errors of 'anthropodenial,' the underestimation of animal intelligence (more on their reply in section 5). My arguments so far show why: Morgan's Canon explicitly aims to correct a bias that *consistently overestimates* intelligence, but, first, intuitive anthropomorphism does not always lead to errors, and second (think of Ham) its errors need not have anything to do with

intelligence. So Morgan's Canon sometimes aims to correct errors that were not made, and sometimes aims to correct the wrong error.

My arguments also suggest another problem: even in cases where intuitive anthropomorphism has lead someone to overestimate the intelligence of an animal, Morgan's Canon may not be effective. Effective interventions to control implicit bias make counter-stereotypical information salient. Morgan's Canon does not. It is more similar to the ineffective interventions: it is a demand to make the 'unbiased' judgment, like the Stewart & Payne (2008) 'think accurate' condition, or the Payne, Lambert, & Jacoby (2002) 'avoid race' instruction. The effect Morgan's Canon has depends on how it is represented by the person using it. If researchers view it as an external demand to avoid anthropomorphism, it could actually *increase* bias (Legault, Gutsell, & Inzlicht 2011). Similarly, if researchers view it as being *about* anthropomorphism, it could activate anthropomorphic stereotypes, and increase bias, like the "use race" and "avoid race" instructions from Payne, Lambert, & Jacoby (2002).

So how this plays out will vary on a case by case basis. But this leaves Morgan's Canon in a bad position. Firstly, intuitive anthropomorphism does not always lead to error, and when it does, it does not always overestimate intelligence. In these cases, Morgan's Canon produces errors in the other direction. Secondly, even in the cases in which it should be most effective, the empirical evidence suggests that it may be ineffective or even increase bias.[4] Morgan's Canon should not be used to control intuitive anthropomorphism.

5. Evidentialism

[4] Also, as mentioned, intuitive anthropomorphism can influence any stage of research. A rule about model choice, like Morgan's Canon, doesn't address anthropomorphism at other stages.

The next control is proposed by Sober (2005) and Fitzpatrick (2008) as replacements for Morgan's Canon. Sober concludes his paper with the memorable line: "The only prophylactic we need is empiricism" (2005 pg. 97). Fitzpatrick frames the claim as a principle he calls 'evidentialism' (2008, pg. 242). The shared idea is that one should wait until there is sufficient evidence before adopting a hypothesis.

This, of course, is good advice. The problem is that inferences about *what is sufficient evidence* are potentially subject to intuitive anthropomorphic bias. This is, arguably, exactly what is at issue in this debate. We need guidelines that *specifically* address intuitive anthropomorphism, so evidentialism alone is not enough.

6. Identifying Errors

Many authors discuss more specific instantiations of anthropomorphic error. This suggests a third strategy. The more specific the errors we can identify, the better positioned we are to avoid them. Simply warning against 'anthropomorphism' in general is too vague to be helpful.

De Waal (1999) argues against *anthropocentric anthropomorphism*, which is attempting to explain behavior by simply imagining how you would approach the task. He emphasizes that researchers need to consider the perspective of the animal itself. Thus, there is an identifiable difference between instances of pernicious anthropocentric anthropomorphism and what he calls *animal-centric* anthropomorphism. Rivas & Burghardt (2002) describe the related *anthropomorphism by omission*, which assumes that the capacities of nonhuman animals are a subset of our own. In reality, nonhuman animals have many capacities that we do not.

Lockwood (1986) distinguishes several kinds of anthropomorphism, the most interesting of these is *explanatory anthropomorphism*. This is the fallacy of thinking we have explained a

behavior simply by giving it a (folk psychological) name. Many raise related worries about the use of folk psychological terminology, which might make it easy to unintentionally slip into intuitive anthropomorphism. Authors worry about this to different degrees. Kennedy (1992) argues for the complete replacement of folk psychological terminology, while de Waal (1999) argues for the use of neutral descriptive terminology when observing and recording data, though folk psychological can be used theoretical interpretation. Finally, Bekoff (2000) argues that general care is all that is needed.

Povinelli lists several specific errors (2012, appendix 1), for instance, the *end-point training effect*: one watching an animal perform some task can be struck by her success, and forget that the behavior required hundreds or thousands of practice trials. Another, he calls *The Erin Moriarty effect* (after a journalist who visited his lab): in tasks in which an animal has a 50/50 chance of 'success,' trials in which it succeeds *feel* more meaningful.

Whether or not these specific proposals stand up, there is a useful idea here. The problem is that there is little reason to believe that we will be able to identify a list of discrete errors that exhausts intuitive anthropomorphic bias. The effects of cognitive bias are subtle, complex, and often unpredictable. And we need controls that will generalize to experiments not yet done. So, like evidentialism, this approach is helpful, but more is needed.

7. Building New Strategies

I have argued that awareness of bias in deliberation is not an effective control itself, but it might seem hasty to argue that it has no role. Either way, though, a proper understanding of bias is essential. So given that the field has implicit anthropomorphism wrong, adopting the view that I have argued for can only help. But we need more than this. Of existing controls, Morgan's

Canon should be eliminated, while evidentialism and identifying errors can be helpful, but are not enough. So I suggest that we look to the literature on controlling social biases for new strategies.

The literature on controlling implicit biases suggests an empirical hypothesis. Imagining intelligent seeming actions performed by a being that is (relatively) unlikely to be anthropomorphised, such as a computer, an insect, or an octopus (Eddy, Gallup, & Povinelli 1993), might reduce subsequent intuitive anthropomorphism. This is analogous to imagining positive interactions with members of stereotyped groups. As an empirical claim, this needs to be tested, and testing it will not likely be easy.

In the meantime, there are proposals better supported by the evidence discussed above. Implementation intentions reduce implicit bias because they help ensure that a specific goal is implemented at the right time (Stewart & Payne 2008). For practicing comparative psychologists, this would be much more difficult because the goals will have to be much more complicated. So, in place of implementation intentions, I suggest checklists. Checklists have been helpful in many settings, including airline takeoffs, engineering, and surgery (Borchard et al. 2012), and have been suggested as a method to reduce implicit bias in judges (Seamone 2006). Checklists can ensure that a large number of complicated intentions are enacted at the right time.

This checklist might include two sets of items. One set should specify alternative hypotheses that might explain the behavior. This should include hypotheses beyond that which is favored in the field and that which is the researchers own hypothesis: Heyes (2015) argues that many well-known effects that could explain behaviors are often ignored in specific debates in comparative psychology (One could also produce new hypotheses by imagining how something unlikely to be

anthropomorphised could perform the task). Another set of items could come from a more systematic program of identifying errors, modelled on the heuristics and biases literature (e.g. Gillovich, Griffin, & Kahneman 2002). A more complete *taxonomy* of intuitive anthropomorphic errors can help in many cases, even if it is not alone sufficient. If there is an alternative hypothesis, or there is reason to believe an error has been made, one must step back and reassess the evidence.

These suggestions of the last section need to be tested, and perhaps better measures can be found, but for now, these are the options best supported by the empirical evidence we have.

8. Conclusion

One might reasonably wonder what this discussion means for the current state of comparative psychology, as I have explicitly avoided this question. This discussion does not tell us whether the field in general overestimates the intelligence of nonhuman animals. While overestimating intelligence is one effect of intuitive anthropomorphism, it is only one of them. It may not even be a dominant effect; it is easily identified, so its apparent prominence may simply be an matter of availability.

One effect that the dynamic between intuitive anthropomorphism and Morgan's Canon might have is to produce a stark divide between sophisticated 'cognitive' capacities, often couched in folk psychological terms, and overly simplistic associative explanations. This concern has been raised elsewhere (Penn 2011, Heyes 2015), but considerations raised here can (partially) explain it. Not only do folk-psychological models seem more intuitively plausible, intuitive anthropomorphism can influence the *building* of models. So, either models are

constructed using the few mathematical tools we have (like associative modeling) or under the heavy influence of folk psychology. Beyond that, I suspect there is little that can be said in general, and progress will have to be made on a case by case basis.

References

Arico, A., Fiala, B., Goldberg, R. F., & Nichols, S. (2011). The Folk Psychology of Consciousness*. Mind & Language, 26(3), 327-352.

Banks, R. R., Eberhardt, J. L., & Ross, L. (2006). Discrimination and implicit bias in a racially unequal society. *California Law Review*, 1169-1190.

Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. Cognitive psychology, 31(3), 219-247.

Bekoff, M. (2000). Animal emotions: exploring passionate natures. *BioScience*, *50*(10), 861-870.

Bertrand, M., & Mullainathan, S. (2003). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination (No. w9873). National Bureau of Economic Research.

Bodenhausen, G. V., & Wyer, R. S. (1985). Effects of stereotypes in decision making and information-processing strategies. Journal of personality and social psychology, 48(2), 267.

Borchard, A., Schwappach, D. L., Barbir, A., & Bezzola, P. (2012). A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. Annals of surgery, 256(6), 925-933.

Caporael, L. R., & Heyes, C. M. (1997). Why anthropomorphize? Folk psychology and other stories. *Anthropomorphism, anecdotes, and animals*, 59-73.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. Journal of Personality and Social Psychology, 44(1), 20.

de Waal, F. B. (1991). Complementary Methods and Convergent Evidence in the Study of Primate Social Cognition1). Behaviour, 118(3), 297-320.

de Waal, F. B. (1999). Anthropomorphism and anthropodenial: consistency in our thinking about humans and other animals. Philosophical Topics, 255-280.Dyer , F.C . 2002 : The biology of the dance language . Annual Review of Entomology , 47 , 917 – 949 .

Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT press.

Eddy, T. J., Gallup, G. G., & Povinelli, D. J. (1993). Attribution of cognitive states to animals: Anthropomorphism in comparative perspective. Journal of Social Issues, 49(1), 87-101.

Fiala, B., Arico, A., & Nichols, S. (2011). On the psychological origins of dualism: Dual-process cognition and the explanatory gap. Creating consilience: Integrating science and the humanities, 88-110.

Fitzpatrick, S. (2008). Doing away with Morgan's Canon. Mind & Language, 23(2), 224-246.

Gallup Jr, G. G., Marino, L., & Eddy, T. J. (1997). Anthropomorphism and the evolution of social intelligence: A comparative approach. In *Anthropomorphism, anecdotes, and animals*,

Gebhard, U., P. Nevers, and E. Billman-Mahecha. 2003. "Moralizing Trees: Anthropomorphism and Identity in Children's Relationships to Nature." In *Identity and the Natural Environment: The Psychological Significance of Nature*, S. Clayton and S. Opotow (Eds.), 91–112. Cambridge, MA: MIT Press.

Govan, C. L., & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. Journal of Experimental Social Psychology, 40, 357 – 365.Graham 1998

Green, Alexander R., et al. "Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients." Journal of general internal medicine 22.9 (2007): 1231-1238.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. The American Journal of Psychology, 243-259.

Heyes, C. (2015). Animal mindreading: what's the problem?. Psychonomic bulletin & review, 22(2), 313-327.

Kennedy, J. S. (1992). The new anthropomorphism. Cambridge: Cambridge University Press.

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. Cortex, 53, 60-77.

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages How motivational interventions Can reduce (but also increase) prejudice. Psychological Science, 0956797611427918.

Lockwood, R. (1986). Anthropomorphism is not a four-letter word. In *Advances in Animal Welfare Science 1985* (pp. 185-199). Springer Netherlands.

Mitchell, S. D. 2005. Anthropomorphism and cross-species modeling. In Thinking with Animals: New Perspectives on Anthropomorphism, 100–117, ed. L. Daston and G. Mitman. New York: Columbia University Press.

Mitchell, T. L., Haw, R. M., Pfeifer, J. E., & Meissner, C. A. (2005). Racial bias in mock juror decision-making: a meta-analytic review of defendant treatment. *Law and Human Behavior*, *29*(6), 621.

Morewedge, C. K., Preston, J., & Wegner, D. M. (2007). Timescale bias in the attribution of mind. *Journal of personality and social psychology*, 93(1), 1.

Morgan, C. L. (1894). Introduction to comparative psychology. London, UK: Walter Scott Ltd.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Science*s, 109(41), 16474-16479.

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. Journal of Experimental Social Psychology, 38(4), 384-396.

Penn, D. (2011). How folk psychology ruined comparative psychology: and how scrub jays can save it. *Animal thinking: contemporary issues in comparative cognition. MIT Press, Cambridge*, 253-266.

Povinelli, D. J. (2012). World without weight: Perspectives on an alien mind. Oxford University Press.

Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does unconscious racial bias affect trial judges?. *notre dame law review*, *84*(3), 09-11.

Rasmussen, J. L., & Rajecki, D. W. (1995). Differences and similarities in humans' perceptions of the thinking and feeling of a dog and a boy. Society & Animals, 3(2), 117-137.

Rivas, J., & Burghardt, G. M (2002). Crotalomorphism: A metaphor for understanding anthropomorphism by omission. In M. Bekoff, C. Allen, & G. M. Burkhardt (Eds.), The cognitive animal: Experimental and theoretical perspectives on animal cognition (pp. 9-17). Cambridge, MA: MIT Press.

Seamone, E. R. (2006). Understanding the person beneath the robe: Practical methods for neutralizing harmful judicial biases. *Willamette L. Rev.*, 42, 1.

Shettleworth, S. (2010) Cognition, evolution, and behavior, 2nd edn. Oxford, New York.

Sober, E. (2005). Comparative psychology meets evolutionary biology. Thinking with Animals, 85-99.

Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. Personality and Social Psychology Bulletin, 34, 1332 – 1345.

Turner, R. N., & Crisp, R. J. (2010). Imagining intergroup contact reduces implicit prejudice. British Journal of Social Psychology, 49, 129 – 142.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.

Uhlmann, E. L., & Cohen, G. L. (2007). "I think it, therefore it's true": Effects of self-perceived objectivity on hiring discrimination. *Organizational Behavior and Human Decision Processes*, 104(2), 207-223.

Wilder, H. 1996. Interpretative cognitive ethology . In M. Bekoff and D. Jamieson (eds), Readings in Animal Cognition . Cambridge MA : MIT Press .

Wynne, C. D. (2007). What are animals? Why anthropomorphism is still not a scientific approach to behavior. Comparative Cognition & Behavior Reviews, 2, 125-135.