ROCZNIKI FILOZOFICZNE
Tom LXXI, numer 3 – 2023
DOI: https://doi.org/10.18290/rf23713.10

MAJA BIAŁEK

THE NEW ANTHROPOMORPHISM DEBATE AND RESEARCHING NON-HUMAN ANIMAL EMOTIONS: A KANTIAN APPROACH

INTRODUCTION

Anthropomorphism is one of the first concepts that spring to mind whenever the subject of non-human animal[1] emotions is brought up. Despite the strong sceptical tradition of denouncing anthropomorphism, it has often been argued that certain forms of anthropomorphism are either inevitable or even helpful in investigating animal behaviour, and this way of thinking is becoming increasingly popular among scientists studying animal emotion.

The current debate focuses on separating the "critical" and "folk" forms of anthropomorphism, often oversimplifying the multilevel nature of anthropomorphising. The idea that anthropomorphising in a self-aware, theoretically informed manner is both inevitable and valuable in animal studies is probably correct—many doubts remain, however, as to what we can do to truly anthropomorphise critically. Debating the pros and cons of anthropomorphising and proposing ways to make it less risky and more disciplined should also consider a more nuanced picture of the phenomenon. Different layers of anthropomorphism are amenable to changes and adjustments to different degrees and serve different purposes.

I put forth that we should approach anthropomorphism in analogy to how Kant teaches us to approach epistemology in general. In the Kantian tradition, we need to embrace the fact that our experience of the world is determined by the way our minds (and, especially in the phenomenological school, our bodies) are made.

MAJA BIAŁEK, PhD—University of Białystok (UwB), Institute of Philosophy; correspondence address: Instytut Filozofii UwB, Pl. NZS, 15-420 Białystok, Poland; e-mail: majasbialek@gmail.com; ORCID: https:// orcid.org/0000-0002-6491-6689.

[1] Throughout the paper, for the sake of brevity, I will be referring to non-human animals as simply animals. This is, however, only a stylistic choice, and I do not intend to suggest that the category of "animals" excludes humans in any way.

We cannot go beyond what is made possible by the deep structures of human cognition. It may even be said that in examining the world we keep examining those very structures, the necessary conditions of possibility for our experiences. Therefore, the Kantian way of approaching anthropomorphism is to accept it as a necessary condition of possibility for our research on other animals—and then to reconstruct how anthropomorphising works, what are the assumptions and mechanisms that make it possible, and what are the relations between the varying levels of anthropomorphism and the roles anthropomorphising plays in our theories, ethics, and daily life.

In my reconstruction I will identify three levels of anthropomorphism. The current debate seems to be focusing on two: the most often criticised narrative level, the anthropomorphic stories we tell to explain animal behaviour, and the cognitive, implicit level, the deeply rooted mechanisms that determine the way we categorise, observe, and understand even the simplest phenomena in the animal world. However, I will show that there is also a third level, overlooked but crucial. This is the metatheoretical level of anthropomorphism—one that truly makes all the other levels possible and mediates between the simple observations and the abstract stories. A Kantian approach will help me describe this level and show how it both fuels and complicates emotion research in general. Importantly, it will also help me provide valuable pointers on how to make this level of anthropomorphism "critical".

The Kantian approach to epistemology is often understood as pessimistic or even sceptical: we *don't* have access to the way things are in themselves, we *only* get to know our own cognitive structures. But there is also an optimistic tradition of interpreting Kant: we *can* examine our own minds, and in doing so, we *can* determine the necessary conditions of possibility of the phenomenal world, the world as we experience it. Moreover, even if we discover certain limitations to the ways we experience the world—for example, that all our experience is necessarily spatiotemporal—we can still examine them and even move beyond them in our theories. Our experience remains necessarily Newtonian, but our physics has become Einsteinian. I believe that the same can be gained by examining anthropomorphism. A Kantian approach would teach us that it is futile or even impossible to renounce anthropomorphism altogether—but we do not have to interpret this pessimistically. By examining anthropomorphism, we do get to understand our research much better. Furthermore, a good grasp of the structure of anthropomorphism in emotion research enables us to move beyond our anthropomorphic way of both experiencing and theorising and thus discover and question our unfounded assumptions and correct errors.

My argument will unfold as follows: in section 2, I will briefly present the contemporary debate on anthropomorphism and its roots. I will offer a case study for anthropomorphism in emotion research and introduce three questions that require answering if we are to be genuinely critical in anthropomorphising animals, one about the structure and nature of anthropomorphising, one about its virtues and dangers, and one about ways of improving our anthropomorphic strategies. The first question will be answered in section 3, in which I present and describe the three levels of anthropomorphism, emphasising the metatheoretical level. The second question will be answered in section 4, in which I shortly discuss the varieties of gains and risks associated with the three levels of anthropomorphism. Finally, section 5 outlines two ways of working on the metatheoretical anthropomorphism by making our theoretical framework easily generalisable and evolutionarily informed.

1. TRADITIONAL ANTHROPOMORPHISM AND THE ANTHROPOMORPHIC TURN

According to Wynne (2007), the term 'anthropomorphism' began to be applied to the study of animals in the second half of the nineteenth century, which is, of course, the time of the Darwinian revolution. Darwin himself is known to attribute human traits to other animals freely, and he is usually praised for the self-conscious and cautious way he did it both by the critics of the anthropomorphic method (Wynne 2004, 2007) and by authors with a more lenient attitude (Crist 1996). In Darwin's works, anthropomorphism was the critical tool to understand and highlight the continuity of evolution (Crist 1996), and emotions became Darwin's point of focus—using the same anthropomorphic vocabulary for humans, and other animals fostered building a unified theory of how life functions and evolves.

The tradition of censuring anthropomorphism and treating it as the greatest sin a scientist researching animal behaviour could commit is equally long, and strongly reinforced by the behaviourist school in psychology. According to Lloyd Morgan, the author of the famously anti-anthropomorphic principle (the so-called "Canon") that is still sometimes taught to comparative psychologists, it should be forbidden to attribute a "higher psychological faculty" to an animal if we could make do with one "which stands lower in the psychological scale" (Morgan 1894, 53). We need to note that for Morgan (as is rightfully pointed out by Cartmill 2000), it was implicitly obvious that evolution was a linear process, that humans were

the final stage of this process, endowed with the highest psychological faculties, and that what came later in evolution was also necessarily more complex.

Dacey (2017) proposes that in the narrowest, perhaps most traditional sense, anthropomorphism has simply been considered a type of error: we ascribe to animals such intellectual properties or capabilities that they do not possess. This understanding also implicitly assumes that humans have the said properties or capabilities and that those properties or capabilities are more advanced, which need not be the case. At least some of the dangers associated with this kind of error have always been clear: we prevent ourselves from understanding animal behaviour, and we create a warped view of the evolution of specific capabilities. Some are vaguer—anthropomorphising is associated with sentimentality and being unscientific, both as causes and effects, our false beliefs can have poor practical and ethical consequences.

The new wave of anthropomorphism and anti-anthropomorphism is partly grounded in Darwin's and Morgan's tradition, but the debate is now much more multi-dimensional. It has become a widely accepted claim that using anthropomorphic terms to understand animal behaviour can bring us closer to understanding other animals and ourselves. This idea has been essential for contemporary ethologists such as Jane Goodall or Frans de Waal, whose vivid descriptions of non-human animals have been central to the development of modern animal studies. However, many more authors subscribe to the idea that anthropomorphism can be defended, albeit only if it is critical and informed (Bruni *et al*. 2018; Williams *et al*. 2020).

Modern proponents of anthropomorphism emphasise that what they have in mind is a rigorous, practical and scientific approach. Burghardt (1991, 1997, 2006, 2016) proposes a distinction between "naive" anthropomorphism and "critical" anthropomorphism, an approach that he deems fruitful in animal studies. His idea has been roughly followed by other authors advocating a theoretically useful form of anthropomorphism—such as constructive anthropomorphism (Arbilly and Lotem 2017), or forms of biocentric, non-anthropocentric anthropomorphism popularised by Frans De Waal and Marc Bekoff. All those concepts describe a cautious theoretical approach, in which the researcher is well-guarded against sentimentalism or naivete. We can allow ourselves to be inspired by anthropomorphic comparisons but only if we are conscious of doing so and prepared to inform our comparisons with as much knowledge as we can gather about the possible analogies and disanalogies between the animals we study and humans.

The defenders of critical anthropomorphism emphasise that if anthropomorphism is to be helpful, it must encourage us to adopt the perspective of the animal

we are researching and focus on its unique qualities rather than allow us to project our own features upon other species directly. We do not want to force non-human animals into our categories; we want to get to know them for who they are, appreciating the vast differences between our species. Some authors, like Bekoff and de Waal, claim that to keep this danger at bay, it might be enough that we denounce anthropo*centrism*. Anthropo*morphism* does not necessarily entail retaining the human perspective, keeping human interests at heart, looking out for human features, always seeking our traits in other animals, and fixating on the constant comparisons between them and us. True, we can never adopt another species' perspective in the sense outlined by Thomas Nagel in the classic paper "What Is It Like to Be a Bat?" (NAGEL 1974). However, we can try to put ourselves in the animal's boots, paws or wings in many ways—learn as much as we can about its ecological niche, way of life, interests, needs, likes, and wants. Other authors propose new terms that would emphasise the focus on the other animal's point of view and denounce anthropocentric forms of anthropomorphism. Timberlake (1997, 2006) suggests "theromorphism" and Milton (2005) opts for "egomorphism", a term coined to express that we should not attribute our traits to other animals in an artificial, abstract way—rather, we can genuinely learn to perceive the similarities, see what is human in other creatures.

A powerful argument for critical anthropomorphism, or, at least, against a straightforward anti-anthropomorphism, is that from the methodological standpoint it is unreasonable to assume that non-human animals do *not* share any human qualities. Especially in the case of animals close to us from the evolutionary perspective, such as chimpanzees, it would be unfounded to interpret their behaviour in a different way than we would were they human beings. Frans de Waal would call this mirror anti-anthropomorphic approach anthropodenialism (DE WAAL, 1999), and his popular book on animal emotion, *Mama's Last Hug*, is indeed a poignant illustration of why anthropodenialism can only be counterproductive if we aim to truly understand animal emotions (2019). Buckner (2013) also warns us against the tendency to treat our capabilities as unachievable for any other animal, calling it 'anthropofabulation'.

It is still in the area of animal emotion that critical anthropomorphism is most often defended as helpful or even indispensable (BEKOFF 2000). For example, Burghardt (2016) offers a good argument for employing critical anthropomorphism in animal studies, showing how it can help resolve the issue of fish experiencing pain. I would like to examine in more detail an example provided by Williams *et al*. (2020) to illustrate their thesis that the anthropomorphic approach can be fruitful, the study of grief in primates, because it is a perfect case study for the latter sections of this paper.

Many observations have been made that some primate mothers, after their offspring's death, tend to carry the corpse of their dead infant on their backs for extended periods (for a detailed cross-species study see (FERNÁNDEZ-FUEYO *et al*. 2021). Various explanations have been offered so far, but one prominent approach is that such behaviour is a form of grief, which, in turn, inspired further speculation concerning the experience and understanding of death in primates. The anthropomorphic idea that chimpanzee mothers experience pain after the death of their loved one is not only helpful in explaining their striking behaviour but also drives fascinating research into animal rituals and the evolution of human ideas about death. Still, nothing about this example is straightforward, and the claim that chimpanzees experience grief requires much unpacking.

(i) What exactly does "anthropomorphising" mean here? Why is ascribing "grief" to a chimpanzee anthropomorphic?
(ii) What is grief? Can we even answer the question in the context of human emotion research?
(iii) What are we gaining by ascribing grief to chimpanzees?

We may already suspect that answering those questions might prove extremely difficult, which suggests another question:

(iv) How can we be "critical" in anthropomorphising if answering questions (i) and (ii) is so difficult?

I will now employ the Kantian strategy to answer those questions, distinguish and describe the three levels of anthropomorphising I have outlined and sketch out an answer to the crucial question (iv).

2. THE THREE LEVELS OF ANTHROPOMORPHISING

Let us begin with question (i): Why is ascribing "grief" to a chimpanzee "anthropomorphic"?

The key problem is that although we have so many concepts at our disposal and such vast theoretical literature on the subject, it has become even less clear what it means to ascribe human traits to non-human creatures. Which features should we consider to be "specifically human" and why? We are not sure what it is, precisely, that anti-anthropomorphism forbids, and anthropomorphism encourages, and, in consequence, what ways of enhancing anthropomorphising practices or criticising them are still valid.

Wynne, a modern critic of anthropomorphism, identifies the essence of anthropomorphism—in his view problematic—as "mentalism" (WYNNE 2007), "a belief that the imaginative projection of one's mentalistic self into the life of a member of another species can lead to the production of hypotheses which may prompt the production of useful objective data" (132). The crux of the matter is, for Wynne, that projecting one's mentalistic self" is something profoundly subjective and cannot be transformed into a standardised, replicable, scientific practice, regardless of how "critical" we might be. Wynne also observes that focusing on mentalist explanations—for example describing the behaviour of a naughty puppy as "showing remorse"—does not explain anything, it may even prevent us from seeking deeper mechanisms behind the dog's behaviour.

Applying mentalist language to non-human animals is obviously difficult for the reasons pointed out by Penn and Povinelli (2007) in a paper arguing against ascribing the Theory of Mind to any non-human creature. There are simply no experimental protocols that could convince us that animals have "beliefs", "representations", or other mental states. Of course, there is also no objective way to know that about any human—however, in the case of humans, we can at least rely on introspective reports. Most humans seem to believe that they experience subjective mental life and express such beliefs in intersubjective language. The research on animal communication is flourishing, but as of now, no known non-human way of communicating can express beliefs about having beliefs. Thus, the anthropomorphic manner of ascribing beliefs or emotions to other animals is inescapably subjective. This is part of the reason why Wynne (2007) claims that no amount of "being critical" can transform mentalism into a standardised, reliable scientific method.

Importantly, there is also a huge variety of "mid-tier" categories that we associate more with behaviour than with complex subjective experiences. Bruni *et al.* 2018 present the fact that Panksepp studied "laughing" rats as an example of (arguably untypical) anthropomorphism because laughing used to be considered a uniquely human activity and what Panksepp actually observed was a regular chirp "seemingly related with positive emotional states". Another not-really-mental example of anthropomorphising is "play", also discussed by Bruni *et al.* (2018). Laughing and playing are not typically considered exclusively human, advanced mental states or abilities—they may (but need not) implicitly imply the existence of such states. In our cases, we could also ask about the adequacy of such "mid-tier", "mildly" anthropomorphic terms and their relation to the more controversial ones. In the following sections of the paper, I will be examining a narrative in which ascribing "smiling" to chimpanzee has proven

anthropomorphic in a problematic way—primarily because we tend to (anthropomorphically) link smiling with happiness.

Keeley (2004) argues that it is incredibly anthropomorphic to think of anthropomorphism as "attributing human qualities to animals". What gives us the right to classify certain features as "human"? We observe a range of qualities shared by all primates, or even by all mammals—attributing them to non-human animals is not "anthropomorphic" but "mammalomorphic" or "primatomorphic". Although, as I have mentioned, in the times of Lloyd Morgan, we might have laboured under the impression that evolution was linear and that humans possessed certain features that are, at the same time, "newest" and "most advanced", we now know the story is much more complex. We share different qualities with different groups of creatures, and those features might have evolved convergently on wildly different branches of the evolutionary tree. We—philosophers especially—are accustomed to the constant search for what makes humans unique, for pinpointing those features that only humans can boast. We are used to thinking that those qualities have to be something "better" than what other animals have—most notably, that they involve possessing special mental or cognitive capacities. We rarely think, however, that what makes us unique is not necessarily something that is "best" or "most complex". Like any other species, humans are a unique combination of features that have developed in our evolutionary history, which we share with so many different creatures. And all those features shape how we understand other creatures and their lives. It is anthropomorphic to ascribe "happiness" to a seal (an example provided by Bliss-Moreau 2017). Still, it is also anthropomorphic in a sense to even say that a seal "swims" because our whole idea of movement, let alone movement in the water, is so deeply determined by how we experience movement in our human mammalian body.

Bigger and more detailed pictures of what anthropomorphising is have been offered by Varella (2018) and Dacey (2017). Varella identifies within anthropomorphism a varied set of cognitive biases, such as teleology, hyper-mentalizing and hyper-theory of mind, and "agenticity" (3). He enumerates four basic modes or stances of thoughts that govern our cognition: the physical stance, the design stance, the basic-goal stance and the belief stance. To explain the source of anthropomorphism, Varella employs Shermer's popular concept of "patternicity" (Shermer 2008), the tendency to seek (and find) meaningful or familiar clusters of various kinds in any clusters of data, even if there aren't any. For Varella, anthropomorphic patternicity consists in overactivating or overextending such stances that are inapplicable or not applicable to that degree to the subject in question. In general, however, anthropomorphising is just an inevitable con-

sequence of how our cognitive system works—it is the set of cognitive mechanisms that shape our observations.

Similarly, Dacey proposes that anthropomorphism is a cognitive bias that can influence all stages of our inquiries, both scientific and every day. However, he goes beyond the idea that specific categories are inappropriately overactivated or overextended. Dacey defines anthropomorphism broadly as the whole process of forming beliefs about animal minds by analogy to our own and observes that 1) it is not always erroneous, and 2) if it is, this is not necessarily because the capabilities we ascribe to the animal are "higher" or overly advanced.

This is, again, especially noticeable in the case of animal emotion. Dacey encourages us to examine the case of Ham, a chimpanzee launched into space in 1961. There is a picture of Ham right before the launch—Ham's facial expression resembles a human smile, which is why lay commentators often assume that he is "happy" or even "proud" (presumably of the vital role in human history he is about to play). In fact, the chimpanzee's expression is most probably one of fear. Dacey claims that the error of attributing happiness to Ham instead of fear is due to "intuitive anthropomorphism"—and please note that it has nothing to do with overestimating the chimpanzee's intellectual abilities. I would like to point out that this is a great example of how our tendency to anthropomorphise is multi-faceted and influences different stages of belief-forming. We see Ham as happy partly due to our semi-automatic interpretation of his specific expression, which is a fast, primary, and low-level cognitive mechanism. It is also grounded in many abstract, high-level (and subjective) simulations of how we might feel in a similar situation, our beliefs about the importance of our undertakings, and possibly our tacit wish not to notice the suffering we are causing.

Ham's case is very similar to our case study of the grieving chimpanzee: we are forming a belief about the chimpanzee's mind by analogy to our own and "projecting our mentalistic selves". We may assume that multiple cognitive processes take place: we interpret certain expressions and behaviours in a semi-automatic way and employ more abstract simulations and hypotheses, perhaps also motivated by other considerations (for example, we may be keen to protect chimpanzees or otherwise biased to seek signs of advanced cognitive development). We end up with a whole narrative that explains Ham's "smile" with a folk-psychological theory about his ambitions and motivations.

I propose that from a cognitive perspective, we can so far distinguish two basic levels or stages in the process of anthropomorphising.

2.1 Cognitive anthropomorphism and narrative anthropomorphism

I'm inclined to accept Dacey's way of thinking about anthropomorphism as the whole process of forming beliefs about other animals in analogy to how we perceive our minds. Dacey focuses on the beginnings of this process—the stage he dubs "intuitive anthropomorphism"—the initial perceptions and categorisations offered by our cognitive apparatus (either innate or acquired early in life), and therefore based on the basic anthropomorphic framework we learn and employ to deal with understanding humans. He does so partly because in debating anthropomorphism we tend to focus overly on the final stage—when we consciously construct anthropomorphic narratives about our observations. We criticise creating sentimental stories about animals and tend to underappreciate the biases that shape our more basic observations. For future reference, I will call the first stage cognitive and the last—narrative.

The relations between those stages are complex and two-way. A helpful analysis has been provided by De Cruz and De Smedt (2015), who tackle anthropomorphism in the context of religion and theology. The original hypothesis put for by some cognitive scientists of religion has been that we only learn to understand and attribute divine properties such as omniscience via cultural learning (the narrative level)—our natural, intuitive tendency (the cognitive level) was said to be "anthropomorphising", that is: implicitly assuming that God is as fallible as all humans are. This was supposed to be why people are often stubbornly "theologically incorrect" (i.e., sceptical about God's total knowledge) despite the efforts of theologians. The whole picture that emerges from empirical evidence is much more complicated. Let me examine it in more detail, as it is instructive and touches directly on the issue of anthropomorphising other animals.

First, it turns out that there is also a form of omniscience ascribed intuitively by children who are not yet able to pass the false belief test (the so-called "reality bias")—not only to God but to any agent. Second, older and religiously educated children have no problem correctly ascribing omniscience to God despite already being aware that humans are fallible. In a particularly interesting experiment in which children were asked to pass a version of the false belief test with regards to their mother, a bear, an ant, a tree and God, it turned out that anthropomorphic intuitions of fallibility are mediated by knowledge (Barrett *et al*., 2001). If taught so by their culture, children do not ascribe any false beliefs to God—although they do ascribe them to their mother. Additionally, children can correctly apply other information to gauge the knowledge of animals such as bears or ants—for example, they consider the animals' sensorimotor abilities. Perhaps

the most exciting result of this experiment was that the tree was the most anthropomorphised agent of all—the one with the most similar rate of false belief ascription to the rate of the mother. The interpretation of this last result is that children anthropomorphise the most what they understand the least, which supports the guidelines that so far have been given by the proponents of critical anthropomorphism.

De Cruz and De Smedt explain this complex picture by referring to the now prominent dual process theories of thinking ("fast" and "slow", as popularised by Kahneman's famous 2011 book). Some schemes and mechanisms are innate or emerge very early in our lives and provide us with fast, automatic, but inflexible answers. They are generally useful for survival but also make us susceptible to illusions and prone to specific errors. It is often costly to overcome the fast responses of the automatic systems. On the other hand, there are also the flexible, reflective, inference-rich, slow processes rooted in our culture and education. In between, there emerges a "practised naturalness", a set of cognitive mechanisms that the inference-rich systems have already informed, but which have become semi-automatic with practice and time. This corresponds well to the distinction between the narrative and the cognitive levels of anthropomorphising I have outlined above and foreshadows the existence of an in-between, metatheoretical level I will describe in the following section. However, before I do so, I need to point out three subtle disanalogies between my analysis and De Cruz and De Smedt's case.

First, in the case of omniscience, the "anthropomorphic" stance—the automatic ascribing of fallibility—is also something not innate but acquired, albeit early on. There is an even deeper, simpler, more automatic attitude—the "reality bias" that needs to be overcome with experience in mentalising. Second, in the case of omniscience, the anthropomorphic stance is then, in turn, overwhelmed by the input from the more conscious, flexible deliberations of the "slow" system. Third, we need to remember that the example of ascribing *divine* omniscience to people, animals, and God is *not* analogous to ascribing typically *human* emotions to non-human animals. In the case of omniscience, anthropomorphism happens between the lowest cognitive levels and the narrative levels of the belief formation process. In the case of animal emotions, it happens on *all* levels—both in the "fast" and in the "slow" systems, and, as I will explain shortly, in between.

2.2 What happens between the cognitive and narrative levels of anthropomorphism? The chaos of psychology and the intermediate level

As we remember, the cognitive and narrative levels corresponded to our simple observations of the chimpanzee's behaviour and expressions based on innate or well-practised mechanisms, and the highly abstract story invoking the complex, high-level, cultural concepts of grief and ritual. The two constitute parts of the same process thanks to a common conceptual framework that binds them together. It is the conceptual framework of psychology and folk psychology. This framework is deeply problematic, and anthropomorphism is only part of its problems.

Ascribing grief to a chimpanzee is problematic not only because our analogies may be subjective, flawed or without a good base, and not only because we can't be sure if we are not engaging in what Varella would call patternicity. It isn't easy primarily because we are not sure what we are ascribing and what the objective criteria should be for ascribing it to any being. Let us briefly turn to question (ii): What does 'grief' mean?

Lonsdorf *et al*. (2020) observe that "grief has been difficult to define and operationalise systematically; even in humans, grief is recognised as a 'highly individualised and dynamic process' (Cowles and Rodgers 1991, 119)" (13). Although the paper Lonsdorf *et al*. refer to is more than thirty years old, much newer reviews of emotion research, such as de Vere and Kuczaj (2016), Paul *et al*. (2020) and Kremer (2020), compound this scepticism. Both de Vere and Kuczaj (2016) and Kremer (2020) point to severe and practically insurmountable difficulties in synthesising research on non-human emotions because of the conceptual and terminological chaos and lack of unifying theories. Paul *et al*. (2020) highlight the merits of the componential view of emotion, according to which emotions comprise no less than five "*loosely* coordinated changes" (p. 6) in feeling (subjective experience), cognition (appraisals), action, expression, and physiology. We can only have a general intuition of what we mean by any specific emotion.

Anthropomorphic terms (of all levels) are problematic in animal emotion research partly because the human psychological framework, in general, is not uniform. Typical mentalism, ascribing folk-psychological notions such as "belief" or "concept" to any being, humans included, is a potentially risky theoretical move, as we have no clear, uncontroversial theories behind any of those notions. We may acquiesce to the fact that our folk psychology is vague and unclear.

However, we tend to overlook that empirical psychology, and cognitive sciences are still *in statu nascendi* and are not unified in any sense. Terms like 'emotion', 'concept', or 'belief' have myriads of technical definitions rooted in specific theories and varied levels of explanation.

Moreover, terminological chaos has at least two dimensions. It is one thing that in emotion research, even the basic terms, such as 'emotion', 'affect', 'feeling', or even 'laugh', can mean vastly different things depending on the theoretical approach, with no synthesis in sight. It is another—that we should not and cannot be overly ambitious as regards the scope of our claims about "non-human animals in general", as is clear from De Vere and Kuczaj's review. Simply put, when mentalist terms are concerned, there is no sense in referring to "animals" in general. And yet, papers and books on "animal emotion" are often filled with examples and anecdotes concerning practically all living creatures and all kinds of behaviour that could be associated with emotions. Grief and happiness in mammals are freely combined with aversive behaviour in molluscs or fish, bees' or bumble bees' moods are thrown into the mix, and even plants, the newest and trendiest area of focus, often make an appearance. How is it possible that we seem to be bent on building any kind of coherent picture from all of this?

This sounds like a Kantian question—and I posit that our tendency to seek synthesis and see patterns in such chaos is rooted in yet another hidden or meta-theoretical level of anthropomorphism. The source of this kind of intuitive, theoretical synthesising is that in humans, it happens that pain, laughter, moods, grief and happiness belong loosely in the same realm of life. This evolutionary fact has shaped our psychology and emotion research—and it is driving our metatheoretical approach. Metatheoretical anthropomorphism provides us with a deeply ingrained framework that determines not only the concepts we apply but also what type of questions we ask and how we attempt to synthesise the answers. It is not enough to acknowledge that we should not nonchalantly mix and match empirical results concerning vastly different species and achieved via entirely different approaches and paradigms. We need to be wary of the fact that our ideas about emotions are anthropomorphic also on the metatheoretical level and that this determines the conditions of possibility of any type of animal research. This intermediate, metatheoretical anthropomorphism ingrained in our psychology and cognitive science is what mediates between fast, simple observations and slow, elaborate narratives.

Having outlined this three-fold structure of anthropomorphism in answer to question (i), we are much better equipped to tackle the crucial questions (iii) and (iv): what are the possible gains and how to make all the levels more "critical"?

3. WHAT ARE THE GAINS AND THE DANGERS?

The anthropomorphism literature often focuses overly on criticising the most high-level, narrative types of anthropomorphism, particularly the "folk" variety, most vibrantly expressed in the narratives expressing how we view and treat our pets. It is almost unanimously perceived as foolish and detrimental both to the welfare of companion animals and, possibly, to animal studies, should the everyday penchant for interpreting animal behaviour in terms of human psychology seep into scientific research. It is associated with sentimentalism and a kind of naivete—and often justly so. An example of this is how animal photos or videos are often misleadingly captioned on social media, as has been the case with Ham's photo. In a viral video, an "empathetic crow" is purportedly "helping" a hedgehog cross the road, which makes it "better than most humans!" (In fact, it is either nudging it to safety simply to eat it in comfort or, on a less bleak interpretation, just being playful.) A puppy is "remorseful" after it has wreaked havoc on the caregiver's living room. Similarly, objections against spaying stray and pet animals or against keeping cats indoors are sometimes grounded in naively anthropomorphic arguments referring to the animals' "wish to have a child", their need for "self-realisation" and "love of freedom". However, the optimistic lesson from the previous section of this paper is that this kind of anthropomorphism is also the one we can influence in the most conscious way. It is flexible and grounded in our cultural practices.

I would also like to point out that although narrative folk anthropomorphism may lead to many harmful misconceptions and mistreatment of pets, even this naive form of anthropomorphism is not without its virtues. In a world in which humans, in general, are still largely unaffected and unbothered by the mass-suffering of non-human animals, it might be the lesser of two evils if they over-attribute human characteristics at least to some groups than if they are further encouraged in the wide-spread practice of commodifying living beings. Maybe the quarrels about whether cats have the existential need to wander freely through the streets of our cities are a reasonable price to pay if this gives us any starting point for the debate about whether hens, cows or pigs genuinely do not need ever to leave their factory pens. Again, although the narrative level is probably the one which could be eliminated, it is also the most flexible. "Folk" anthropomorphic stories can be easily corrected and turned into ones that are informed and grounded in our knowledge. Narratives of this kind are also, as it has been shown, inspiring and valuable in animal studies.

Regarding the inflexible cognitive level, Dacey (2017) suggests that we take heed of what psychology teaches us about controlling any kind of implicit bias. For this level, it is not enough to simply demand criticism or accuracy. Going against our biases, as has also been shown in the experiments on religious anthropomorphism, is costly and requires extraordinary effort. The implicit bias literature suggests that the most effective strategy for diminishing bias is to make counter-stereotypical information salient to the participant. Dacey proposes that this idea can take the form of a checklist of counter hypotheses and alternative explanations of behaviours that are less accepted in the field—making such a checklist requires, in my view, diving straight into the in-between, meta-theoretical Kantian level of anthropomorphism.

The primary role of metatheoretical anthropomorphism is to enable us to synthesise our observations and narratives. To talk about whether certain behaviours are indicative of grief in a chimpanzee, we need the intermediary of human psychology—not only to grasp the concept of grief alone but to have a theoretical idea of what emotions are, that they comprise certain physiological, psychological and cognitive phenomena as well as subjective experiences. Also, Nemati (2022) has recently argued that anthropomorphic attribution is, on a metatheoretical level, a creative link between observation and building testable hypotheses, thus enabling scientific discovery . We can't even attempt to operationalise grief in animal studies without first adopting the underlying anthropomorphic psychological framework—any modifications can only be introduced afterwards.

Of course, we can and should worry that metatheoretical anthropomorphism can lead us astray, often without us even realising. We have some control over it—but only if we make a conscious effort and only to some degree. To paraphrase Thomas Nagel, we could never know what it's like to be a bat psychologist. However, I find Dacey's (2017) suggestion inspiring. We can make counter-stereotypical information salient for ourselves, check other possibilities, and question all the elements of the framework we are adopting. Importantly, it is not enough to provide a checklist of alternate explanations. In our example, it would not be enough to ask ourselves whether perhaps the chimpanzee's behaviour is not a sign of happiness and celebration rather than grief. We need to carefully examine the framework we apply to other species and the assumptions we implicitly make about the structure of emotions.

In the last section of this paper, I will answer question (iv) by presenting two strategies for how we can develop a more critical metatheoretical anthropomorphic strategy for researching animal emotion.

4. UNIFICATION AND EVOLUTION: WAYS TO PERFECT METATHEORETICAL ANTHROPOMORPHISM

4.1 FLEXIBLE AND UNIFIED APPROACHES TO EMOTIONS

As I have mentioned, the wealth of conflicting theoretical approaches to emotions has proved to be another obstacle in our attempts at understanding non-human animals. However, if our goal is to counter our metatheoretical assumptions, there are some promising directions.

Bliss-Moreau claims that the research on animal emotions has been dominated by a "classic view", involving the assumptions that there exists a particular set of basic emotions which are biologically hardwired, evolutionarily conserved and have discrete, specific, behavioural and biological outputs (BLISS-MOREAU 2017). Barrett (2016) argues that the aforementioned "classic view" has been empirically disproved, at least to a degree, because there is too much variation even in human emotional behaviour, including the ways reported emotions are expressed both outwardly and in the nervous system. Any attempts at forcing non-human animals into the moulds that do not even consistently match human emotions seem unfounded, to say the least. Bliss-Moreau suggests that we must turn to theories of constructed emotions (of which Barrett's theory is the most prominent example). This family of theories brings together some of the classic intuitions—that emotions emerge from many ingredients or components—but offers more leeway to labelling them and seeking consistent types and, crucially, combinations.

The theory of constructed emotions says that we have no general, biologically engrained kinds of emotions—instead, they are dynamically constructed by organisms during their interactions with the environment. The "matter" from which emotions are constructed is termed 'affect', which is a state of arousal with negative or positive valence. The phenomenon of experiencing specific types of emotions is constituted by the fact that our brains construct them via culturally, linguistically and socially grounded concepts. Why is this construction at all needed? Barrett employs the predictive processing theory, which she applies to interoceptive inference. Predictive processing theories, in general, point to a unified explanation of most of our cognitive activity as an attempt to perfect our predictions and thus minimise possible surprise. This model has been initially used to explain visual perception—as the result of constantly modifying our beliefs to explain better/predict incoming perceptual data. For Barrett's purposes, however, we need to concentrate on the inner workings of our organism. Wilkinson et al. (2019) describe her theory succinctly as "James's old subtraction theory with a predictive twist". It is, indeed, similar to the classic Jamesian

approach since it traces the origins of emotional experience to bodily reactions and interoceptive signals our organism receives. Then comes the twist: to explain the incoming signals, predict what might happen next, and motivate quick responses that could modify those predictions, the predictive system needs to employ concepts of emotions. In the Kantian idiom, this is the necessary condition of possibility of a conscious emotional experience.

The most significant advantage of this approach to how the human emotional experience is generated is that it is a potentially general account. All components—the interoceptive, the predictive and the conceptual parts of the story—are very flexible, also in the metatheoretical sense. We not only can examine, at least to a degree, how other species might experience their own bodies with their specific sets of senses and ways of interacting with the environment, but we can also imagine different ways in which emotions could be constructed and the kinds of "concepts" that could mediate this process.

Predictive theories are, if anything, almost too universal as they do not put any substantial limitations on how their basic notions such as concepts or perceptions should be understood. What is certain is that in this theory, we are not forced to think of emotion concepts as something necessarily linguistic in the human sense. Rather, we can and should investigate them in a theromorphic, non-anthropocentric approach based on carefully considering a particular species' ways of life and ecological niche. Moreover, although predictive processing is typically treated as a brain-centric theory, with many researchers attempting to seek neural correlates of specific elements of the predictive mechanism, we are not forced to make any human-centric assumptions that have often plagued comparative research on emotions, for example concerning the crucial role of typically human brain structures. We are encouraged to think of emotions as something happening to the whole organism, regardless of how it is built. We could also imagine that all three components can be combined differently in different species. At the same time, this theory is specific enough about the structure of emotions and how the organism implements them to allow for creating detailed checklists containing counter-stereotypical hypotheses.

In our case study, we deal with an animal we can expect to be very similar to humans. This approach would prevent us from using the same labels for emotions, such as "grief" or "sadness"—but the intuition that we deal with what in our species would be categorised this way can still be a valid theoretical inspiration. We can, then, attempt to model the predictive process that ties the chimpanzee's interoceptive experiences of its body with its interaction with its ecological niche. Our checklist can, therefore, include different models of the process.

4.2 Evolutionary anthropomorphism

The second opportunity for informing and improving our checklists of alternatives goes back to the roots of this whole debate. The best way to examine whether our Kantian metatheoretical assumptions about the components of emotions can be called into question is to consider empirical arguments provided by the study of evolution.

The "James's story with a predictive twist" discussed above is just a revamped version of one of the oldest ways of understanding emotions: our body's appraisal of what is happening to it. For any evaluative experience, there must be a being that has this experience, a perspective, and a self which may consider the goings on as dangerous, harmful or positive and beneficial. This understanding of emotions is undoubtedly a part of those deeply engrained "Kantian" structures that we apply in our research, regardless of whether the focus is on bodily expression, neural correlates, or concepts and labels. We should expect, however, that in modelling emotions as appraisals of this sort, we inevitably assume a very human notion of how this can happen. We have an anthropomorphic idea of appraising and an anthropomorphic idea of a "self". Naturally, neither of those ideas has to apply to other species. Our notion of self is not anthropomorphic just because it may involve self-consciousness, subjectivity,[2] self-narratives, or language. Our whole idea of what it is like to be a self is not universal—we have the body of a primate, a mammal, a very particular animal. Our way of being a self is just one of the ways devised by evolution.

In his newest work, Peter Godfrey-Smith (2019, 2020) sketches just how complex the evolution of selves has been, and one of his aims is to explain the evolution of consciousness via evaluative and sensory experience. The sensory side is related to perception and having a point of view—the evaluative side to pain, pleasure, and judging events as good or bad. As Godfrey-Smith (2020) observes, both types of experience would fit into the traditional descriptions of consciousness within the philosophy of mind. They involve "how things seem to an animal", and they both can be used to describe "what it's like" to be a particular animal (247–48). Godfrey-Smith's fundamental thesis is that the "evaluative" and "sensory" paths are independent of the evolutionary point of view.

This is contrary to the profound anthropomorphic intuition we have as mammals that to evaluate something is to experience it as good or bad for us and that to experience something is to have a sensory perception of it. It is part of the

[2] For a discussion of how our assumptions about the link between consciousness and subjectivity influence ethology see Marchesini 2016.

metatheoretical anthropomorphism in emotion research that emotional experience has to combine subjectivity, sensory perception and positive or negative affect or even judgement. We are no longer as quick to deny other animals emotions on the grounds that they do not possess specific linguistic capabilities—we are still reluctant, however, to think of emotions without sensory experience. Godfrey-Smith, however, makes a strong case for his claim that from an evolutionary standpoint, those things have emerged separately and did not converge in many creatures.

His case studies—mainly insects and molluscs—led him to formulate two essential claims. First, we can find empirical examples of creatures highly capable of evaluative experience and not sensory experience and vice versa. Many insects, for instance, are perfectly capable of perceiving/sensing their surroundings but are known to show no behavioural signs of evading pain, which has been often cited as proof of their not being able to feel it at all and could be interpreted as their not being able to judge whether those experiences are good or bad. On the other hand, as Godfrey-Smith points out (2019), slugs seem capable of complex evaluative experiences despite having relatively poor sensory capacities, which can be shown in experiments involving instrumental learning. Second, in accordance with our modern approach to evolution, we learn again that there is no point in seeking evolutionary hierarchies or straight developmental paths. Insects or molluscs are not just "previous steps in the history of evolution", as we would say in the times of Lloyd Morgan—instead, they are huge, diverse groups of animals that have been evolving for millions of years. Godfrey-Smith, a famous cephalopod specialist and octopus-lover, is well aware that some molluscs are well endowed with both kinds of experience. Even among slugs, various experiments bring unclear results.

Moreover, Godfrey-Smith is not suggesting that the stiff division between sensory and evaluative sides can become our new go-to, hardwired framework for reconstructing the evolutionary history of emotional experience. Quite the contrary: he openly points out further difficulties. For example, insects, most notably bees and bumblebees, despite being, as we have just claimed, relatively incapable of evaluative experience, seem to be able to experience *moods*, defined as a longer-lasting emotional state. This suggests that disentangling sensory and evaluative sides might not be enough—we may still be blind to differences in modes of experience because of our Kantian anthropo- or mammalomorphism.

Moreover, we will never understand non-human animals if we stop at the animal kingdom. Another of Godfrey-Smith's tools that can help us disentangle our deepest anthropomorphic assumptions is the emphasis on the self—not as we

are programmed to understand it, but as it has evolved. Again, we should not think about it as a linear process—from the first coherent multicellular organisms to primates with their only slightly lateralised brains. Instead, we can trace how selves appeared and which changes influenced the emergence of new types of bodies (and new types of experiences).

Jékely, Godfrey-Smith and Keijzer (2021) describe one such particular form of body-self—one that includes motility and sensing, and one in which, notably, action and sensing are tied together through reafference, enabling it to act as a single unit. From the metatheoretical anthropomorphic point of view we adopt when researching emotions, it is almost difficult to realise that we implicitly assume that to experience emotions is to be a self and that a self is something that—as we do—can act as a single unit thanks to feedback loops between its senses and its motor systems. Realising that "selves" can be varied requires much more than simply leaving behind our ideas about basic emotions experienced by all and any species—but can also open our horizons in a much more exciting way.

Godfrey-Smith's work on the genesis of the "self" is especially inspiring in the context of the increasingly debated possibility that plants also could experience something that could be classified as emotions. From our "Kantian anthropomorphic" perspective, we may acknowledge all the differences between plants and animals but still, find clear evolutionary paths through the tightly woven branches. There is still some coherence in the plant "self", although many other features are lacking. Moreover, this coherence is also tied to how plants are capable of some forms of action and movement, albeit those forms are very difficult for us to classify. We cannot change the time frame and the time rate we live in, and we are possibly unable to perceive the slow processes that take place in plants' roots in response to chemical danger as "action" or "emotion". However, we can work out the evolutionary and conceptual framework that could help us understand the barely coherent plant-self whose parts nevertheless communicate with each other to avoid danger and a grieving chimpanzee.

In our case study, although we are dealing with an animal which is very close to humans in the tree of evolution, we can still pose valid questions about the structure of the self and the proportions of the "evaluative" and "sensory" components in the experience. Naturally, considerations of this kind are crucial whenever we attempt to study animals who are much more different to ourselves.

CONCLUSION

In this paper, I have supported a Kantian strategy toward anthropomorphism. First, we need to examine it as something that determines how we perceive ourselves and other animals. It is not just a naïve way of (mis)understanding animal behaviour. It is also not just an uncontrollable cognitive bias, nor is it simply a helpful narrative. It is all of those things at once—and more. Above everything, it is the implicit metatheoretical framework that makes it possible for us to build coherent theories about our observations and our narratives.

How anthropomorphism permeates our implicit theoretical assumptions remains underappreciated and thus has not received enough attention. However, if we are to reap the benefits promised by the proponents of modern forms of anthropomorphism, we need, first and foremost, to tackle the metatheoretical level of anthropomorphising. It is not enough to be critical, nor is it enough to simply apply as much knowledge about the perspective and the ecological niche of the animal we research as we can. The alternative hypotheses that we should examine must consider not only how certain behaviours could be explained differently but also how our theoretical assumptions on what emotions are can be challenged.

We can use two sources of inspiration to challenge them fruitfully and helpfully. First, such theories are universally applicable, allowing for flexible interpretations of the mechanisms they postulate but specific enough to remain informative. Second—lessons from evolution, mainly focusing on the evolution of sensory and evaluative experience and the self.

REFERENCES

ARBILLY, Michal, and Arnon LOTEM. 2017. "Constructive Anthropomorphism: A Functional Evolutionary Approach to the Study of Human-Like Cognitive Mechanisms in Animals." Proceedings of the Royal Society B: Biological Sciences 284 (1865): 20171616. https://doi.org/10.1098/rspb.2017.1616.

BARRETT, Justin L., Rebekah A. RICHERT, and Amanda DRIESENGA. 2001. "God's Beliefs versus Mother's: The Development of Nonhuman Agent Concepts." *Child Development* 72 (1): 50–65. https://doi.org/10.1111/1467-8624.00265.

BARRETT, Lisa F. 2016. "The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization." *Social Cognitive and Affective Neuroscience* 12 (1): 1–23. https://doi.org/10.1093/scan/nsw154.

BEKOFF, Marc. 2000. "Animal Emotions: Exploring Passionate Natures." *BioScience* 50 (10): 861. https://doi.org/10.1641/0006-3568(2000)050[0861:AEEPN]2.0.CO;2

Bliss-Moreau, Eliza. 2017. "Constructing nonhuman animal emotion." *Current Opinion in Psychology* 17 (October): 184–88. https://doi.org/10.1016/j.copsyc.2017.07.011

Bruni, Domenica, Pietro Perconti, and Alessio Plebe. 2018. Anti-Anthropomorphism and Its Limits. *Frontiers in Psychology* 9, 2205. https://doi.org/10.3389/fpsyg.2018.02205.

Buckner, Cameron. 2013. "Morgan's Canon, Meet Hume's Dictum: Avoiding Anthropofabulation in Cross-Species Comparisons." *Biology & Philosophy* 28 (5): 853–871. https://doi.org/10.1007/s10539-013-9376-0.

Burghardt, Gordon. 1991. "Cognitive Ethology and Critical Anthropomorphism: A Snake with Two Heads and Hognose Snakes That Play Dead." In *Cognitive Ethology: The Minds of Other Animals: Essays in Honor of Donald R. Griffin*, edited by Carolyn A. Ristau, 53–90. Hillsdale, NJ: Lawrence Erlbaum Associates.

Burghardt, Gordon. 1997. "Amending Tinbergen: A Fifth Aim for Ethology. In *Anthropomorphism, Anecdotes, and Animals*, edited by Robert W. Mitchell, Nicholas S. Thompson, and H. Lynn Miles, 254–76. SUNY Press.

Burghardt, Gordon. 2006. Critical Anthropomorphism, Uncritical Anthropocentrism, and Naïve Nominalism. *Comparative Cognition & Behavior Reviews* 2. https://doi.org/10.3819/ccbr.2008.20009.

Burghardt, Gordon. 2016. "Mediating Claims through Critical Anthropomorphism." *Animal Sentience* 3 (17). https://doi.org/10.51291/2377-7478.1063.

Cartmill, Matt. 2000. "Animal Consciousness: Some Philosophical, Methodological, and Evolutionary Problems." *American Zoologist* 40 (6): 835–46. https://doi.org/10.1093/icb/40.6.835.

Cowles, Kathleen V., and Beth L. Rodgers. 1991. "The Concept of Grief: A Foundation for Nursing Research and Practice." *Research in Nursing & Health* 14 (2): 119–27. https://doi.org/10.1002/nur.4770140206.

Crist, Eileen. 1996. "Darwin's Anthropomorphism: An Argument for Animal-Human Continuity." *Advances in Human Ecology* 5:33–83.

De Cruz, Helen D., and Johan De Smedt. 2015. *A Natural History of Natural Theology: The Cognitive Science of Theology and Philosophy of Religion*. MIT Press.

de Vere, Amber J., and Stan A. Kuczaj. 2016. "Where Are We in the Study of Animal Emotions?" *Wiley Interdisciplinary Reviews: Cognitive Science* 7 (5): 354–62. https://doi.org/10.1002/wcs.1399.

de Waal, F. 1999. "Anthropomorphism and Anthropodenial: Consistency in Our Thinking about Humans and Other Animals." *Philosophical Topics* 27 (1): 255–80.

de Waal, F. 2019. *Mama's Last Hug: Animal Emotions and What They Tell Us about Ourselves*. W. W. Norton & Company.

Fernández-Fueyo, Elisa, Yukimaru Sugiyama, Takeshi Matsui, and Alecia J. Carter. 2021. "Why Do Some Primate Mothers Carry Their Infant's Corpse? A Cross-Species Comparative Study." *Proceedings of the Royal Society B: Biological Sciences* 288 (1959): 20210590. https://doi.org/10.1098/rspb.2021.0590.

Godfrey-Smith, Peter. 2019. "The Evolution of Consciousness in Phylogenetic Context." In *The Routledge Handbook of Animal Minds*, edited by Kristin Andrews and Jacob Beck, 216–26. London: Routledge.

Godfrey-Smith, Peter. 2020. *Metazoa: Animal Life and the Birth of the Mind*. New York: Farrar, Straus and Giroux.

JÉKELY, Gáspár, Peter GODFREY-SMITH, and Fred KEIJZER. 2021. "Reafference and the Origin of the Self in Early Nervous System Evolution." *Philosophical Transactions of the Royal Society B: Biological Sciences* 376 (1821): 20190764. https://doi.org/10.1098/rstb.2019.0764.

KAHNEMAN, Daniel. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

KEELEY, Brian L. 2004. "Anthropomorphism, Primatomorphism, Mammalomorphism: Understanding Cross-Species Comparisons." *Biology & Philosophy* 19 (4): 521–40. https://doi.org/10.1007/sBIPH-004-0540-4.

KREMER, L., S. E. J., KLEIN HOLKENBORG, I. REIMERT, J. E. BOLHUIS, and L. E. WEBB. 2020. "The Nuts and Bolts of Animal Emotion." *Neuroscience & Biobehavioral Reviews* 113:273–86. https://doi.org/10.1016/j.neubiorev.2020.01.028.

LONSDORF, Elisabeth V., Michael L. WILSON, Emily BOEHM, Josephine DELANEY-SOESMAN, Tessa GREBEY, Carson MURRAY, Kaitlin WELLENS, and Anne E. PUSEY. 2020. "Why Chimpanzees Carry Dead Infants: An Empirical Assessment of Existing Hypotheses." *Royal Society Open Science* 7 (7): 200931. https://doi.org/10.1098/rsos.200931.

MARCHESINI, Roberto. 2016. Philosophical Ethology and Animal Subjectivity. *Angelaki* 21 (1): 237–52. https://doi.org/10.1080/0969725X.2016.1163859.

MILTON, Kay. 2005. "Anthropomorphism or Egomorphism? The Perception of Non-human Persons by Human Ones." In *Animals in Person. Cultural Perspectives on Human-Animal Intimacy*. London: Routledge. https://doi.org/10.4324/9781003135883.

MORGAN, C. Lloyd. 1894. *An Introduction to Comparative Psychology*. London: The Walter Scott Publishing Co.

NAGEL, Thomas. 1974. "What Is It Like to Be a Bat?" *The Philosophical Review* 83 (4): 435. https://doi.org/10.2307/2183914.

NEMATI, Farshad. 2022. "Anthropomorphism in the Context of Scientific Discovery: Implications for Comparative Cognition. Foundations of Science." https://doi.org/10.1007/s10699-021-09821-1.

PAUL, Elisabeth S., Shlomi SHER, Marco TAMIETTO, Piotr WINKIELMAN, and Piotr MENDL. 2020. "Towards a Comparative Science of Emotion: Affect and Consciousness in Humans and Animals." *Neuroscience & Biobehavioral Reviews* 108:749–70. https://doi.org/10.1016/j.neubiorev.2019.11.014.

PENN, Derek C., and Daniel J. POVINELLI. 2007. "On the Lack of Evidence That Non-Human Animals Possess Anything Remotely Resembling a 'theory of mind'." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362 (1480): 731–44. https://doi.org/10.1098/rstb.2006.2023.

SHERMER, Michael. 2008. "Patternicity." *Scientific American* 299 (6): 48–48. https://doi.org/10.1038/scientificamerican1208-48.

TIMBERLAKE, W. 1997. "An Animal-Centered, Causal-System Approach to the Understanding and Control of Behavior." *Applied Animal Behaviour Science* 53 (1–2): 107–29. https://doi.org/10.1016/S0168-1591(96)01154-9.

TIMBERLAKE, W. 2006. "Anthropomorphism Revisited." *Comparative Cognition & Behavior Reviews* 2. https://doi.org/10.3819/ccbr.2008.20010.

VARELLA, Marco A. C. 2018. "The Biology and Evolution of the Three Psychological Tendencies to Anthropomorphize Biology and Evolution." *Frontiers in Psychology* 9, 1839. https://doi.org/10.3389/fpsyg.2018.01839.

WILKINSON, Sam, George DEANE, Kathryn NAVE, and Andy CLARK. 2019. "Getting Warmer: Predictive Processing and the Nature of Emotion." In *The Value of Emotions for Knowledge*,

edited by Laura Candiotto, 101–19. Springer International Publishing. https://doi.org/10.1007/978-3-030-15667-1_5.

WILLIAMS, Lisa A., Sarah F. BROSNAN, and Zanna CLAY. 2020. "Anthropomorphism in Comparative Affective Science: Advocating a Mindful Approach." *Neuroscience & Biobehavioral Reviews* 115:299–307. https://doi.org/10.1016/j.neubiorev.2020.05.014.

WYNNE, Clive D. L. 2004. "The Perils of Anthropomorphism." *Nature* 428 (6983): 606. https://doi.org/10.1038/428606a.

WYNNE, Clive D. L. 2007. "What Are Animals? Why Anthropomorphism Is Still Not a Scientific Approach to Behavior." *Comparative Cognition & Behavior Reviews* 2 (1): 125–35.

THE NEW ANTHROPOMORPHISM DEBATE AND RESEARCHING NON-HUMAN ANIMAL EMOTIONS: A KANTIAN APPROACH

Summary

Researchers of non-human animal emotions tend to defend some forms of anthropomorphism and seek ways to make it more critical, self-aware, and useful for scientific purposes. I propose that to achieve this goal, we need first to conduct a Kantian investigation into the deeper structure of anthropomorphism. I argue that we can distinguish at least three levels of anthropomorphising: a narrative level, a cognitive level and an in-between, metatheoretical level which is the deeper structure determining how we anthropomorphise. Because the current debate tends to focus either on the narrative level or on the cognitive level, this paper concentrates on the metatheoretical level, discusses its role in emotion research, the possible errors it may cause, and how we can work on it, drawing on predictive processing-based theories of emotions and an evolutionary approach. The key to being critical in anthropomorphism is to be aware of the complexity of this whole structure, as well as to be able to challenge and put into question all and any of its elements.

Keywords: anthropomorphism; animal emotion; non-human animals; emotion research; predictive processing.

WSPÓŁCZESNY SPÓR O ANTROPOMORFIZM A BADANIE EMOCJI ZWIERZĄT POZALUDZKICH. PODEJŚCIE KANTOWSKIE

Streszczenie

Badacze emocji zwierząt pozaludzkich mają tendencję do obrony niektórych form antropomorfizmu i szukają sposobów na uczynienie go bardziej krytycznym, samoświadomym i użytecznym dla celów naukowych.

Proponuję, że aby osiągnąć ten cel, musimy najpierw przeprowadzić kantowskie badanie głębszej struktury antropomorfizmu. Twierdzę, że możemy wyróżnić co najmniej trzy poziomy antropomorfizacji: poziom narracyjny, poziom poznawczy i poziom pośredni, metateoretyczny, który jest głębszą strukturą określającą sposób, w jaki antropomorfizujemy. Ponieważ obecna debata zazwyczaj skupia się wyłącznie albo na poziomie narracyjnym, albo na poziomie poznawczym, niniejszy artykuł koncentruje się na poziomie metateoretycznym, omawia jego rolę w badaniach nad emocjami, możliwe błędy, które może powodować, oraz sposób, w jaki możemy nad nim pra-

cować, opierając się na teoriach emocji opartych na przetwarzaniu predykcyjnym i podejściu ewolucyjnym. Kluczem do krytycznego podejścia do antropomorfizmu jest świadomość złożoności całej tej struktury, a także umiejętność kwestionowania i podawania w wątpliwość wszystkich jej elementów.

Słowa kluczowe: antropomorfizm; emocje zwierząt; zwierzęta pozaludzkie; badanie emocji; przetwarzanie predykcyjne

Informacja o Autorce: Dr Maja Białek – Uniwersytet w Białymstoku, Instytut Filozofii; adres do korespondencji: Instytut Filozofii UwB, Pl. NZS 15-420 Białystok, Polska; majasbialek@gmail.com; ORCID: https://orcid.org/0000-0002-6491-6689.