

## NON-HUMAN ANIMAL EMOTIONS: HOMOLOGICAL OR FUNCTIONAL KINDS?<sup>1</sup>

*Emoções Animais Não-Humanas: Tipo Homológicos ou Funcionais?*

---

Juan R. Loaiza<sup>2</sup>

### ABSTRACT

In our daily lives, we attribute emotions to non-human animals. However, the ontological commitments this implies are still in discussion. Particularly, philosophers still debate whether considerations about the mechanisms underlying emotions are necessary or not to attribute emotions to non-human animals. Here, I argue that such considerations are not sufficient, and that a functionalist perspective is more fruitful than its main contender, the homology thinking view. To do this, I consider findings from experimental psychology on emotion attribution to non-human animals and distinguish two questions concerning such emotions. I then discuss functionalism and homology thinking, claiming that homology thinking precludes promising empirical hypotheses from the outset, resulting in a more limited position than functionalism. Hence, functionalism inherits many benefits of homology thinking while providing more productive grounds.

**Key-words:** Non-human emotions. Functionalism. Homology thinking. Multiple realization.

### RESUMO

Em nossa vida diária, atribuímos emoções a animais não-humanos. No entanto, os compromissos ontológicos que isto implica ainda estão em discussão. Particularmente, os filósofos ainda debatem se as considerações sobre os mecanismos subjacentes às emoções são necessárias ou não para atribuir as emoções a animais não humanos. Aqui, eu defendo que tais considerações não são suficientes e que uma perspectiva funcionalista é mais frutífera do que seu principal concorrente, a visão do pensamento homológico. Para fazer isto, considero os resultados da psicologia experimental sobre a atribuição de emoções a animais não humanos e distingo duas questões relativas a tais emoções. Em seguida, discuto o funcionalismo e o pensamento homológico, afirmando que o pensamento homológico exclui desde o início hipóteses empíricas promissoras, resultando em uma posição mais limitada

---

<sup>1</sup> DOI: <https://doi.org/10.51359/2357-9986.2022.256754>

<sup>2</sup> Universidad del Rosario in Bogotá, Colombia. E-mail: [juan.loaiza@urosario.edu.co](mailto:juan.loaiza@urosario.edu.co). ORCID: <https://orcid.org/0000-0003-0570-3832>. Juan R. Loaiza is a lecturer at the Escuela de Ciencias Humanas, Universidad del Rosario in Bogotá, Colombia. He is interested in the intersection between the philosophy of emotions and philosophy of science, particularly in issues concerning natural kinds in emotion research. Other areas of interest include the philosophy of visual perception and social ontology.

do que o funcionalismo. Assim, o funcionalismo herda muitos benefícios do pensamento homológico, ao mesmo tempo em que fornece bases mais produtivas.

**Palavras-chave:** Emoções não-humanas; Funcionalismo; Pensamento homológico; Realização múltipla.

## **INTRODUCTION**

It is common to describe some non-human animals—at least those close to us, such as dogs or cats—as having some emotions. We can talk about dogs being happy to see us, cats being scared, or horses being angry. This suggests that in our folk vocabulary, some emotion concepts apply to non-human animals. Yet, it is unclear how much of our emotion vocabulary applies successfully and how much does not. While it is easy to think of a happy dog, a scared cat or an angry horse, it is unclear how we could talk about an ashamed insect or an embarrassed fish. Hence, we do not know which emotions apply to which non-human animals and how these concepts are construed to allow such application.

Among philosophers, there are a number of views on this question. Some philosophers think that applying emotion concepts to non-human animals, i.e., stating that a non-human animal genuinely has an emotion, is not entirely unwarranted. If this is the case, we can indeed make sense of how emotion concepts apply to non-human animals, even if in a restricted sense. How we should restrict the application of emotion concepts to non-human animals remains a contentious point of debate, though. Defenders of this view usually defend one of two claims. Some argue that since emotion categories can be characterized in functional terms, we can apply emotion concepts to non-human entities (not just animals) as long as the general causal patterns obtain (e.g., Adolphs 2016; Adolphs and Andler, 2018; Loaiza 2021b). Others think that this view is overly liberal, allowing too many organisms to be characterized as having an emotion, and hence we must introduce some restrictions as to the kinds of mechanisms that can implement a genuine emotion (Griffiths 1997; Clark 2010; 2013a; 2013b).

In what follows, I will defend a functionalist view on non-human animal emotions, that is, I will argue that since emotions can be characterized functionally, we can (1) apply some emotion concepts to non-human animals without falling into problematic kinds of anthropomorphizing, and (2) we can overcome the difficulties pointed out by critics of functionalist views without needing to go into the specifics of the mechanisms implementing emotions. As a result, there are reasons to hold a functionalist view as a more overarching, yet powerful alternative to other views in the debate.

To defend my claim, I shall start by presenting some empirical research showing how we apply emotion concepts to non-human animals. This will allow us to restrict which concepts we are actually talking about, and the kinds of vernacular talk that we intend to characterize and understand. Most importantly, this will allow me to formulate two questions which I call the Species Question and the Emotions Question. Second, I will introduce the main idea of the functionalist view I will defend by contrasting functionalism with type-identity theories in philosophy of mind. Third, I defend my functionalist view against the contender of homology thinking, a view which claims that any proper theory of emotions and any use of observations of emotions in non-human animals must delve into specifics about implementing mechanisms. This will result in a more nuanced functionalist view which remains explanatorily useful and ontologically neutral. Lastly, I will explore some consequences for the question of how to account for emotions in non-human animals and for emotion theories in general.

## **1. Vernacular Emotion Concepts and Non-Human Animals**

Before we engage in philosophical discussion, it is important to be clear about the phenomenon we intend to characterize. How do we actually apply emotion concepts to non-human animals? How does this attribution of emotion work? In this section, I will present some empirical studies in psychology showing different ways in which we apply emotion concepts to non-human animals. At this point, I am only interested in describing the phenomenon that I will discuss later, without any relevant assumptions. In other words, I will not assume that said attribution is correct or meaningful.

What I am looking for is setting the stage by showing that, as a matter of fact, people do attribute some emotions to non-human animals in a variety of ways.

There are several empirical studies showing to which animals we do in fact attribute emotions and what sorts of emotions we are willing to project. Wilkins, McCrae and McBride (2015) investigated different factors that might explain why we attribute emotions to animals and which emotions do we attribute the most. The animals they included in their list were mammals (ranging from rats and squirrels to cats, dogs, and horses), birds (pigeons, chickens, and parrots), reptiles (cobras, crocodiles, and tortoises), fish, and invertebrates (cockroaches, fruit flies, and honey bees). They found that belief in the idea that animals have minds is the most important predictor of emotion attribution in animals in general. In other words, if we believe that animals have minds to a higher degree (contrast mammals with insects, for instance), then we tend to be willing to attribute a certain degree of emotionality. More interestingly though, they found that primary emotions (sadness, anger, joy, and fear) are attributed with significant frequency. Secondary emotions (guilt, pride, and jealousy), although less frequently, are also attributed in some cases. This already suggests that humans think of many non-human animals as having at least some emotions, and that not all emotions will have the same range of application.

Morris, Doe and Godsell (2008) replicated similar findings. In this study, the investigators compared attributions between primary and secondary emotions to animals ranging from birds and rats to cats and dogs. They included not only the emotions mentioned above, but also primary emotions like anxiety and surprise, and secondary emotions such as grief, empathy, embarrassment, and shame. Their findings show that primary emotions are more frequently attributed than secondary emotions, except for jealousy. Nevertheless, all emotions tested were attributed at least with some frequency and with relative confidence.

These studies show that we do in fact attribute emotions to animals. What drives these attributions? As I explained above, Wilkins, McCrae and McBride found that belief in animals having minds is an important predictor, although this may be easily explained by the fact that when we project

our mental lives onto animals, we are open to the idea that they share many of our mental capacities. However, the fact that only primary emotions are attributed with a relevant degree of frequency and confidence shows that we do not hold animals capable of the same range of emotions as ourselves. What these studies show is that we project those emotions that we consider primary, that is, less cognitively demanding and less ingrained in our social lives. For instance, in Morris, Doe, and Godsell, social and moral emotions such as embarrassment and shame are the least frequent. In a similar vein, animals with fewer similarities to us such as birds and invertebrates are often considered less emotional than mammals.

Other factors that influence emotion attribution to animals include familiarity and ownership. Wilkins, McCrae, and McBride showed that if participants are asked to rate emotions in animals intended for use (e.g. lab mice, pigs for meat) or pests (e.g. cockroaches, rats), they attribute less emotionality overall in comparison to pets. The researchers explain this effect by appealing to the cognitive dissonance between our attribution of mental states and emotions, and our willingness to use and kill these animals. Whatever the explanation may be, this shows that the closer we feel to animals, the more psychological capacities we are willing to attribute. This, however, does not depend on the number of animals one owns or is related to. Morris, Knight, and Lesley (2012) showed that pet owners attribute a wide range of emotions to a number of animals including dogs, cats, and horses, but this effect is present if participants have at least one animal (there is no difference if they own more than two and there is too much heterogeneity in responses when they own no animals at all, with a slight tendency towards attributing fewer emotions).

Lastly, it is interesting to see that the neural mechanisms behind these attributions are the same as those involved when we attribute emotions to humans. Spunt, Ellsworth and Adolphs (2017) investigated these mechanisms using fMRI. They presented subjects with pictures of human faces, non-human primate faces, dogs, and a control scrambled picture. In one implicit task, they just presented the stimuli inside the scanner. In another explicit condition, they presented the stimuli and asked subjects to tell whether the stimulus matched with a description. In one condition, the description

describes the emotion, asking whether the face is one of boredom, sadness, excitement, etc. In another condition, the description describes the expression itself, asking whether the human or animal is baring teeth, gazing up, opening its mouth, etc. They found that subjects did match the emotional description in all cases, and that the neural mechanisms involved in attributions to humans and animals were relevantly similar. These include activation in the dorsomedial prefrontal cortex (dmPFC), lateral orbitofrontal cortex (LOFC), and anterior superior temporal sulcus (aSTS). These areas are involved in theory of mind and reward processing.<sup>3</sup> Hence, the researchers conclude that there is no human-unique mechanism underlying emotional attribution, but rather a general mechanism that also applies to animals.

What can we conclude from these studies? For our purposes, we can draw two lessons from the findings above. First, we do apply emotion concepts to non-human animals, but not to all animals equally. While there are animals we can easily conceive of as having emotions, particularly those we are familiar with, there are species which we don't think of as emotional. This suggests that whatever the criteria people use for said attribution, the degree to which they are satisfied varies across species. As a result, we should not expect that if one species is described as having an emotion, this must also apply to all species.

A second lesson we can draw from these studies is that just as not all animals have emotions according to our folk attribution, neither are all emotions present in all animals. There are emotions which are easy to extend to non-human animals, such as fear or anger, but there are others that we do not extend as easily, such as shame or embarrassment. Specifically, it seems that social and moral emotions do not apply to most non-human animals. This may not seem surprising at first, but it is an important finding nonetheless, since this shows that for some emotions we expect certain degrees of social and cognitive complexity, while not for others. In other words, we do not think of all emotions as requiring the same capacities, but rather we think of them as a psychologically heterogeneous kind. This will be vital for what I will discuss below.

---

<sup>3</sup> The researchers found some differences in activation between conditions, but they explain them as differences in conceptual familiarity, rather than in attribution.

Now that we have some idea about how people attribute emotions to non-human animals, a number of questions arise. On the one hand, we have evidence that some species, but not all species, are described as having some emotions. This raises the question of which species have emotions and which do not. Call this the *Species Question*. A second question we can formulate concerns the ontology of emotions. We have evidence that among the organisms to which we attribute emotions, not all emotions are present in the same way. Consequently, we can ask which emotions can be said to be present in non-human animals and which emotions are exclusively human. Call this the *Emotions Question*.

Both the Species Question and the Emotion Question must be answered empirically. We must observe the behavior of other species and determine whether these behaviors classify as emotional or not, and which emotion category do they fall under. Yet, this procedure presupposes the possibility that when we observe a given behavioral pattern, we have license to formulate a hypothesis that this pattern is one corresponding to an emotional state. If the way in which we conceptualize psychological phenomena in general, or at least emotions in particular, makes it nonsensical to attribute emotions to non-human animals beyond metaphorical uses, then these hypotheses cannot get off the ground. This means that our ontological commitment and our conceptual framework must be one that admits such hypotheses. What can this framework look like? I will turn to this question below.

## 2. Criteria to Attribute Emotions to Non-Human Animals

The question concerning the criteria to apply emotion concepts to non-human animals mirrors a question largely discussed in philosophy of mind: what are the criteria to apply mental concepts to an organism (carbon-based or otherwise)? There are many answers to this question, but we can identify a continuum along two poles. On one side, there is the view that mental states are identical to some sort of neurophysiological activity. Traditionally, one way of understanding this claim is as stating that mental states are identical to physical states, i.e., type-identity theory (Place, 1956; Smart,

1959). According to this view, a candidate mental concept applies to a non-human organism just in case there is some concrete neurophysiological activity occurring.

In the case of emotions, we can think of examples such as identifying fear with activity in the amygdala. For a type-identity theorist, since fear is said to be identical to activity in the amygdala (perhaps accompanied by physiological effects such as pupillary dilation or increased heart rate), an organism is said to be afraid if and only if it shows such type of activity. This already constrains the set of organisms that are capable of this emotion, as only organisms with an amygdala would be capable of fear, and any other organism that does not have an amygdala cannot be said to be afraid in a genuine sense.

Debates in philosophy of mind have already shown several limitations of such an approach. One common response to type-identity theory is that it is too restricted as to the kinds of organisms capable of mental states. For instance, octopuses and similar species which have largely different neural structures may nonetheless be said to have some mental states akin to those of humans, even if the mechanisms enabling such states is quite different. Such an argument suggests that mental states may be realized in different systems, and that what is important about a given mental concept is not its realization but rather some higher-level description (see e.g., Putnam, 1967/1997).

In addition to these classical responses in philosophy of mind, discussions on emotions specifically also present arguments against type-identities. For type-identity theories to work in the case of emotions, we must find one-to-one correspondences between emotion and neurophysiological mechanisms. However, empirical research has shown that even in humans there is a great variety of mechanisms underlying our emotional responses. For example, Lindquist and colleagues (2012) have argued that canonical correspondences between activity in specific brain regions and emotion categories do not hold. Such cases are the amygdala, which is not only involved in fear but also in detection of salient stimuli in general, or activity in the anterior insula, traditionally associated with disgust but found to be also involved in sensations of bodily awareness and even digestion. As a result,

maintaining type-identity theories would make our emotion concepts collapse, as we have found that the neurophysiological mechanisms corresponding to emotions are not homogeneous.

Arguments such as these give rise to classical functionalism, the view that mental states can be individuated by their causal role rather than physical structures. In philosophy of mind, this position is most famously proposed by Putnam (1967/1997, 1960/1997); in the case of emotions, I have defended such a position elsewhere (Author, 2021b), but the most known approach is Adolphs' (2016; Adolphs & Andler, 2018). According to functionalism, what determines whether an organism can be in a certain mental state is not whether a specific mechanism is at play. Instead, functionalists argue that a given mental state is to be identified with some input conditions, some output conditions, and any number of internal states which explain the process whereby a system reacts in a certain way under given conditions.

In the case of emotions, let us consider again the example of fear. Rather than focusing on amygdala activity, functionalists claim that we should think of fear as a reaction to some type of stimulus, perhaps an event which is perceived by the organism as dangerous. This process is then characterized as having the perception of the event as an input, and the reaction of fleeing or fighting as a response. Between these two states (perceiving the event, fleeing or fighting), we can propose a number of internal states such as evaluating the stimulus as dangerous, considering different reaction strategies, and deciding which strategy might better fit the situation given previous learning. All of these states are described at a higher level of abstraction as the level on which we describe their implementation, and they may be thought of as being potentially implemented by any system capable of these sorts of processes.<sup>4</sup>

In previous work, I have sketched out functionalism as a plausible view on the ontology of emotions. Functionalism can help us understand variability at the neurophysiological level (Author 2021a) and can also be de-

---

<sup>4</sup> A common way of understanding functionalism is under the guise of computational functionalism, the claim that functional descriptions must be cashed out as computations. While I am sympathetic to this view, it is worth noting that computationalism is an additional claim to classical functionalism. See Piccinini (2004, 2010)

fended as an explanatorily interesting view (Author 2021b). I do not wish to go into all of the arguments for functionalism here, but only its consequences for the question of which emotion concepts can be applied to non-human animals. For a functionalist, any system capable of instantiating the relevant processes may be said to have a given emotion, regardless of how it implements these processes. Following the case of fear, if a system can perceive an event and evaluate it as dangerous, and subsequently consider and adopt a fleeing or fighting strategy, then it can be said to be afraid.<sup>5</sup> It is probable that the actual functional description is much more nuanced as the one I am offering, but the general spirit of the view remains the same: identify the causal patterns mapping an input to an output, along with the relevant internal states.

One important consequence of functionalism is that in principle, any system that is complex enough can be said to have any given emotion. Put differently, there are no in-principle constraints as to which animals may have a given emotion, as what matters is that they have the relevant capacity regardless of their neural makeup. For some, this is a problem. According to its critics, functionalism is just too liberal about what constrains our application of emotion concepts, making it explanatorily inert. One such argument is the one offered by Griffiths (1997).

For Griffiths, a complete explanation of emotions must not only consider functional characterizations, that is, descriptions at what he calls the *ecological level* or the level of *task description*, but they must also explain *how* a given task is solved. Functional descriptions, even if they work, can only provide information about the problem an organism solves, but not how evolution has made it such that the organism actually solves it. Instead, we should opt for more constrained explanations that not only focus on the abstract problem-solving aspects of a given behavioral trait, but that also include information about how a specific solution is implemented. This infor-

---

<sup>5</sup> LeDoux (2013) has argued that we should distinguish fear as an emotion from the activation of defense circuits. An organism can activate physiological and behavioral patterns without an intermediate inner “fear” state. Such activation would not be emotional, according to LeDoux, since there would be no inner feeling state. Even though a proper discussion of this argument is beyond the scope of this paper, it is worth considering that for the functionalist, either the feeling state is reducible to a functional state (i.e., it is due to some change that can be characterized functionally) or, if not, then the feeling state is not scientifically tractable.

mation will include facts about what mechanisms underlie a specific solution in a given species, how these mechanisms have evolved, and how they develop ontogenetically. In other words, we should opt for a *homological explanation* of psychological states, rather than a merely analogical one. As Ereshefsky (2007) puts the argument:

The virtues of investigating a trait as a homology rather than an analogy are generic. Griffiths (1994, 1997, 2007) and Matthen (1998, 2000) argue that a phylogenetic approach to psychological categories provides a better basis for explaining the properties of those categories than functionalism. Indeed, citing the phylogenetic history of a psychological category, and the ontogenetic factors within that phylogeny, provides a richer explanation of that category than a mere analogy-based, functional explanation. Consider the example of fear. Attending to the different developmental factors that affect instances of non-doxastic fear explains the different expressions of fear within *Homo sapiens*. It also explains the different expressions among primate and mammalian species. Those different developmental factors, often cultural differences, are themselves bound by distinct histories. More speculatively, if doxastic fear is a less inclusive phylogenetic unit within non-doxastic fear, then a homology-based explanation may explain the differences between those two types of fear (p. 668),

As I understand the argument, functionalist explanations are not taken as inherently wrong, but rather incomplete. Functionalism only gets part of the story of how an organism behaves, as it only describes abstract processes that must be somehow implemented without offering a story about said implementation. Homological explanations, in turn, have the virtues of functionalist explanations in that they can include an abstract level of description specifying a task, but they also give us information about how that task is solved in concrete species, opening the door for phylogenetic and ontogenetic considerations.

On the view Griffiths is defending, the answer to the Emotions Question laid out above must be constrained by what Ereshefsky calls *homology thinking*. Whatever emotions can be said to be present in non-human animals, they must not only satisfy some functional description probably developed by observations in humans, but they must also correspond to some homological structure in other species in order to qualify as the same emoti-

on. Continuing with the case of fear, any organism that is said to be afraid must not only react in the appropriate ways as functionalism would dictate, but these responses must also be implemented in systems that are homologous to those in humans. If a system reacts in a fear-like fashion, but the implementation is largely different, then we must speak of a different emotion altogether, and the application of a fear concept to these species should be qualified as somehow artificial. In Griffiths's (1997) words:

People may agree that an octopus discharges its defensive cloud of black ink "when it is afraid," but the same speakers back off if asked whether the octopus is "really afraid" or if it "experiences emotion." The tendency to disavow these applications or regard them as somehow metaphorical is even stronger in the case of insects and, for some people, artificial intelligences. The application of the concept in its paradigm, human cases, relies on the coincidence of a whole set of features, some at the level of task description, some at lower levels. As some of these features fall away, people become unsure that they are dealing with the same phenomena and unsure whether to apply the concept (p. 233).

The resulting debate can be summarized as follows. For functionalists, details about the implementation of a given emotion are unnecessary to apply that concept to an organism. If an organism satisfies a functional description for a given emotion, then the organism genuinely has that emotion. For defenders of homological thinking, this would be an incomplete application of the emotion concept in question, and perhaps one that is too liberal. Instead, information about the implementation of the functional description, including evolutionary and developmental considerations, are necessary for a proper application of said emotion concept.

In the next section, I will defend functionalism against these arguments. In my view, the worries defenders of homology thinking raise against functionalism are part of its advantages. And just as defenders of homology thinking claim that homology thinking is richer than functionalism, I intend to flip the argument and claim that it is functionalism which provides a richer view. To be clear, this does not mean that homology thinking plays no role in a science of emotions, or in answering the questions I have laid out above. What I will argue is that homology thinking, if it must be put to

work, leads to an overly committed view on the ontology of emotions, falling into problems similar to those pressed against type-identity theories.

### **3. Functionalism and Homology Thinking**

Let us take stock of the state of the debate so far. It seems that humans attribute some emotions to some non-human animals. These emotions tend to be basic (or primary) emotions, and the animals to which they are attributed are generally mammals phylogenetically close to us. To answer the question concerning which emotions can we attribute to non-human animals, we must consider different criteria to say that an organism has an emotion. These criteria are either independent of the mechanisms realizing an emotion (functionalism) or they must consider homologous mechanisms to provide a complete explanation (homology thinking).

To understand the virtues of functionalism, let us first take a look at the points where this view meets homology thinking. Classical functionalism is often presented in contrast to type-identity theories. Type-identity theories claim that types of mental states are identical with physical types, usually cashed out in terms of types of neurophysiological activity. Both functionalism and homology thinking reject type-identity theories, and they do so for more or less the same reasons: type-identity theories are too restrictive as to the possible variability realizers of a given mental state can have. Since the mechanisms underlying a type of mental state are typically varied, type-identity theories have trouble attributing mental states to non-human organisms.

At this stage, both functionalism and homology thinking views allow for some degree of multiple realization. However, the difference between the two lies precisely in how much variability they admit before splitting a given mental or emotional category. Put more clearly, while these views admit that a specific mental type or emotion category can have various underlying mechanisms, they differ on how much these underlying mechanisms can differ before we decide that the putative mental or emotional category should be split in two. For example, let us suppose that an emotion category  $E$  has as underlying mechanisms  $M_1$  and  $M_2$ . How different can  $M_1$

and  $M_2$  be in order for us to consider them two mechanisms of the same emotion type? For functionalists, there is no restriction on the physical level, that is, the two mechanisms may differ completely in terms of their material composition, as long as they instantiate the same abstract process describing E. For defenders of homology thinking however, the two mechanisms must share a specific relation, namely, they must be homologous mechanisms.

What does it mean for two mechanisms to be homologous? As a general definition, two traits are homologous if they are produced by a common ancestral origin. This is the definition adopted for example by Griffiths (1997). However, current debates in philosophy of biology distinguish two types of homology concepts. These are what Griffiths (2006) calls a *taxic* or *Darwinian* concept, and a *developmental* or *biological* concept. According to Griffiths:

‘Taxic’ or ‘Darwinian’ approaches to homology treat characters in two or more organisms as homologous if they are descended from a single character in an ancestral organism. ‘Developmental’ or ‘biological’ (Wagner 1989) approaches, however, treat characters as homologous if the preferred theory of how organisms develop identifies them as instances of the same developmental phenomenon at some level of analysis [...] (2006, p. 9).

According to these two concepts then, two traits are homologous either in case they are descended from a single character, or in case the preferred theory of how these organisms develop the trait treats them as instances of the same phenomenon. If we adopt this definition of homologous traits, then two underlying mechanisms for a given emotion category can differ as long as they are descended from the same character (taxic concept) or if theories of how these mechanisms develop consider them the same phenomenon (developmental concept).

This understanding of homologies, particularly the taxic concept, seems to be in line with the general definition found in Griffiths (1997). On such a view, then, multiple realization is allowed as long as we are still on the same evolutionary line. If we are to apply this strategy to the question of which emotion concepts can be applied to non-human animals, then we can only apply emotion concepts for which we can find homologous underlying mechanisms, i.e., mechanisms that have evolved from a common origin.

Griffiths thinks that homology thinking is preferable to functionalism for two reasons we have already discussed. One is that intuitively, people apply emotion concepts to organisms beyond our own evolutionary line only metaphorically, but they are rather timid when it comes to committing themselves to the view that these organisms, for example octopuses, genuinely have the emotions attributed to them. The second is that homology thinking provides a richer explanation as to why non-human animals have or lack emotions, namely, evolutionary history. In spite of these arguments, I believe that functionalism is a better position than homology thinking.

Concerning the first argument, while I am sympathetic to taking folk attribution of emotion concepts seriously, I do not think we should be completely bound by such conditions. It is important that science can progress beyond our folk understanding, and that we can discover facts that can be unintuitive. In this case, it is plausible that people are shy to commit themselves to the view that organisms outside our ancestry line have emotions due to intuitions that can turn out to be false. If we have reasons to accept that behaviors such as discharging ink in octopuses count as fear behaviors—and this might be the case, given that it functionally mirrors fleeing or fighting behavior—then we should not reject this finding because it may be unintuitive. Rather, we can expect folk intuitions to shift due to scientific findings, and we can expect a future where we do not find such attributions puzzling.

More critical, however, is the second argument. If explanations based on homologies are indeed richer than functional explanations, then there is good scientific reason to reject functionalism in lieu of the former. Yet, there are two ways to resist this line of argument. On the one hand, functionalism is compatible with homology thinking, while being less ontologically restricted. It is true that functionalism is a view that is ontologically neutral regarding the realizers of a given mental state. This does not imply however that it is incompatible with theorizing about the realizers. If biologists discover that a functionally identified behavior is present in other species, but that these behaviors have homologous underlying mechanisms, this is perfectly consistent with the functionalist's claim. This is because functionalism can be understood as claiming that functional identification is *sufficient* to apply

a mental concept, but not that it must be the only way to investigate mental states. Put differently, functionalism does not claim that we should completely ignore facts about the realizers, but only that these facts are not necessary to apply a given mental concept to a system.

So far, this argument only means that functionalism and homology thinking can coexist, but defenders of homology thinking can still press claiming that homology-based identification contains more information. Such line of argument in favor of homology thinking has been proposed by Clark (2010, 2013a, 2013b). Following Griffiths, Clark also thinks that “functional categories are by comparison relatively shallow, capturing only the commonalities in terms of which they are defined” (2010, p. 76). Using the cases of shame (2010) and pride (2013a), Clark has claimed that functionalism is in dire straits, as homologies provide a better framework to understand what is common between emotional behavior across species.

Clark’s argument is based on a slightly different understanding of homologies than Griffiths. While Griffiths understands homologies in general as mechanisms with a common ancestry, Clark follows Ereshefsky’s (2007) account of homologies designed to analyze psychological categories. On this account, two traits are homologous in case they satisfy three criteria:

Position: the two traits share a spatio-temporal position or functional role within a more general pattern of organization.

Special quality: the more specialized the traits in question are, the less likely it is that they evolved independently.

Continuity: there is a continuous line of evolution connecting a simpler, more primitive state to a more complex, derived state of the traits. (adapted from Ereshefsky, 2007)

Clark also adds a fourth criterion not found in Ereshefsky (2007), namely:

Shared developmental pathways: the use of common developmental mechanisms or constraints establishes the traits. (adapted from Clark, 2013a).

Let us consider Clark’s analysis of pride as a homology. According to Clark, dominance pride has various commonalities in different species and across cultures. One such commonality is bodily posture, a posture dis-

playing tilting of the head, expansion of the chest, arms raised, among other characteristics. This would suggest that there is a similar spatio-temporal position or functional role between pride displays in other species and across human groups. Clark also presents a number of neurophysiological mechanisms showing common mechanisms, and he argues that there is reason to believe that pride has evolved continuously from pride displays in non-human animals such as macaques. Hence, it seems that pride can be understood as a homologous trait that evolved other species.

While I find Clark's argument and analysis very informative, it is difficult to see how it is incompatible with functionalist accounts. First, recall that functionalism is ontologically neutral with regards to realizers of a mental state, although it is compatible with research on said realizers. It is not that realizers are not interesting, but that they are additional information that is not necessary to apply a given mental concept to an organism. Consequently, findings such as the ones presented by Clark are not inconsistent with a functionalist approach, since we can argue that the application of a pride concept via functional criteria is possible independently of these findings while accepting these findings as true.

Second, we should consider the drawback of limiting identification of mental states to homologous mechanisms. While it is interesting to see that two mechanisms underlying a given mental state are homologous, it may come to our hindrance if it turns out that there are organisms which act exactly in the ways that our mental concepts suggest, but that do so based on different mechanisms. If an organism acts in ways that show that they escape quickly when they perceive an event that is dangerous to them, we should be open to the idea that they are afraid, even before we investigate whether the mechanisms whereby they realize this process are the same as in our own biological makeup. Moreover, the application of the emotion concept of "fear" is what makes it possible and interesting to formulate hypotheses regarding similarities in realizers, and thus there is reason to think that functional criteria are logically prior to considerations concerning realizers. Additionally, if we discover that their mechanisms are largely different from ours, this may be a fascinating finding in its own right, the fact that the processes we identify as "fear" are solved in different ways in different or-

ganisms. As a result, requiring that the mechanisms are homologous would be to miss out on a number of promising empirical hypotheses. Functionalism thus opens the door to more empirical hypotheses concerning which emotions can be said to be possessed by which organisms, while being compatible with information provided by homology thinking.

Lastly, if we take a closer look at Ereshefsky's criteria of homologous classification, it seems that the type of homologies Clark is after are ultimately functional criteria. It is important to note, first, that Ereshefsky's criteria are designed to spell out the concept of a *behavioral* homology. Ereshefsky is interested in explaining the evolution of behaviors such as web building in spiders. The kind of explanation Ereshefsky wants is indeed one that can make use of information about realizers, as commonalities at the level of realizers can help make a case for homologous classification of these behavioral traits.

However, Ereshefsky does not require that common mechanisms are strictly necessary to understand two behavioral traits as homologous. This is why Ereshefsky's criteria for behavioral homologies accept that two traits can be homologous if they share a common functional role within a system (as per the first criterion). As a result, Ereshefsky's criteria for homologies can be understood as functional criteria, as they admit that we can identify a behavior functionally and look for functional homologies. In his view, the difference between homology thinking and functionalism is mostly one of emphasis, but they are largely complementary rather than opposing views. In his words:

Functionalism and homology thinking offer different ways of investigating psychological categories. Functionalism focuses on the function or design problem of a psychological category; homology thinking examines the phylogeny and ontogeny of a psychological category. (Ereshefsky, 2007, p. 667)

In spite of the acceptance that functionalism and homology thinking may be compatible, both Clark and Ereshefsky, just like Griffiths, insist that homology thinking is more informative than functionalism. Again, they argue that this is because homology thinking leads to richer explanations. Yet, we need to take these arguments with a grain of salt. As I explained above,

there are reasons to think that even if homology thinking is interesting, it may be overly committal to specificity at the level of realizers which may turn out excluding other interesting empirical hypotheses worth exploring, hypotheses as to how organisms can solve the same problem in mechanically different ways (i.e., differing only in terms of the mechanics of the realizers but being similar in their functional description). This is true as long as homologies are limited to specific realizers. If defenders of homology thinking relax this criterion, as Ereshefsky seems to do by accepting that two traits can be homologous by sharing a functional role, then both functionalist and homology thinking views can co-exist, and both can be similarly informative depending on concrete empirical findings. At this point, homology thinking can be understood as a version of functionalism, one that makes emphasis on evolution, but that still cashes out psychological categories by their functional role. This is why it is not the case that homology thinking is opposed to functionalism, but that, in Ereshefsky's words, "applying insights from ethology to functionalist and adaptationist accounts of psychology strengthens those accounts" (Ereshefsky, 2007, p. 660).

#### **4. A Framework to Attribute Emotions to Non-Human Animals**

In this article, I began by considering how we attribute emotions to non-human animals. Based on findings from psychology, we noticed that people do attribute emotions to animals, but this attribution is not homogeneous. In a nutshell, not all animals are thought of as having all emotions. First, animals that are more familiar to us, and those with more psychological complexity, are thought to have more complex emotions. Second, only some emotions are projected onto non-human animals, particularly emotions such as fear, sadness or anger, in contrast to more socially complex emotions such as shame or embarrassment. This led us to formulate two questions concerning non-human animal emotions. The first, which I called the Species Question, concerns the species to which we can attribute emotions. The second, which I called the Emotions Question, concerns which emotion concepts we can apply to these organisms.

After raising these two questions, I considered two ways in which we can account for emotions in non-human animals, namely, functionalism and homology thinking. For functionalists, we can apply emotion concepts to animals in case the relevant causal patterns obtain, that is, in case animals behave in ways that would lead us to think of them as having that emotion, regardless of whether their neural makeup is similar to ours. For example, we can say that an animal is afraid in case it can be described as evaluating a stimulus as dangerous and developing a fleeing or fighting strategy (including not only behavioral patterns, but also possible intermediate cognitive states). For defenders of homology thinking, the same attribution holds as long as there is some homologous structure between the organism and human emotions. Between these positions, I argued that functionalism was a better choice, since constraints about realizers lead to more fruitful empirical hypotheses which should be considered before we restrict our vocabulary to organisms with similar physical makeup. In other words, functionalism is more empirically productive than homology thinking.

If my arguments work, this has important consequences for answering both the Species Question and the Emotions Question, as well as for our general theories of emotions. To conclude, I shall go over the most significant consequences of the view and some questions that remain up for debate.

First, let us draw some conclusions concerning the Species Question and the Emotions Question. Regarding the Species Question, a functionalist account leaves open the possibility that any organism can be thought of as having emotions, as long as their behavior and the appropriate mediating mental states can be meaningfully described as emotional. This means that there is no a priori restriction as to which species have emotions. I believe this to be an important advantage over homology thinking, as it does not impose limitations on empirical research from the outset. If organisms such as octopuses can be described as afraid, angry, sad, etc., then we have good reason to think of them as genuinely having those emotions. This is not to say that all animals have all emotions, but rather that conceptually speaking, there is no reason to limit our vocabulary before we observe an organism's behavior.

This leads us to a second relevant consequence of the view I am defending, namely, that behavioral descriptions are taken to be logically prior to considerations about realizers. As I explained before, concerns about realizers, while pertinent in their own right, constitute questions and hypotheses that can be made meaningful only after we describe an organism's behavior as emotional. Recall that functionalism does not hold that we cannot delve into details about realizers, but only that these questions are to be raised only after we apply a mental concept based on functional criteria. This allows us to preserve the benefits of researching neural and physiological mechanisms underlying emotions, while leaving it open that there are mechanisms different from our own which realize the same functional descriptions.

Such a claim, in turn, has implications concerning the Emotions Question. On the one hand, my claim implies that since functional descriptions are prior to concerns about realizers, we must first determine how we can describe each emotional pattern in order to investigate whether these patterns are present in non-human animals or not. In other words, this means that the Emotions Question is prior to the Species Question. Only after we have a functional description for a given emotion category can we ask whether that description applies to a given species. Consequently, research on emotions should prioritize characterizing each emotion's functional pattern, developing abstract but tractable accounts of how our emotion categories work.

This raises the question of how we can obtain such descriptions, as it seems that they are paramount to formulate further hypotheses concerning emotions in non-human animals. In my view, we can do this based, but not constrained by, folk emotion concepts. As I have claimed elsewhere (Loaiza 2021a, 2021b), we can examine our vernacular concepts and extract functional descriptions which can then be empirically confirmed and refined. This is not to say that we can stipulate definitions of emotion, but rather that we can formulate hypotheses based on our current practices. As practices shift, our science of emotion will shift accordingly, even influencing changes in folk concepts. Implementing this strategy in the case of non-human animals

means taking folk attribution seriously, but taking it as the grounds on which we can begin scientific theorizing.

Lastly, let us consider some consequences for general theories of emotions. On the one hand, the appeal of functionalism when it comes to non-human animal emotions should inform debates between major theories of emotions. The debate between basic emotion theories (BET), appraisal theories, and psychological constructionism is one that is concerned with such discussions. For example, BET has traditionally appealed to emotions in non-human animals as part of what makes an evolutionary account important, while constructionists have discussed non-human animal emotions with some degree of suspicion (see e.g., Fugate, 2015). If we can make sense of emotions in non-human animals in a functionalist framework, this may provide arguments to decide from different theories. Such a case would be developing a functionalist account of basic emotions, in contrast to locutionist accounts such as Ekman's (1992; Ekman & Cordaro, 2011), Izard's (1992, 2007) or Panksepp's (1998, 2011). This may give BET new avenues that can meet the challenges posed by constructionists who emphasize neural and physiological variability.

To sum up, I have argued that functionalism is a promising alternative to other approaches, in this case concerning how to account for emotions in non-human animals. Even though many of these questions must be answered empirically, it is vital that we develop conceptual frameworks that can make sense of predictions and that can allow the formulation of promising hypotheses. It will remain open however whether these approaches are ultimately successful, as only future research will decide their merits. For the time being, if I am correct, this is a research avenue worth pursuing.

*Recebido em 30/12/2021*

*Aprovado em 28/02/2022*

## REFERÊNCIAS

Adolphs, R. (2016). How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences. *Social Cognitive and Affective Neuroscience*, 49(4), 24–31. <https://doi.org/10.1093/scan/nsw153>

Adolphs, R., & Andler, D. (2018). Investigating Emotions as Functional States Distinct From Feelings. *Emotion Review*, 10(3), 191–201. <https://doi.org/10/gfn9rs>

Clark, J. A. (2010). Relations of homology between higher cognitive emotions and basic emotions. *Biology & Philosophy*, 25(1), 75–94. <https://doi.org/10.1007/s10539-009-9170-1>

Clark, J. A. (2013a). Integrating basic and higher-cognitive emotions within a common evolutionary framework: Lessons from the transformation of primate dominance into human pride. *Philosophical Psychology*, 26(3), 437–460. <https://doi.org/10.1080/09515089.2012.659168>

Clark, J. A. (2013b). Intersections between development and evolution in the classification of emotions. *Developmental Psychobiology*, 55(1), 67–75. <https://doi.org/10.1002/dev.21063>

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3), 169–200. <https://doi.org/10.1080/02699939208411068>

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, 3(4), 364–370. <https://doi.org/10.1177/1754073911410740>

Ereshefsky, M. (2007). Psychological categories as homologies: Lessons from ethology. *Biology & Philosophy*, 22(5), 659–674. <https://doi.org/10.1007/s10539-007-9091-9>

Fugate, J. M. B. (2015). Evolutionary Constraints and Cognitive Mechanisms in the Construction of an Emotion. In L. F. Barrett & J. A. Russell (Eds.), *The Psychological Construction of Emotion* (pp. 399–426). The Guilford Press.

Griffiths, P. E. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. University of Chicago Press.

Griffiths, P. E. (2006). Function, Homology, and Character Individuation. *Philosophy of Science*. <https://doi.org/10.1086/510172>

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, 99(3), 561–565.

Izard, C. E. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, 2(3), 260–280. <https://doi.org/10.1111/j.1745-6916.2007.00044.x>

LeDoux, J. E. (2013). The slippery slope of fear. *Trends in Cognitive Sciences*, 17(4), 155–156. <https://doi.org/10.1016/j.tics.2013.02.004>

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(03), 121–143. <https://doi.org/10.1017/S0140525X11000446>

Loaiza, J. R. (2021a). Emotions and the problem of variability. *Review of Philosophy and Psychology*, 12(2), 329–351. <https://doi.org/10.1007/s13164-020-00492-8>

- Loaiza, J. R. (2021b). Functionalism and the emotions. *The British Journal for the Philosophy of Science*. <https://doi.org/10.1086/715207>
- Morris, P., Doe, C., & Godsell, E. (2008). Secondary emotions in non-primate species? Behavioural reports and subjective claims by animal owners. *Cognition & Emotion*, 22(1), 3–20. <https://doi.org/10.1080/02699930701273716>
- Morris, P., Lesley, S., & Knight, S. (2012). Belief in Animal Mind: Does Familiarity with Animals Influence Beliefs about Animal Emotions? *Society & Animals*, 20(3), 211–224. <https://doi.org/10.1163/15685306-12341234>
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford Univ. Press.
- Panksepp, J. (2011). The basic emotional circuits of mammalian brains: Do animals have affective lives? *Neuroscience and Biobehavioral Reviews*, 35(9), 1791–1804. <https://doi.org/10.1016/j.neubiorev.2011.08.003>
- Piccinini, G. (2004). Functionalism, computationalism, and mental states. *Studies in History and Philosophy of Science Part A*, 35(4), 811–833. <https://doi.org/10.1016/j.shpsa.2004.02.003>
- Piccinini, G. (2010). The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism. *Philosophy and Phenomenological Research*, 81(2), 269–311. <https://doi.org/10.1111/j.1933-1592.2010.00356.x>
- Place, U. T. (1956). Is Consciousness a Brain Process? *British Journal of Psychology*, 47(1), 44–50. <https://doi.org/10.1111/j.2044-8295.1956.tb00560.x>
- Putnam, H. (1960/1997). Mind and machines. In *Mind, language and reality* (pp. 362–385). Cambridge University Press.
- Putnam, H. (1967/1997). The nature of mental states. In *Mind, language and reality* (pp. 429–440). Cambridge University Press.
- Smart, J. J. C. (1959). Sensations and Brain Processes. *The Philosophical Review*, 68(2), 141–156. JSTOR. <https://doi.org/10.2307/2182164>
- Spunt, R. P., Ellsworth, E., & Adolphs, R. (2017). The neural basis of understanding the expression of the emotions in man and animals. *Social Cognitive and Affective Neuroscience*, 12(1), 95–105. <https://doi.org/10/gf26mm>
- Wilkins, A. M., McCrae, L. S., & McBride, E. A. (2015). Factors affecting the Human Attribution of Emotions toward Animals. *Anthrozoös*, 28(3), 357–369. <https://doi.org/10.1080/08927936.2015.1052270>

