

Mazor, M., Brown, S., Ciaunica, A., Demertzi, A., Fahrenfort, J. J., Faivre, N., Francken, J., Lamy, D., Lenggenhager, B., Moutoussis, M., Nizzi, M., Salomon, R., Soto, D., Stein, T. & Lubianiker, N. The scientific study of consciousness cannot, and should not, be morally neutral

In press in *Perspectives on Psychological Science*

# The scientific study of consciousness cannot, and should not, be morally neutral

Matan Mazor<sup>1,2</sup>, Simon Brown<sup>3</sup>, Anna Ciaunica<sup>4</sup>, Athena Demertzi<sup>5,6</sup>, Johannes Fahrenfort<sup>7,8</sup>, Nathan Faivre<sup>9,10</sup>, Jolien C. Francken<sup>11</sup>, Dominique Lamy<sup>12,13</sup>, Bigna Lenggenhager<sup>14</sup>, Michael Moutoussis<sup>2,15</sup>, Marie-Christine Nizzi<sup>16,17,18</sup>, Roy Salomon<sup>19</sup>, David Soto<sup>20,21</sup>, Timo Stein<sup>7</sup> & Nitzan Lubianiker<sup>12,13</sup>

<sup>1</sup>Department of Psychological Sciences, Birkbeck, University of London

<sup>2</sup>Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, London, UK

<sup>3</sup>Department of Philosophy, Johns Hopkins University, Baltimore, USA

<sup>4</sup>Centre for Philosophy of Science, University of Lisbon

<sup>5</sup>Physiology of Cognition, GIGA Consciousness Research Unit, Université de Liège, Liège, Belgium

<sup>6</sup>Fund for Scientific Research - FNRS, Bruxelles, Belgium

<sup>7</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

<sup>8</sup>Department of Experimental and Applied Psychology, Vrije Universiteit, Amsterdam, The Netherlands

<sup>9</sup>Centre for Neuroprosthetics and Brain Mind Institute, Faculty of Life Sciences, Swiss Federal Institute of Technology (EPFL), Geneva, Switzerland

<sup>10</sup>University Grenoble Alpes, University Savoie Mont Blanc, CNRS, LPNC, Grenoble, France

<sup>11</sup>Faculty of Philosophy, Theology and Religious Studies, Radboud University, Nijmegen, The Netherlands

<sup>12</sup>Sagol Brain Institute, Tel-Aviv Medical Centre, Tel-Aviv, Israel

<sup>13</sup>School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel

<sup>14</sup>Department of Psychology, University of Zurich, Zurich, Switzerland

<sup>15</sup>Max Planck-UCL Centre for Computational Psychiatry and Ageing Research, University College London, London, UK

<sup>16</sup>UCLA, Semel Institute for Neuroscience and Human Behaviour, Department of Psychiatry & Biobehavioral Sciences, David Geffen School of Medicine, Los Angeles, CA, United States.

<sup>17</sup>Dartmouth College, Cognitive Science Program, Hanover, NH, United States.

<sup>18</sup>Chapman University, Institute for Interdisciplinary Brain and Behavioural Sciences, Orange, CA, United States.

<sup>19</sup>Gonda Multidisciplinary Brain Research Centre, Bar-Ilan University, Tel Aviv, Israel

<sup>20</sup>Basque Centre on Cognition, Brain and Language, San Sebastian, Spain

<sup>21</sup>Ikerbasque, Basque Foundation for Science, Bilbao, Spain

---

Authors appear in Alphabetical order, except for M. Mazor and N. Lubianiker who co-wrote the original version of this article. Correspondence should be addressed to Matan Mazor, [mtnmzor@gmail.com](mailto:mtnmzor@gmail.com)

A target question for the scientific study of consciousness is how dimensions of consciousness, such as the ability to feel pain and pleasure or reflect on one's own experience, vary in different states and animal species. Considering the tight link between consciousness and moral status, answers to these questions have implications for law and ethics. Here we point out that given this link, the scientific community studying consciousness may face implicit pressure to carry out certain research programmes or interpret results in ways that justify current norms rather than challenge them. We show that since consciousness largely determines moral status, the use of non-human animals in the scientific study of consciousness introduces a direct conflict between scientific relevance and ethics – the more scientifically valuable an animal model is for studying consciousness, the more difficult it becomes to ethically justify compromises to its well-being for consciousness research. Lastly, in light of these considerations, we call for a discussion of the immediate ethical corollaries of the body of knowledge that has accumulated, and for a more explicit consideration of the role of ideology and ethics in the scientific study of consciousness.

# Moral status and consciousness are coupled

Philosophers and theologians have long debated which beings deserve moral consideration, and to what extent. Many see *moral status*, or the degree to which an entity deserves moral consideration, as dependent on certain mental capacities, and specifically on consciousness, broadly defined here as perceptual, cognitive, and emotional states that are experienced by a subject<sup>1</sup>. In ancient traditions, moral obligations towards non-human animals often rested on conscious aspects of experience, such as the capacity to suffer. For instance, according to Dhārmic religions, the principle of *ahimsā* (non-violence) is applied to animals based on their capacity to suffer and to experience desire (Carpenter, 2018; Finnigan, 2017), and the moral priority of humans rests on unique characteristics of the human soul (Jena, 2019). In Greek natural philosophy, the Pythagorean school held that animals were capable of suffering (and were also capable of a degree of rationality), which in turn entailed a moral obligation to minimise animal suffering (Campbell, Gordon Lindsay, 2014). In contrast, Aristotelians, who dominated mediaeval philosophy, took the limited rationality of animals as evidence that they did not have a mind but only ‘locomotive souls’, and hence had no moral status (Allen & Trestman, 2020; Sorabji, 2018). This dominant view coexisted with a privileged moral standing for some working animals, such as hunting dogs, which were assumed to have rich mental lives (Crane, 2015). The tight link between consciousness and moral status was not limited to non-human animals: Aristotle justified slavery by alluding to some non-Greeks as ‘not having reason’ and ‘live by perception alone, like non-human animals’ (Heath, 2008).

In post-medieval Western thought, different philosophical traditions identify different mental capacities as the determining factor for moral status. For Immanuel Kant, it was autonomy (Kant, 1785/2002). In contrast, utilitarian philosophers identified the origin of moral

---

<sup>1</sup>Our focus in this paper is on consciousness as a term used in the scientific literature, in contrast to consciousness as a mental state or capacity. As such, this definition is intended to be agnostic with regard to different theoretical approaches to consciousness in the scientific literature.

status in the capacity to experience suffering (Bentham, 1789), or pain and pleasure more generally (Mill, 1863/2015). More recent debates have seen broad agreement that moral status rests, at least in part, on mental capacities (Carruthers, 2019; Danaher, 2020; Levy, 2014; Shepherd, 2018). Although there is debate about the relative weight of different capacities (be it cognitive complexity, functional and representational aspects of the mind, or felt experience), most have included consciousness as amongst the most important. Even those denying that consciousness itself is morally valuable commonly acknowledge that specific qualities of conscious experience are likely to contribute to intrinsic value (Glover, 2006; Lee, 2019). Further supporting this dependence of ethics on consciousness, the philosophical view that posits that consciousness is merely an illusion has been challenged by philosophers for its potentially dangerous implications for ethics and society (Kammerer, 2019; Strawson, 2018).

Due to this tight link between consciousness and ethics, beliefs about the mind often mirror cultural practices and norms. If moral status depends on properties of the mind, differences in moral status between individuals can be justified based on presumed differences in consciousness (see next Section for a discussion of what falls under this umbrella term in scientific writing). For example, French psychologist Ribot described savages as not capable of sustained attention, together with “vagabonds, thieves, and prostitutes” (Daston, Lorraine & Galison, Peter, 2010; Ribot, 1889). More recently, white children and adults attributed reduced emotions to black compared to white people in a lab experiment, and this was especially the case for emotions that are perceived as ‘uniquely human’ (Costello & Hodson, 2014). People also ascribe lower levels of consciousness to individuals who were pushed to the margins of society: in one study, participants attributed lower levels of intention and cognition to an individual if they learned he had lost his job and could not afford to pay rent and bills (Kozak et al., 2006).

These effects are not restricted to the attribution of mental properties to human beings. For example, participants attributed reduced mental properties to lambs and sheep after being reminded that they will later be used as food (Bastian et al., 2012). Bastian and Loughnan

(Bastian & Loughnan, 2016) proposed that the denial of mind to certain animals resolves the cognitive dissonance between the practice of meat-eating and the belief that animals are sentient beings that are capable of suffering. More generally and related to our focus here, the denial of conscious experience can be used by individuals and societies to justify pre-existing moral attitudes and practices.

## Scientific attribution of consciousness

Cognitive science, and more specifically the scientific study of consciousness, is concerned with the study of the mind via behavioural and physiological measurement. The term 'consciousness' is notoriously hard to define (Giacino et al., 2014), and is used in the scientific literature to refer to a wide range of mental states and their corresponding behavioural and neural markers. For example, in a recent review of consciousness in the animal kingdom, Birch, Schnell and Clayton (2020) list as dimensions of consciousness the capacity for self awareness, unity of experience across time and senses, perceptually rich experience, and the capacity for experiencing affective states and emotions. Others have suggested that consciousness is tightly linked to attention (Graziano & Webb, 2015; Posner, 1994), learning (Birch, Ginsburg, et al., 2020), and self-referential thinking (Hofstadter, 2007; Rosenthal, 2005). In clinical settings, consciousness is additionally linked to wakefulness and awareness (Giacino et al., 2014). Developments in experimental design and neuroimaging methods now bring us closer than ever to a systematic investigation of these states and capacities, as well as their neural correlates, not only in adult humans, who are capable of reporting their internal states, but also in non-communicating patients, preverbal infants, and non-human animals. To date, all suggested non-verbal markers of conscious states that have been identified in humans, behavioural and neural, have also been observed in other animals.

Invasive experiments on corvid birds (Nieder et al., 2020), rodents (Sachidhanandam et al., 2013), and non-human primates (De Lafuente & Romo, 2006; Leopold & Logothetis, 1996; Vugt et al., 2018) reveal percept-yoked neural activation patterns similar to what is typically interpreted as neural correlates of visual consciousness in humans (Koch et al., 2016). Candidate measures of self-awareness are also observed in other animal species: the capacity to identify one's reflection in the mirror as 'self' was reported in elephants (Plotnik et al., 2006), birds (Prior & Schwarz, 2008), and arguably also in some fish (Kohda et al., 2019), among other animals. Other measures are also used to document self-awareness in non-human animals with better experimental control (Wada et al., 2016). Behavioural and brain experiments provide converging evidence for something akin to episodic memory and future thinking in corvids and rodents (Clayton & Dickinson, 1998; Panoz-Brown et al., 2018). Rats were able to monitor the accuracy of their decisions, a capacity that in humans is associated with metacognition and higher-order thinking (Yuki & Okanoya, 2017). Insects were able to integrate information across different sensory modalities (Solvi et al., 2020) and form egocentric representations of the world (Barron & Klein, 2016). Critically, this list is not meant to convince readers that elephants, rodents, birds or insects are conscious (indeed, the validity of some of these markers is disputed, for example see Gallup & Anderson, 2018, for a critical review of the mirror test literature). Instead, it is meant to show that as of April 2022, no single non-verbal behavioural or neural marker of consciousness has been shown to be uniquely human.

Inherent to the scientific study of conscious experience is a leap from observable behaviour and physiological processes to conjectures about private conscious experiences. While for humans, few researchers question drawing associations between subjective conscious states and their measured correlates, many researchers raise the question "but are they really conscious?" when it comes to non-human animals. Examples include debates about the interpretation of the mirror test (Kohda et al., 2019), and about the true nature of metacognitive behaviour (Jozefowicz et al., 2009), episodic memory (Suddendorf & Corballis, 2007; Tulving,

2005), and emotional experiences (LeDoux, 2021) in non-human animals. These debates over the presence of 'true experience' are telling. Scientists come with strong priors regarding the presence or absence of consciousness in human and non-human agents. These priors can result in liberal biases to identify consciousness more readily in some cases, and conservative biases to be more sceptical in other cases. While some of these priors are based on previous literature, others may be guided by pre-scientific tendencies to ascribe agentic and experiential states based on behavioural (Heider & Simmel, 1944) and physical (Sherman & Haidt, 2011) cues. Critically, as we show in the previous section, pre-scientific priors are susceptible to influences from moral outlook and lifestyle. This way, beliefs about consciousness interact with ethics in a two-way fashion: they shape ethics, and they are constrained by it.

For example, a scientist may believe that 1) beings are conscious if and only if they can integrate information across different senses, 2) consciousness is a determinant of moral status, and 3) insects deserve no moral consideration. This scientist will need to revise at least some of their beliefs if they learn that bees can integrate information across different senses (Solvi et al., 2020; see Fig. 1 for a general scheme). The scientist may hold to their specific moral intuitions that bees do not possess moral status, and revise their global beliefs about a link between consciousness and moral status. For example, evidence for belief-desire psychology in bees has led philosopher Peter Carruthers to question basic axioms of utilitarian ethics on the grounds that bees cannot possibly be subjects of moral concern (Carruthers, 2007). Alternatively, they may update their ethical norms regarding the treatment of some entities in light of scientific evidence. An example of this second option is the change to the legal status of cephalopods (including octopuses, squids, and cuttlefish) in European law in light of evidence for a capacity to "experience pain, suffering, distress or lasting harm" in these marine creatures (AHAW, 2005). The accumulation of scientific evidence informed a decision to provide cephalopods with the same level of protection as vertebrates in scientific experiments (Smith et al., 2013). A third option is to question the validity of the theory of consciousness at hand. For example, the



scientist may decide that cross-modal sensory integration cannot be a sufficient condition for consciousness if it is evident in bees. Critically, all three courses of action involve an interaction between ethics and scientific practice.

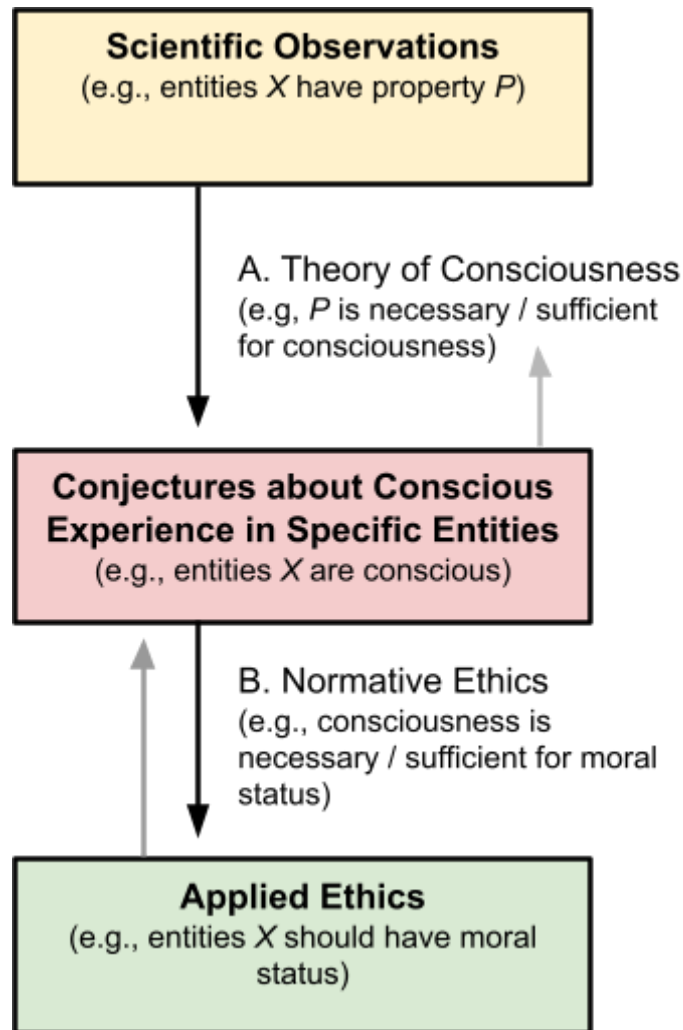


Fig. 1: A schematic description of the link between scientific observations, conjectures about conscious experience, and applied ethics. Descending black arrows: scientific observations are translated into conjectures about conscious experience in specific entities via a theory of consciousness. These conjectures then go on to affect applied ethics via general beliefs about the relation between consciousness and ethics. Ascending grey arrows: applied ethics shapes conjectures about conscious experience: entities that are not currently held to have moral status are less likely to be perceived as conscious. This in turn puts pressure on theories of consciousness to align with current intuitions about the attribution of consciousness to specific entities.

Some readers may worry that accepting that the scientific study of consciousness cannot, and should not, be morally neutral might jeopardise the field's objectivity, and even its status as

science. This need not be so. Scientists being open about what they take to be the moral implications of their research, explicitly considering whether their paradigms are ethical rather than merely currently legally accepted, arguably does not undermine, but enhances, scientific trustworthiness and legitimacy (Kitcher, 2003; Longino, 2020; Oreskes, 2021). Historians and sociologists have documented cases where a veneer of supposed value-free objectivity and refusal to discuss the role of values and interests has sometimes served to mask their influence. These include biases leading some researchers to be too ready to accept theories about gender on the basis of weak evidence in sociobiology (Longino, 2020), or economic interests making some scientists unduly reticent to accept well-established results in climate science (Oreskes, & Conway, 2011). In a similar way, institutional or personal biases with respect to the treatment of non-human animals could unduly influence consciousness science, especially if not addressed and discussed.

## The animal-models-of-consciousness paradox

An instance where the scientific community has failed to acknowledge the intimate link between consciousness and ethics is in the use of animal models of consciousness. Our focus here is on the use of animals that are assumed to be conscious as an opportunity to probe the underlying mechanisms of consciousness in ways that would not be ethically acceptable with human subjects. In such studies, animals are often captive, deprived of basic needs, and undergo invasive procedures. At the same time, for these animals to be appropriate models for the study of consciousness, it has to be assumed that they are conscious. As conscious capacities play a pivotal role in the attribution of moral status to animals, in these experiments scientific validity and moral justification are in direct conflict. This conflict is particularly acute in the study of consciousness and subjective experience: that an animal is an adequate model for the study of consciousness makes it more likely to be capable of experiencing rich phenomenal states,

self-awareness or suffering, and its life deserving the appropriate protection, much more than being an appropriate model for the study of the immune system does.

In a recent study of the Neural Correlates of Consciousness, researchers contrasted brain activation in awake, sleeping, and anaesthetised macaque monkeys (Redinbaugh et al., 2020). For this study, two monkeys were kept in captivity, implanted with brain electrodes, and were immobilised by sticking rods in a head implant during electrophysiological recordings. In another study from 2021, a behavioural measure of conscious awareness was reported in four caged rhesus monkeys (Ben-Haim et al., 2021). Scientists surgically implanted subjects with a metal extension to their skull for the purpose of restraining movement during experimental sessions, and restricted subjects' access to water at testing so that they are motivated to participate in the task for juice droplets. In a study from 2019 on the neural basis of introspection, researchers abolished parts of the prefrontal cortex of six caged macaque monkeys, which were killed at the end of the study (Kwok et al., 2019). In another study published in *Science* in 2020 (Nieder et al., 2020), a neural correlate of sensory consciousness was demonstrated in the brains of two male crows by implanting electrodes in their brains. These are mere examples of typical research practice in the field of invasive electrophysiology, conforming with current ethical guidelines in place at a national level and which are commonplace in many fields of study. Yet, common to these studies is that their scientific relevance rests on the animal being conscious, while their ethical justification rests on the animal not deserving the same protection from suffering as a human subject.

Animal models of psychopathologies exhibit a similar paradox, specific to one particular dimension of conscious experience, namely the capacity to suffer. To be clinically relevant, model animals must show behaviours that in humans are interpreted as indicating a mental illness or distress. For example, in research about depression, non-human animals can be led to express behavioural markers such as passivity and anhedonia when exposed to long periods of social (Hollis & Kabbaj, 2014) or physiological stress (Willner, 2017), or to stress inducers that

cannot be avoided (Maier, 1984). The ethical justification for such experiments is in their clinical potential for suffering patients. However, for most psychiatric disorders such as major depression, schizophrenia, and post-traumatic stress disorder, conscious suffering is central to, even a defining feature of, the disorder. For example, according to the Diagnostic and Statistical Manual of Mental Disorders, a diagnosis of depression depends on the patient having a depressed mood or a loss of interest or pleasure in activities. Without at least one of these experiential symptoms, they should not be diagnosed with depression, even if they walk slowly, lose their appetite, and show signs of tiredness. Do animal models of psychiatric disorders induce the same kind of experiential suffering in non-human animals? An affirmative answer to this question would call into question the morality of such experiments, and a negative answer their scientific validity and clinical utility<sup>2</sup>.

We do not mean to deny that there can be huge benefits from research on animals. In the case of consciousness research, a great deal has been learned from studying animal models, including knowledge which has been applied to alleviate the suffering of humans and other animals (for example, fear conditioning experiments on rodents led to the development of exposure therapy for anxiety disorders; Kirlic et al., 2017). However, we note that this tension between ethics and scientific relevance in the case of consciousness studies cannot be fully resolved by cost-benefit considerations such as “it is better to cause some suffering now in order to prevent a lot of suffering in the future”. Indeed, cost-benefit considerations are used in practice to motivate and constrain the use of animals in scientific research (Banner, 2003; Bateson, 1986; Grimm et al., 2019; Nuffield Council on Bioethics, 2005). For example, in the United Kingdom, harm-benefit analysis is required before determining whether and on what terms to grant a licence to carry out scientific procedures on animals, where harm is defined in terms of pain, suffering, distress, and lasting harm (Animals in Science Regulation Unit, 2015).

---

<sup>2</sup>We thank an anonymous reviewer for pointing out a third option: animal models of psychiatric disorders may be both unethical (in that they cause suffering to animals) and invalid (in that they neglect to model the subjective experience itself; see Taschereau-Dumouchel et al., 2022).

Similar requirements appear in American, European and international regulations (Grimm et al., 2019). Furthermore, legal regulations require researchers to minimise unnecessary suffering (for example, by using anaesthetics when possible), in an effort to minimise harm (Banner, 2003). Yet, the scope of cost- or harm-benefit considerations is not unlimited. We would not consider similar cost-benefit calculations to justify experiments on non-consenting, captive humans, even if properly anaesthetised and under strict regulations, although for many basic science and medical questions, the benefit for humanity could far outstrip the harm caused to these individuals.

It might be thought that approaching these questions via cost-benefit considerations is appropriate when the costs are to non-human animals rather than to humans (“utilitarianism for animals, Kantianism for people”; Nozick, 1974). But more needs to be said about why this might be: what is the critical feature that makes cost-benefit ethics inappropriate in the human case, and is it really true that the animals we would like to use as model organisms lack this feature? Note that even if we were inclined to think that human suffering is more morally important than animal suffering, it does not follow that trading off costs and benefits is the correct approach to these questions, any more than it is in the human case. Without necessarily taking a stand on this matter, and given that current discussions focus on cognitive and experiential capacities as key to justifying the different ethical treatment of human and non-human animals (Johnson et al., 2020; Sunstein & Nussbaum, 2004), we believe consciousness scientists should acknowledge and address this tension between scientific validity and ethical justification in the case of animal models of consciousness.

## What steps should we take?

As we have shown, the scientific study of consciousness is not ethically neutral in that a) it informs ethical decisions, b) it is particularly susceptible to societal and normative biases, and c)

in some cases it introduces a conflict between scientific validity and morality, as in the case of non-human models of consciousness. By this we do not mean to imply that scientists should leave consciousness in the hands of philosophers and theologians. Instead, we believe that the link between the study of consciousness and ethics should be made more explicit, both in the way scientific research is practised and in the way its findings are communicated to non-scientific audiences.

As a first step, the field should prioritise conceptual clarity with respect to the words 'consciousness' and 'awareness' in scientific writing. We suspect that different scientists mean vastly different things when claiming that one is conscious or aware, and that these differences further translate to different attitudes toward what consciousness means for ethics. What do we mean when we say of an animal that it is capable of 'conscious experience' or 'visual consciousness'? What do we mean by the definition of a 'minimally conscious' patient or by saying that a foetus at some gestational age is 'aware' of sounds? While rigorously defining "consciousness" is notoriously difficult, researchers can and should indicate what they take the ethical upshots of consciousness as they understand it to be. Being more explicit about the ethical connotations of the words we use will not only make for a more responsible science, but will also facilitate better communication in this jargon-laden field.

Second, it is essential to have an open scientific discussion about the relation between consciousness and the capacity to suffer. As we review in the article, various aspects of consciousness are considered relevant to ethics, including the presence of phenomenal consciousness (Siewert, 1998), self-awareness and rationality (Kant, 1785/2002), rich phenomenal states (Shepherd, 2018) and functional aspects of consciousness (Danaher, 2020; Levy, 2014). Among the most central ones is the capacity to experience valenced phenomenal states, and suffering more specifically (Bentham, 1789; Mill, 1863/2015). Perhaps surprisingly, current leading scientific theories of consciousness have fairly little to say about the relation between suffering and other dimensions of conscious experience. An open discussion would

bring to the surface hidden preconceptions and their impact on finding interpretation and on theorising . For example, in the second Section we discussed a recent proposal for a classification of different dimensions of animal consciousness, including selfhood and richness of visual experience (Birch, Schnell, et al., 2020). According to the authors, creatures can independently vary on each of those dimensions, giving rise to different consciousness profiles. Which of these dimensions contributes to a capacity to experience suffering is an empirical question with far-reaching ethical implications.

To facilitate an open discussion, we envision a requirement for research papers that make claims about consciousness, awareness, or introspection in model organisms to include a short statement, explaining 1) the degree to which the choice of the model organism rests on aspects of their conscious awareness, 2) the degree to which the study's results shed light on whether the animal is indeed conscious (and if so, in what ways), and 3) the way the first two statements interact with the ethics of the methods used. For example, a statement for a study on neural markers of consciousness in scrub jays may read *“We chose scrub jays as our model organism for demonstrating sophisticated perceptual and cognitive behaviours that suggest conscious experiences. Our finding of stimulus-evoked activity in scrub jays increases the likelihood that their visual awareness resembles that of primates. We do not think the presence of conscious experience by itself should matter for moral standing, and for the ethics of keeping such animals in captivity and performing invasive experiments on them. We believe that this study would not have been ethically defensible if scrub jays were shown to have self-consciousness or metacognition.”* This hypothetical statement entails a qualitative distinction between perceptual consciousness and other forms of consciousness — an important distinction that should be open to scientific and societal criticism.

Finally, we believe consciousness researchers, including those working only with consenting humans, should take an active role in the ethical discussion about these issues, including the use of animal models for the study of consciousness. Studying consciousness, our field has the

responsibility of leading the way on these ethical questions, and of making strong statements where such statements are justified by empirical findings. Recent examples include discussions of ethical ramifications of neuronal signs of foetal consciousness (Lagercrantz, 2014) and a consolidation of evidence for consciousness in vertebrate animals, with a focus on livestock species, ordered by the European Food and Safety Authority (Neindre, C. & others, 2017). In these cases, the science of consciousness provided empirical evidence to weigh on whether a foetus or a livestock animal is conscious or not. The question of animal models of consciousness is simpler, as the presence of consciousness is a prerequisite for the model to be valid. Here we can skip the difficult question of whether or not the entity is indeed conscious, and directly ask: do we believe that consciousness, or some specific form or dimension of consciousness, entails moral status?

It is useful to remind ourselves that ethical beliefs and practices are dynamic: things that were considered acceptable in the past are no longer acceptable today. A relatively recent change is that to the status of non-human great apes (gorillas, bonobos, chimpanzees, and orangutans), such that research on great apes is banned in some countries today, including all European Union member states and New Zealand. In these countries, drilling a hole in the head of a chimpanzee, keeping them in isolation, or restricting their access to drinking water, are forbidden by law. It is a fundamental question of the utmost importance which differences between animals make some practises acceptable with respect to some animals and not others. If consciousness is a determinant of moral status, consciousness researchers have a responsibility in taking an active part in this discussion - by providing scientific observations that either justify current ethical standards or induce the scientific and legal communities to revise these standards.



## *Acknowledgements*

We thank Jonathan Birch, Cormac Dickson, Steve Fleming, Liad Mudrik, Leon Deouell, Lucie Charles, Ian Phillips & Chaz Firestone for their helpful comments.

## References

- Allen, C., & Trestman, M. (2020). Animal Consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2020). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/win2020/entries/consciousness-animal/>
- Animals in Science Regulation Unit. (2015). *The Harm-Benefit Analysis Process: New Project Licence Applications*. UK Home Office. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/487914/Harm\\_Benefit\\_Analysis\\_\\_2\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/487914/Harm_Benefit_Analysis__2_.pdf)
- Banner, M. (2003). *Review of cost-benefit assessment in the use of animals in research* (p. 106). The Animal Procedure Committee.
- Barron, A. B., & Klein, C. (2016). What insects can tell us about the origins of consciousness. *Proceedings of the National Academy of Sciences*, *113*(18), 4900–4908. <https://doi.org/10.1073/pnas.1520084113>
- Bastian, B., & Loughnan, S. (2016). Resolving the Meat-Paradox: A Motivational Account of Morally Troublesome Behavior and Its Maintenance. *Personality and Social Psychology Review*, *21*. <https://doi.org/10.1177/1088868316647562>
- Bastian, B., Loughnan, S., Haslam, N., & Radke, H. R. M. (2012). Don't Mind Meat? The Denial of Mind to Animals Used for Human Consumption. *Personality and Social Psychology Bulletin*, *38*(2), 247–256. <https://doi.org/10.1177/0146167211424291>
- Bateson, P. (1986). When to experiment on animals. *New Scientist*, *109*(1496), 30–32.
- Ben-Haim, M. S., Monte, O. D., Fagan, N. A., Dunham, Y., Hassin, R. R., Chang, S. W. C., &

- Santos, L. R. (2021). Disentangling perceptual awareness from nonconscious processing in rhesus monkeys (*Macaca mulatta*). *Proceedings of the National Academy of Sciences*, 118(15). <https://doi.org/10.1073/pnas.2017543118>
- Bentham, J. (1789). *The Principles of Morals and Legislation*.
- Birch, J., Ginsburg, S., & Jablonka, E. (2020). Unlimited Associative Learning and the origins of consciousness: A primer and some predictions. *Biology & Philosophy*, 35(6), 56. <https://doi.org/10.1007/s10539-020-09772-0>
- Birch, J., Schnell, A. K., & Clayton, N. S. (2020). Dimensions of Animal Consciousness. *Trends in Cognitive Sciences*, 24(10), 789–801. <https://doi.org/10.1016/j.tics.2020.07.007>
- Campbell, Gordon Lindsay. (2014). *The Oxford Handbook of Animals in Classical Thought and Life*. Oxford Handbooks.
- Carpenter, A. D. (2018). Illuminating Community: Animals in Classical Indian Thought. In *Animals*. Oxford University Press. <https://doi.org/10.1093/oso/9780199375967.003.0005>
- Carruthers, P. (2007). Invertebrate Minds: A Challenge for Ethical Theory. *The Journal of Ethics*, 11(3), 275–297. <https://doi.org/10.1007/s10892-007-9015-6>
- Carruthers, P. (2019). *Human and Animal Minds: The Consciousness Questions Laid to Rest* (1st ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780198843702.001.0001>
- Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395(6699), 272–274. <https://doi.org/10.1038/26216>
- Costello, K., & Hodson, G. (2014). Explaining dehumanization among children: The interspecies model of prejudice. *British Journal of Social Psychology*, 53(1), 175–197. <https://doi.org/10.1111/bjso.12016>
- Crane, S. (2015). *Medieval Animal Studies: Dogs at Work*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199935338.013.103>
- Danaher, J. (2020). Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism. *Science and Engineering Ethics*, 26(4), 2023–2049.

<https://doi.org/10.1007/s11948-019-00119-x>

Daston, Lorraine, & Galison, Peter. (2010). *Objectivity*. Zone books.

De Lafuente, V., & Romo, R. (2006). Neural correlates of subjective sensory experience. *Nature Neuroscience*, 8, 1698–1703. <https://doi.org/10.1038/nn1587>

Finnigan, B. (2017). Buddhism and animal ethics. *Philosophy Compass*, 12(7), e12424. <https://doi.org/10.1111/phc3.12424>

Gallup, G. G., & Anderson, J. R. (2018). The “olfactory mirror” and other recent attempts to demonstrate self-recognition in non-primate species. *Behavioural Processes*, 148, 16–19. <https://doi.org/10.1016/j.beproc.2017.12.010>

Giacino, J. T., Fins, J. J., Laureys, S., & Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: The state of the science. *Nature Reviews Neurology*, 10(2), 99–114. <https://doi.org/10.1038/nrneurol.2013.279>

Glover, J. (2006). *The sanctity of life*. 2, 267--72.

Graziano, M. S., & Webb, T. W. (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology*, 6, 500.

Grimm, H., Olsson, I. A. S., & Sandøe, P. (2019). Harm–benefit analysis – what is the added value? A review of alternative strategies for weighing harms and benefits as part of the assessment of animal research. *Laboratory Animals*, 53(1), 17–27. <https://doi.org/10.1177/0023677218783004>

Heath, M. (2008). Aristotle on Natural Slavery. *Phronesis*, 53(3), 243–270. <https://doi.org/10.1163/156852808X307070>

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, 57(2), 243–259. <https://doi.org/10.2307/1416950>

Hofstadter, D. R. (2007). *I am a strange loop*. Basic books.

Hollis, F., & Kabbaj, M. (2014). Social Defeat as an Animal Model for Depression. *ILAR Journal*, 55(2), 221–232. <https://doi.org/10.1093/ilar/ilu002>

- Jena, N. P. (2019). Moral dilemmas of Buddhism on animal suffering. *Asian Philosophy*, 29(3), 248–263. <https://doi.org/10.1080/09552367.2019.1662589>
- Johnson, L. S. M., Fenton, A., & Shriver, A. (2020). *Neuroethics and Nonhuman Animals*. Springer International Publishing.  
<https://link.springer.com/book/10.1007/978-3-030-31011-0>
- Jozefowicz, J., Staddon, J. E. R., & Cerutti, D. T. (2009). Metacognition in animals: How do we know that they know? *Comparative Cognition & Behavior Reviews*, 4.  
<https://doi.org/10.3819/ccbr.2009.40003>
- Kammerer, F. (2019). The Normative Challenge for Illusionist Views of Consciousness. *Ergo, an Open Access Journal of Philosophy*, 6(20201214).  
<https://doi.org/10.3998/ergo.12405314.0006.032>
- Kant, I. (2002). Groundwork of The Metaphysic of Morals. In *Immanuel Kant: Groundwork of the Metaphysic of Morals in focus*. Routledge.
- Kirlic, N., Young, J., & Aupperle, R. L. (2017). Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behaviour Research and Therapy*, 96, 14–29. <https://doi.org/10.1016/j.brat.2017.04.010>
- Kitcher, P. (2003). *Science, Truth, and Democracy*. OUP USA.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, 17(5), 307–321.  
<https://doi.org/10.1038/nrn.2016.22>
- Kohda, M., Hotta, T., Takeyama, T., Awata, S., Tanaka, H., Asai, J., & Jordan, A. L. (2019). If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? *PLOS Biology*, 17(2), e3000021.  
<https://doi.org/10.1371/journal.pbio.3000021>
- Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*,

10–1037.

- Kwok, S. C., Cai, Y., & Buckley, M. J. (2019). Mnemonic Introspection in Macaques Is Dependent on Superior Dorsolateral Prefrontal Cortex But Not Orbitofrontal Cortex. *Journal of Neuroscience*, *39*(30), 5922–5934.  
<https://doi.org/10.1523/JNEUROSCI.0330-19.2019>
- Lagercrantz, H. (2014). The emergence of consciousness: Science and ethics. *Seminars in Fetal and Neonatal Medicine*, *19*(5), 300–305. <https://doi.org/10.1016/j.siny.2014.08.003>
- LeDoux, J. E. (2021). What emotions might be like in other animals. *Current Biology*, *31*(13), R824–R829. <https://doi.org/10.1016/j.cub.2021.05.005>
- Lee, A. Y. (2019). Is consciousness intrinsically valuable? *Philosophical Studies*, *176*(3), 655–671. <https://doi.org/10.1007/s11098-018-1032-8>
- Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565).  
<https://philpapers.org/rec/LEOACI>
- Levy, N. (2014). The Value of Consciousness. *Journal of Consciousness Studies*, *21*, 127–138.
- Longino, H. E. (2020). Science as Social Knowledge: Values and Objectivity in Scientific Inquiry. In *Science as Social Knowledge*. Princeton University Press.  
<https://doi.org/10.1515/9780691209753>
- Maier, S. F. (1984). Learned helplessness and animal models of depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *8*(3), 435–446.  
[https://doi.org/10.1016/S0278-5846\(84\)80032-9](https://doi.org/10.1016/S0278-5846(84)80032-9)
- Mill, J. S. (2015). *On Liberty, Utilitarianism, and Other Essays*. Oxford University Press.
- Neindre, C. & others. (2017). *Animal consciousness*.
- Nieder, A., Wagener, L., & Rinnert, P. (2020). A neural correlate of sensory consciousness in a corvid bird. *Science*, *369*(6511), 1626–1629. <https://doi.org/10.1126/science.abb1447>
- Nozick, R. (1974). Constraints and Animals. In *Anarchy, State and Utopia* (p. 4). Basic Books.

Nuffield Council on Bioethics (Ed.). (2005). *The ethics of research involving animals*. Nuffield Council on Bioethics.

Opinion of the Scientific Panel on Animal Health and Welfare (AHAW) on a request from the Commission related to the aspects of the biology and welfare of animals used for experimental and other scientific purposes. (2005). *EFSA Journal*, 3(12), 292.  
<https://doi.org/10.2903/j.efsa.2005.292>

Oreskes, N. (2021). Why Trust Science? In *Why Trust Science?* Princeton University Press.  
<https://doi.org/10.1515/9780691222370>

Oreskes, Naomi & Conway, Erik M. (2011). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Publishing USA.

Panoz-Brown, D., Iyer, V., Carey, L. M., Sluka, C. M., Rajic, G., Kestenman, J., Gentry, M., Brotheridge, S., Somekh, I., Corbin, H. E., Tucker, K. G., Almeida, B., Hex, S. B., Garcia, K. D., Hohmann, A. G., & Crystal, J. D. (2018). Replay of Episodic Memories in the Rat. *Current Biology*, 28(10), 1628-1634.e7. <https://doi.org/10.1016/j.cub.2018.04.006>

Plotnik, J. M., Waal, F. B. M. de, & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences*, 103(45), 17053–17057.  
<https://doi.org/10.1073/pnas.0608062103>

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences*, 91, 6.

Prior, H., & Schwarz, A. (2008). Mirror-Induced Behavior in the Magpie (*Pica pica*): Evidence of Self-Recognition. *PLoS Biology*, 6(8), 9.

Redinbaugh, M. J., Phillips, J. M., Kambi, N. A., Mohanta, S., Andryk, S., Dooley, G. L., Afrasiabi, M., Raz, A., & Saalman, Y. B. (2020). Thalamus Modulates Consciousness via Layer-Specific Control of Cortex. *Neuron*, 106(1), 66-75.e12.  
<https://doi.org/10.1016/j.neuron.2020.01.005>

- Ribot, T. (1889). *Psychologie de l'attention*.
- Rosenthal, D. (2005). *Consciousness and Mind*. Clarendon Press.
- Sachidhanandam, S., Sreenivasan, V., Kyriakatos, A., Kremer, Y., & Petersen, C. (2013). Membrane potential correlates of sensory perception in mouse barrel cortex. *Nature Neuroscience*, 16. <https://doi.org/10.1038/nn.3532>
- Shepherd, J. (2018). *Consciousness and Moral Status* (1st ed.). Routledge. <https://doi.org/10.4324/9781315396347>
- Sherman, G. D., & Haidt, J. (2011). Cuteness and Disgust: The Humanizing and Dehumanizing Effects of Emotion. *Emotion Review*, 3(3), 245–251. <https://doi.org/10.1177/1754073911402396>
- Siewert, C. (1998). The Significance of Consciousness. In *The Significance of Consciousness*. Princeton University Press. <https://doi.org/10.1515/9781400822720>
- Smith, J. A., Andrews, P. L. R., Hawkins, P., Louhimies, S., Ponte, G., & Dickel, L. (2013). Cephalopod research and EU Directive 2010/63/EU: Requirements, impacts and ethical review. *Journal of Experimental Marine Biology and Ecology*, 447, 31–45. <https://doi.org/10.1016/j.jembe.2013.02.009>
- Solvi, C., Al-Khudhairy, S. G., & Chittka, L. (2020). Bumble bees display cross-modal object recognition between visual and tactile senses. *Science*, 367(6480), 910–912. <https://doi.org/10.1126/science.aay8064>
- Sorabji, R. (2018). Animal Minds and Human Morals: The Origins of the Western Debate. In *Animal Minds and Human Morals*. Cornell University Press. <https://doi.org/10.7591/9781501717888>
- Strawson, G. (2018). The Consciousness Deniers. *The New York Review of Books*, 10.
- Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, 30(3), 299–313. <https://doi.org/10.1017/S0140525X07001975>

- Sunstein, C. R., & Nussbaum, M. C. (2004). *Animal Rights: Current Debates and New Directions*. Oxford University Press.
- Taschereau-Dumouchel, V., Michel, M., Lau, H., Hofmann, S. G., & LeDoux, J. E. (2022). Putting the “mental” back in “mental disorders”: A perspective from research on fear and anxiety. *Molecular Psychiatry*, 1–9. <https://doi.org/10.1038/s41380-021-01395-5>
- Tulving, E. (2005). Episodic Memory and Autonoesis: Uniquely Human? In *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3–56). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195161564.003.0001>
- Vugt, B. van, Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, 360(6388), 537–542. <https://doi.org/10.1126/science.aar7186>
- Wada, M., Takano, K., Ora, H., Ide, M., & Kansaku, K. (2016). The Rubber Tail Illusion as Evidence of Body Ownership in Mice. *Journal of Neuroscience*, 36(43), 11133–11137. <https://doi.org/10.1523/JNEUROSCI.3006-15.2016>
- Willner, P. (2017). The chronic mild stress (CMS) model of depression: History, evaluation and usage. *Neurobiology of Stress*, 6, 78–93. <https://doi.org/10.1016/j.ynstr.2016.08.002>
- Yuki, S., & Okanoya, K. (2017). Rats show adaptive choice in a metacognitive task with high uncertainty. *Journal of Experimental Psychology: Animal Learning and Cognition*, 43(1), 109–118. <https://doi.org/10.1037/xan0000130>