aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
1–9

DOI: 10.1177/17456916221110222
www.psychologicalscience.org/PPS

SAGE

The Scientific Study of Consciousness Cannot and Should Not Be Morally Neutral

Matan Mazor[1,2], Simon Brown[3], Anna Ciaunica[4], Athena Demertzi[5,6], Johannes Fahrenfort[7,8], Nathan Faivre[9,10], Jolien C. Francken[11], Dominique Lamy[12,13], Bigna Lenggenhager[14], Michael Moutoussis[2,15], Marie-Christine Nizzi[16,17,18], Roy Salomon[19], David Soto[20,21], Timo Stein[7], and Nitzan Lubianiker[13,22]

[1]Department of Psychological Sciences, Birkbeck, University of London; [2]Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London; [3]Department of Philosophy, Johns Hopkins University; [4]Centre for Philosophy of Science, University of Lisbon; [5]Physiology of Cognition, GIGA Consciousness Research Unit, Université de Liège; [6]Fund for Scientific Research, Bruxelles, Belgium; [7]Department of Psychology, University of Amsterdam; [8]Department of Experimental and Applied Psychology, Vrije Universiteit; [9]Centre for Neuroprosthetics and Brain Mind Institute, Faculty of Life Sciences, Swiss Federal Institute of Technology; [10]University Grenoble Alpes, University Savoie Mont Blanc, CNRS, LPNC; [11]Faculty of Philosophy, Theology and Religious Studies, Radboud University; [12]Sagol School of Neuroscience, Tel Aviv University, Tel-Aviv, Israel; [13]School of Psychological Sciences, Tel Aviv University; [14]Department of Psychology, University of Zurich; [15]Max Planck-University College London Centre for Computational Psychiatry and Ageing Research, University College London; [16]Semel Institute for Neuroscience and Human Behavior, Department of Psychiatry & Biobehavioral Sciences, David Geffen School of Medicine, University of California Los Angeles; [17]Cognitive Science Program, Dartmouth College; [18]Institute for Interdisciplinary Brain and Behavioral Sciences, Chapman University; [19]Gonda Multidisciplinary Brain Research Centre, Bar-Ilan University; [20]Basque Centre on Cognition, Brain and Language, San Sebastian, Spain; [21]Ikerbasque, Basque Foundation for Science, Bilbao, Spain; and [22]Sagol Brain Institute, Tel-Aviv Medical Centre, Tel Aviv, Israel

Abstract

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 programs or interpret results in ways that justify current norms rather than challenge them. We show that because consciousness largely determines moral status, the use of nonhuman 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. Finally, 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.

Keywords

allied field: philosophy, application: policy, comparative psychology, consciousness, ethics, metascience

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.[1] In ancient traditions, moral obligations toward nonhuman animals often rested on conscious aspects of experience, such as the capacity to suffer. For instance, according

Corresponding Author:
Matan Mazor, Department of Psychological Sciences, Birkbeck, University of London, and Wellcome Centre for Human Neuroimaging, University College London
Email: mtnmzor@gmail.com

to Dhārmic religions, the principle of ahiṃsā (nonviolence) is applied to animals on the basis of 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 minimize animal suffering (Campbell, 2014). In contrast, Aristotelians, who dominated medieval 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 nonhuman 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 postmedieval Western thought, different philosophical traditions identify different mental capacities as the determining factor for moral status. For Immanuel Kant (1785/2002), it was autonomy. In contrast, utilitarian philosophers identified the origin of moral status in the capacity to experience suffering (Bentham, 1789/1996) 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 among the most important. Even people 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).

Because of 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 on the basis of 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 & Galison, 2010; Ribot, 1889). More recently, White children and adults attributed reduced emotions to Black compared with 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 (2017) proposed that the denial of mind to certain animals resolves the cognitive dissonance between the practice of eating meat 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 preexisting 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 behavioral 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 behavioral and neural markers. For example, in a recent review of consciousness in the animal kingdom, Birch, Schnell, and Clayton (2020) listed 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, & Jablonka, 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 researchers 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 noncommunicating patients,

preverbal infants, and nonhuman animals. To date, all suggested nonverbal markers of conscious states that have been identified in humans, behavioral and neural, have also been observed in other animals.

Invasive experiments on corvid birds (Nieder et al., 2020), rodents (Sachidhanandam et al., 2013), and nonhuman primates (De Lafuente & Romo, 2006; Leopold & Logothetis, 1996; Vugt et al., 2018) have revealed 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 have also been used to document self-awareness in nonhuman animals with better experimental control (Wada et al., 2016). Behavioral and brain experiments have provided 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; e.g., for a critical review of the mirror-test literature, see Gallup & Anderson, 2018). Instead, it is meant to show that as of April 2022, no single nonverbal behavioral 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 behavior and physiological processes to conjectures about private conscious experiences. Whereas 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 nonhuman animals. Examples include debates about the interpretation of the mirror test (Kohda et al., 2019) and about the true nature of metacognitive behavior (Jozefowiez et al., 2009), episodic memory (Suddendorf & Corballis, 2007; Tulving, 2005), and emotional experiences (LeDoux, 2021) in nonhuman 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 nonhuman agents. These priors can result in liberal biases to identify consciousness more readily in some cases and conservative biases to be more skeptical in other cases. Although some of these priors are based on previous literature, others may be guided by prescientific tendencies to ascribe agentic and experiential states on the basis of behavioral (Heider & Simmel, 1944) and physical (Sherman & Haidt, 2011) cues. Critically, as we show in the previous section, prescientific 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, scientists may believe that (a) beings are conscious if and only if they can integrate information across different senses, (b) consciousness is a determinant of moral status, and (c) insects deserve no moral consideration. Scientists 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; for a general scheme, see Fig. 1). Scientists 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 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 (Animal Health and Welfare, 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.

Some readers may worry that accepting that the scientific study of consciousness cannot and should not be morally neutral might jeopardize 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

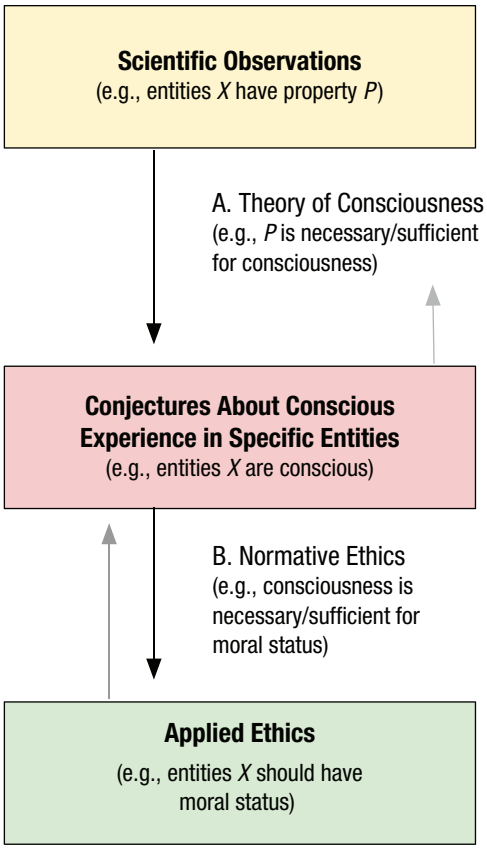

Fig. 1. A schematic description of the link between scientific observations, conjectures about conscious experience, and applied ethics. Descending black arrows: scientific observations are translated into conjectures about conscious experience in specific entities via a theory of consciousness. These conjectures then go on to affect applied ethics via general beliefs about the relation between consciousness and ethics. Ascending gray 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.

currently legally accepted, arguably does not undermine but instead enhances scientific trustworthiness and legitimacy (Kitcher, 2003; Longino, 2020; Oreskes, 2021). Historians and sociologists have documented cases in which 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 nonhuman animals could unduly influence consciousness science, especially if not addressed and discussed.

The animal-models-of-consciousness paradox

An instance in which 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 and 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. Because 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 to have its life considered to be deserving of 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 anesthetized macaque monkeys (Redinbaugh et al., 2020). For this study, two monkeys were kept in captivity, implanted with brain electrodes, and immobilized by sticking rods in a head implant during electrophysiological recordings. In another study from 2021, a behavioral 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 were 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 that conform with current ethical guidelines in place at a national level and are commonplace in many fields of study. Yet common to these studies is that their scientific relevance rests on the animal being conscious,

whereas 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: the capacity to suffer. To be clinically relevant, model animals must show behaviors that in humans are interpreted as indicating a mental illness or distress. For example, in research about depression, nonhuman animals can be led to express behavioral 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 posttraumatic 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* (American Psychiatrist Association, 2013), a diagnosis of depression depends on patients 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 nonhuman animals? An affirmative answer to this question would call into question the morality of such experiments, and a negative answer would call into question their scientific validity and clinical utility.[2]

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 that has been applied to alleviate the suffering of humans and other animals (e.g., 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 license to carry out scientific procedures on animals. The term "harm" encompasses pain, suffering, distress, and lasting harm (Animals in Science Regulation Unit, 2015). Similar requirements appear in American, European, and international regulations (Grimm et al., 2019). Furthermore, legal regulations require researchers to minimize unnecessary suffering (e.g., by using anesthetics when possible) in an effort to minimize 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 nonconsenting, captive humans, even if properly anesthetized 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 nonhuman 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 researchers would like to use as model organisms lack this feature? Note that even if researchers 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 nonhuman 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 researchers 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 nonhuman 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 practiced and in the way its findings are communicated to nonscientific audiences.

As a first step, the field should prioritize 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 a subject is conscious or aware and that these differences further translate to different attitudes toward what consciousness means for ethics. What do scientists mean when they say of an animal that it is capable of "conscious experience" or "visual consciousness"? What do they mean by the definition of a "minimally conscious" patient or by saying that a fetus at some gestational age is "aware" of sounds? Although 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 scientists 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, 1789/1996; 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 theorizing. 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, & Clayton, 2020). According to Birch, Schnell, and Clayton (2020), 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 articles that make claims about consciousness, awareness, or introspection in model organisms to include a short statement explaining (a) the degree to which the choice of the model organism rests on aspects of its conscious awareness, (b) the degree to which the study's results shed light on whether the animal is indeed conscious (and if so, in what ways), and (c) 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 as follows:

> We chose scrub jays as our model organism for demonstrating sophisticated perceptual and cognitive behaviors 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, the field has the responsibility of leading the way on these ethical questions and of making strong statements when such statements are justified by empirical findings. Recent examples include discussions of ethical ramifications of neuronal signs of fetal 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 (Le Neindre et al., 2017). In these cases, the science of consciousness provided empirical evidence to weigh on whether a fetus or a livestock animal is conscious. The question of animal models of consciousness is simpler because the presence of consciousness is a prerequisite for the model to be valid. Here, researchers can skip the difficult question of whether 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 nonhuman 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 chimpanzees' heads, 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 practices 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.

Transparency

Action Editor: Timothy J. Pleskac
Editor: Klaus Fiedler
Author Contributions
Authors appear in alphabetical order, except for M. Mazor and N. Lubianiker, who cowrote the original version of this article.
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

ORCID iDs

Matan Mazor https://orcid.org/0000-0002-3601-0644
Simon Brown https://orcid.org/0000-0002-5606-6979
Nathan Faivre https://orcid.org/0000-0001-6011-4921
Marie-Christine Nizzi https://orcid.org/0000-0002-2062-8992

Acknowledgments

We thank Jonathan Birch, Cormac Dickson, Steve Fleming, Liad Mudrik, Leon Deouell, Lucie Charles, Ian Phillips, and Chaz Firestone for their helpful comments, and Bat-Sheva Mazor for inspiring us to write this article.

Notes

1. Our focus in this article is on consciousness as a term used in the scientific literature, in contrast to consciousness as a mental state or capacity. Thus, this definition is intended to be agnostic regarding different theoretical approaches to consciousness in the scientific literature.
2. 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)).

References

Allen, C., & Trestman, M. (2020). Animal consciousness. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2020/entries/consciousness-animal/

American Psychiatrist Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Animal Health and Welfare. (2005). 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. *EFSA Journal*, *3*(12), 292. https://doi.org/10.2903/j.efsa.2005.292

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

Banner, M. (2003). *Review of cost-benefit assessment in the use of animals in research*. 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, USA*, *113*(18), 4900–4908. https://doi.org/10.1073/pnas.1520084113

Bastian, B., & Loughnan, S. (2017). Resolving the meat-paradox: A motivational account of morally troublesome behavior and its maintenance. *Personality and Social Psychology Review*, *21*(3), 278–299. 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, USA*, *118*(15), Article e2017543118. https://doi.org/10.1073/pnas.2017543118

Bentham, J. (1789/1996). *The collected works of Jeremy Bentham: An introduction to the principles of morals and legislation*. Clarendon Press.

Birch, J., Ginsburg, S., & Jablonka, E. (2020). Unlimited Associative Learning and the origins of consciousness: A primer and some predictions. *Biology & Philosophy*, *35*(6), Article 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, G. L. (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 P. Adamson & G. F. Edwards (Eds.), *Animals: A history* (pp. 63–86). 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*. 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, L., & Galison, P. (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), Article 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. In H. Kuhse & P. Singer (Eds.), *Bioethics: An anthology* (pp. 266–275). Blackwell.

Graziano, M. S., & Webb, T. W. (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology*, 6, Article 500. https://doi.org/10.3389/fpsyg.2015.00500

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. https://link.springer.com/book/10.1007/978-3-030-31011-0

Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. T. (2009). Metacognition in animals: How do we know that they know? *Comparative Cognition & Behavior Reviews*, *4*, 29–39. 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*. https://doi.org/10.3998/ergo.12405314.0006.032

Kant, I. (2002). Groundwork of the metaphysic of morals. In L. Pasternack (Ed.), *Immanuel Kant: Groundwork of the metaphysic of morals in focus* (pp. 17–98). Routledge. (Original work published 1785)

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*. Oxford University Press.

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), Article 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*, *90*(4), 543–555. https://doi.org/10.1037/0022-3514.90.4.543.

Kwok, S. C., Cai, Y., & Buckley, M. J. (2019). Mnemonic introspection in macaques is dependent on superior dorsolateral prefrontal cortex but not orbitofrontal cortex. *The 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

Le Neindre, C., Bernard, E., Boissy, A., Boivin, X., Calandreau, L., Delon, N., Deputte, B., Desmoulin-Canselier, S., Dunier, M., Faivre, N., Giurfa, M., Guichet, J.-L., Lansade, L., Larrère, R., Mormède, P., Prunet, P., Schaal, B., Servière, J., & Terlous, C. (2017). *Animal consciousness*. European Food Safety Author. https://efsa.onlinelibrary.wiley.com/doi/pdfdirect/10.2903/sp.efsa.2017.EN-1196

Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565), 549–553. https://doi.org/10.1038/379549a0

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*. 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

Mill, J. S. (2015). *On liberty, utilitarianism, and other essays*. Oxford University Press. (Original work published 1863)

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). *Anarchy, state and Utopia*. Basic Books.

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

Oreskes, N. (2021). *Why trust science?* Princeton University Press. https://doi.org/10.1515/9780691222370

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

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., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences, USA*, *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, USA*, *91*(16), 7398–7403. https://doi.org/10.1073/pnas.91.16.7398

Prior, H., & Schwarz, A. (2008). Mirror-induced behavior in the magpie (Pica pica): Evidence of self-recognition. *PLOS Biology*, *6*(8), Article e202. https://doi.org/10.1371/journal.pbio.0060202

Redinbaugh, M. J., Phillips, J. M., Kambi, N. A., Mohanta, S., Andryk, S., Dooley, G. L., Afrasiabi, M., Raz, A., & Saalmann, 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. H. (1889). *Psychologie De L'attention* [The psychology of attention]. Paris: Librairie Félix Alcan

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*, 1671–1677. https://doi.org/10.1038/nn.3532

Shepherd, J. (2018). *Consciousness and moral status*. 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*. 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*. Cornell University Press. https://doi.org/10.7591/9781501717888

Strawson, G. (2018). The consciousness deniers. *The New York Review*, March 13. https://www.nybooks.com/daily/2018/03/13/the-consciousness-deniers/#:~:text=One%20of%20the%20strangest%20things,%2C%20in%20fact%2C%20an%20illusion.&text=What%20is%20the%20silliest%20claim%20ever%20made%3F

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*, *27*, 1322–1330. https://doi.org/10.1038/s41380-021-01395-5

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human? In H. S. Terrace & J. Metcalfe (Eds.), *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. *The 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