

Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., & Skudlarski, P. (2003). The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 358, 415–427.

Skuse, D., Morris, J., & Lawrence, K. (2003). The amygdala and development of the social brain. *Annals of the New York Academy of Science*, 1008, 91–101.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, 3, 764–769.

Vuilleumier, P. (2000). Faces call for attention: Evidence from patients with visual extinction. *Neuropsychologia*, 38, 693–700.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1–34.

Wright, C. I., Martis, B., Shin, L. M., Fischer, H., & Rauch, S. L. (2002). Enhanced amygdala responses to emotional versus neutral schematic facial expressions. *Neuroreport*, 13, 785–790.

Reading 6.3

Animal Neuroethics and the Problem of Other Minds¹

Martha J. Farah

The “problem of other minds” is a central problem in the philosophy of mind. It refers to the difficulty of knowing whether someone or something, other than oneself, has a mind. What is the relevance of the problem of other minds to ethics and neuroscience? Its relevance to ethics rests on the relation between moral standing and capacity for mental life, particularly the capacity to suffer. If a being is capable of suffering, then it deserves protection from suffering. How and whether we can know about the mental lives of others is therefore an epistemological question with direct relevance to ethics. The relevance of this question to neuroscience rests on the potential value of neuroscience evidence for informing us about a being's mental life. In this article I will argue that, within the context of a certain class of metaphysical assumptions concerning mind–brain relations, neuroscience evidence is different from the kinds of evidence traditionally used to infer mental life and that it is in principle more informative. I will then discuss the potential implications of neuroscience evidence for animal ethics.

6.3.1 From Behavior to Mental States: The Argument from Analogy

The problem of other minds is a consequence of mind–body dualism, specifically the idea that there is no necessary relation between physical bodies and their behavior, on the one hand, and mental processes, on the other. Descartes' famous “I think therefore I am” expresses a basis for certainty concerning the existence of our own mental life. But on what basis can we infer that other people have minds? Attempts to justify our belief that other people have minds have generally rested on a kind of analogy, also discussed by Descartes (1968) and emphasized by Locke (1959) and other British empiricists such as J.S. Mill (1979). The analogy uses the known relation between physical and mental events in

oneself to infer the mental events that accompany the observable physical events for someone else. For example, as shown in Analogy 1, when I stub my toe, this causes me to feel pain, which in turn causes me to say “ouch!” When I see Joe stub his toe and say “ouch,” I infer by analogy that he feels pain.

Analogy 1

You stub your toe	→ You feel pain	→ You say “ouch!”
Joe stubs his toe	→ <i>Joe feels pain</i>	→ Joe says “ouch!”

The problem with this analogy is that it begs the question. Why should I assume that the same behavioral–mental relations that hold in my case also hold in Joe’s? Joe could be acting and not really feel pain. He could even be a robot without thoughts or sensations at all. The assumption that analogous behavior–mental state relations hold for other people is essentially what the analogy is supposed to help us infer.

The question of whether someone is actually in pain or is acting or a robot might seem academic. After all, common sense tells us that there are no robots human-like enough to fool us, and barring very special circumstances, there is little reason to suspect anyone of acting. However, the question is not academic when applied to non-human animals whose behavioral repertoires are different from ours and who lack language.

6.3.2 Behavior and Brain Activity as Evidence of Mental Life

Brain imaging may be able to provide evidence of mental life when behavior cannot. Here I will argue that brain activity is not simply a more sensitive measure of cognitive processing than behavior but is qualitatively different from behavior in the inferences that it allows. Consider the possibility that brain activity and behavior play analogous roles. The diagram of Analogy 2 illustrates this possibility, by replacing the “ouch” behavior of Analogy 1 with activation of the anterior cingulate cortex (ACC), part of the brain’s pain network.

Analogy 2

You stub your toe	→ You feel pain	→ Your ACC activates
Joe stubs his toe	→ <i>Joe feels pain</i>	→ Joe’s ACC activates

The problem with Analogy 2 is that it implies that feeling pain causes ACC activation, just as it causes saying “ouch.” However, the relations between mental states and behaviors are different in kind from the relations between mental states and brain states. Mental states and behaviors

are contingently related. What one means by a term like *feeling pain* is not a behavior, or even a behavioral disposition. Although this possibility was explored in earnest by some behaviorist philosophers several decades ago, for example by Ryle (1949), it is not now regarded as a viable approach to the meaning of mental state terms.

For purposes of knowing mental states, behavior is like an indicator light. Indicator lights can be disabled or disconnected, or they can be turned on by other means. Their relation to the thing indicated is contingent on being hooked up a certain way. Inferences based on indicator lights and, analogously, behavior are therefore fallible. In contrast, virtually all contemporary approaches to the mind–body problem regard the relation between mental states and brain states as noncontingent.

The predominant view of the relation between mental states and brain states in cognitive neuroscience and contemporary philosophy of mind is one of identity: mental states *are* brain states. According to one version of this view, “type identity,” each type of mental event is a type of physical event (Churchland, 1988; Smart, 1959). According to a weaker version, “token identity,” every instance of a mental event is an instance of a physical event. The most widely accepted version of token identity is based on “functionalism,” which identifies the functional role of a physical state, in mediating between the inputs and outputs of the organism, as the determinant of its corresponding mental state (Block & Fodor, 1972). Functionalism has many versions of its own, some of which blur the line between type and token mind–brain identity theory (e.g., Armstrong, 1968; Lewis, 1980).

There is an alternative to mind–brain identity based on the idea that mental states “supervene” on brain states, which avoids substance dualism yet stops short of equating mental states with brain states (Davidson, 1970; Kim, 2005). Analogy 2 is incompatible with supervenience theories as well as identity theories. This is because, despite the nonidentity of mental and brain states according to supervenience theories, the relation between the two is stronger than mere causality. According to supervenience, mind–brain relations are noncontingent. In the words of Davidson, “there cannot be two events alike in all physical respects but differing in some mental respects [and] an object cannot alter in some mental respects without altering in some physical respects.”

In sum, across all these different contemporary metaphysical positions on the mind–body problem, the relationship between mental states and brain states is not contingent, as with the causal relations in Analogy 1. For type identity theories as well as functionalist theories, the ACC

activation of the example is identical to a pain. For supervenience theories, the ACC activation cannot exist without there being pain. Thus it makes more sense to diagram the inferences from brain activity to mental state as in Analogy 3. The gist of this analogy is that, however sure you are of the ACC activation in Joe's brain, you can be that sure that Joe is in pain. In sum, the argument from analogy with brain activity is immune to the alternative interpretations that plague the behavioral analogy.

Analogy 3

You stub your toe	→ You feel pain
	Your ACC activates
Joe stubs his toe	→ <i>Joe feels pain</i>
	Joe's ACC activates

6.3.3 The Problem of Other, Nonhuman, Minds

Non-human animals have limited communicative abilities, and this limitation deprives us of the usual methods for learning about their mental states. Although few people today would agree with Descartes' conclusion that animals lack mental states altogether (1968), most of us feel uncertain about the extent and nature of animals' mental lives. On the one hand, many of us anthropomorphize certain animals, especially our pets, attributing complex thoughts and expectations to them on the basis of what a human in the same situation might think. On the other hand, the mental life of animals is often treated by us as hypothetical, incomparably different from our own, or even nonexistent. How else to explain our acceptance of glue traps for rodents and boiled lobster dinners?

As shown in the diagram of Analogy 4, nonhuman animals present us with a version of the problem of other minds for which the usual problematic analogy is even more problematic because of differences between human and animal behavioral repertoires. Animals cannot talk, and may not even express distress in nonverbal ways that are analogous to ours. For example, they may not vocalize at all and may freeze rather than struggle when afraid.

Analogy 4

You stub your toe	→ You feel pain	→ You say "ouch!"
Bat stubs his toe	→ <i>Bat feels pain</i>	→ Bat squeaks

Can the neuroscience approach provide traction for exploring the mental life of other species? To a degree it already has, yet according to the current analysis it could provide even more. Ethicists have previously

brought physiologic data to bear on the question of animal suffering, specifically the similarities between human and animal pain systems. For example, Singer (1990, pp. 12–13) quotes at length from the writings of a pain researcher to the effect that pain processing is a lower level brain function that differs little between humans and other animals. This use of physiologic data differs in two ways from the current one.

First, according to the current analysis, physiologic data are not simply one more source of evidence about a being's mental life, to be weighed together with behavioral evidence, valuable as they might be in that role. Rather, physiologic data can play a qualitatively different and more definitive role because of their noncontingent relation to mental states, as argued in the last section. In terms of the inferences diagrammed earlier, this is the difference between Analogies 2 and 3.

The second difference results from the relatively new ability of cognitive neuroscience to parse brain processes into psychologically and ethically meaningful categories. In the current case, it has revealed the neural basis of the distinction between what could be called "mere pain" and suffering. Pains vary along many dimensions, and one dimension of particular ethical relevance is the psychological quality of the pain (Dawkins, 1985; Dennett, 1996; Hardcastle, 1999). Some pain experiences are primarily physical whereas others are psychologically distressing. The latter, characterized by Dawkins as both unpleasant and intense, warrant the term *suffering*. The neural states corresponding with pain states appear to respect this important distinction, demarcating the physical and psychologic components of pain experience by the involvement of different brain areas.

Research with animals and humans has revealed a widespread network of brain areas that become active in response to pain-inducing stimuli, including thalamic and somatosensory cortical regions as well as regions further removed from the sensory input such as the insula and anterior cingulate cortex (Porro, 2003). When the physical intensity of pain is varied, for example in an imaging experiment by having human subjects touch a painfully hot surface that varies in temperature, the level of activity throughout this network varies (Becerra et al., 2001). Taking advantage of the human ability to report their mental states (and in principle the possibility of first person research in which one introspects on one's own mental states), it is possible to vary independently the physical and psychological dimensions of pain and map the brain states that correspond with each. Morphine, for example, is known to diminish the psychological component of pain. Patients commonly report

that they still feel the “physical” pain but that they are less bothered by it. The same is reported by patients whose pain is treated with hypnotic suggestion. Both interventions have their neural effects primarily in the ACC (Kupers, Faymonville, & Laureys, 2005; Lidstone & Stoessl, 2007). When people who are not being subjected to pain are empathizing intensely with someone who is, their ACCs become activated in the absence of physical pain (Singer et al., 2004). These findings indicate that ACC activation reflects suffering rather than “mere” pain.

Shriver (2006) points out that mammals have ACCs and are thus neurally equipped for psychological as well as physical pain. Following Shriver, we can substitute an animal for Joe in Analogy 4. However, because brain states can only be as similar as the brains that have them, we must amend Analogy 4 to specify a human ACC in one’s own case and an animal ACC in the animal’s.

Analogy 5

You stub your toe	→ You feel pain
	Your human ACC activates
Bat stubs his toe	→ <i>Bat feels pain</i>
	Bat’s bat ACC activates

This raises the question of how could one determine whether mind–brain relations established with one species’ brain generalize to other species. Behind this question is a more fundamental one about how degrees and types of variation in brain states correspond with degrees and types of variation in mental states, a question that will arise even within a given species because no two brains are identical. In principle, one could manipulate human brains (including one’s own) to systematically vary all the different biophysical characteristics by which brains differ, in order to discover what the relevant aspects of the brain state are for determining the mental state. Of course in practice this is not even remotely possible.

At best we can suppose that similarity of psychological state will fall off as similarity of brain state falls off, without knowing which aspects of brain state similarity are relevant or how sharply the one falls off relative to the other. Edelman, Baars, and Seth (2005) provide an example of the attempt to identify functional similarities in brain architecture across species, including nonmammalian species. Shriver (2006) attempts to address the problem of generalization from human to non-human in the case of pain by citing evidence that the ACC plays a similar role in rat and human pain experience (although this evidence is admittedly

based on behavior, which the current appeal to brain evidence was intended to replace): LaGraize and colleagues (2004) compared the behavior of rats with and without lesions of the ACC when forced to choose between staying in the dark, which rats generally prefer, and avoiding electric shocks to their feet. All of the rats reacted similarly when shocked, by withdrawing the shocked foot, thus indicating preserved perception of pain. However, the lesioned rats were more willing to experience the shocks for the sake of staying in a dark region of the experimental apparatus. Like patients on morphine, they appeared to be less distressed by the pain. This implies that rat ACCs play a role similar to human ACCs.

Am I suggesting that neuroscience can tell us “what it is like to be a bat?” Yes and no. When Nagel (1974) framed this question, he chose the bat as his nonhuman animal because bats use echolocation to perceive the world, a sense that humans lack. Knowing what it is like to perceive the world with a sense we lack remains a problem, even with the help of neuroscience, because the neural systems that perform echolocation in bats have no obvious homolog in the human brain. However, given that we do share the same general pain physiology with bats, including an ACC, we can know certain things about what it’s like to be one of them. Specifically, we can infer that to be a bat with an injured toe is more like being a human with an injured toe and no pain relief than it is like being a human with an injured toe who has been given morphine.

The problem of animal minds has not thus far figured prominently in the field of neuroethics. One reason may be that neuroethics is young and has yet to engage all of the subject matter that will eventually comprise the field. Another reason may be that the personal and political rancor associated with animal ethics has discouraged scholars from approaching this topic. Given the real-world importance of animal ethics, and the special role that neuroscience evidence can play in this endeavor, the study of animal neuroethics would seem to have great promise.

6.3.4 Assumptions and Conclusions

The idea that neuroscience can reveal ethically relevant information about non-human animals rests on a number of assumptions. One assumption that has not been examined in this article is that our ethical obligations toward a being depend at least in part on the mental life of that being. Although this assumption hardly needs defending, there is

much more that could be said about which specific aspects of mental life have which specific ethical implications. Perhaps the most important further clarification concerns those aspects of mental life that obligate us to prevent suffering and those that obligate us to protect life.

The current article has focused on the question of whether another being has the capacity for relatively simple mental states, those with some consciously experienced content and affective valence. This mental capacity has more limited ethical implications than the mental capacity to conceive of oneself and one's life and have an explicit preference to continue living (Levy, 2008; Singer, 1999). The neuroscience evidence discussed so far pertains only to the capacity of animals to experience the former kind of mental state, and the relevant ethical implications are therefore confined to preventing suffering rather than protecting life. However, this is not an in-principle limitation of neuroscience data. Given the appropriate research program, there is no reason why we could not identify the neural systems, and states thereof, corresponding with the self-concept and the desire to continue living. This knowledge would have implications for many aspects of end-of-life decision making and might obligate us to refrain from killing certain animals.

Another assumption that deserves explicit discussion concerns the relation between cognitive processing, of the kind that cognitive neuroscientists correlate with brain activation in imaging experiments, and consciousness. This is an important assumption in the current context because our ethical concern is with conscious mental life, and conscious suffering in particular, rather than with unconscious information processing. Based on most of the views of mind–brain relations reviewed earlier, certain types or instances of neural processing are identical to, or are necessarily associated with, certain mental states, including conscious mental states. Therefore the problem is one of determining empirically which brain states correspond with which conscious mental states. This is not a trivial problem, but it is in principle solvable. Indeed, if one is willing to accept other normal humans' reports of conscious experience as evidence, we are on our way to solving it in practice. (Skeptics unwilling to accept others' reports of conscious experience would have to be scanned themselves, which could be done to verify specific findings but would not be feasible as a means of verifying all cognitive neuroscience knowledge.)

A final assumption concerns the accuracy and completeness of cognitive neuroscience. For purposes of exploring the in-principle prospects

and limitations of neuroscience evidence as a solution to the problem of other minds, I have written as if we know the brain states associated with specific mental states. Unfortunately, this is not true. Although cognitive neuroscience has made tremendous progress in the past few decades, the current state of our knowledge is far from complete. For many mental states, including suffering, we have good working hypotheses about the brain regions that are relevant, but future research will undoubtedly call for the revision of some of these hypotheses. In addition, we know little about the specific mechanisms by which these regions implement the relevant mental states. "Activation" observed in brain imaging studies is closely related to neural activity measured at the single cell level, but does not map perfectly onto a specific aspect of neuronal behavior such as action potentials (Logothetis & Wandell, 2004). Furthermore, any single measure of brain activity, be it single cell or aggregate, electrical or chemical, will omit potentially important features of neuronal function. It is possible that activation as measured by our current methods is not diagnostic of the relevant neuronal activity and that under some circumstances it will be misleading. Knowing more about the specific computations performed by neurons in the brain regions implicated by brain imaging, including their interactions with neurons in other regions, will be particularly important as we attempt to evaluate cross-species homologies.

Notes

1. Editor's note: This reading was excerpted from an article that appeared in 2008 in the journal *Neuroethics*, volume 1, pages 9–18, under the title "Neuroethics and the problem of other minds: Implications of neuroscience for the moral status of vegetative patients and nonhuman animals," and is used with permission. The omitted section concerned the problem of other minds in relation to vegetative and minimally conscious patients. I thank Liz Camp, Neil Levy, Adrienne Martin, and Susan Schneider for their helpful comments on an earlier draft of the paper. The writing of this paper was supported by NIH grants R21-DA01586, R01-HD043078, R01-DA14129, and R01-DA18913 and by the MacArthur Project on Neuroscience and the Law.

References

- Armstrong, D. M. (1968). *A materialistic theory of the mind*. London, UK: RKP.
- Becerra, L. R., Breiter, H. C., Wise, R., Gonzalez, R. G., & Borsook, D. (2001). Reward circuitry activation by noxious thermal stimuli. *Neuron*, 32, 927–946.

- Block, N., & Fodor, J. (1972). What psychological states are not. *Philosophical Review*, 81, 159–81.
- Churchland, P. M. (1988). *Matter and consciousness*, 2nd edition. Cambridge, MA: MIT Press.
- Davidson, D. (1970). Mental events. Reprinted in *Essays on actions and events*. Oxford: Clarendon Press.
- Dawkins, M. S. (1985). The scientific basis for assessing suffering in animals. In P. Singer (Ed.), *In defense of animals* (pp. 27–40). New York: Basil Blackwell.
- Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness*. New York: Basic Books.
- Descartes, R. (1968). *Discourse on method and the meditations* (trans. F. E. Sutcliffe). New York: Penguin, 1968.
- Edelman, D. B., Baars, B. J., & Seth, A. K. (2005). Identifying hallmarks of consciousness in non-mammalian species, *Consciousness and Cognition*, 14, 169–187.
- Hardcastle, V. G. (1999). *The myth of pain*. Cambridge, MA: MIT Press.
- Kim, J. (2005). *Physicalism, or something near enough*. Princeton, NJ: Princeton University Press.
- Kupers, R., Faymonville, M.E., & Laureys, S. (2005). The cognitive modulation of pain: Hypnosis- and placebo-induced analgesia. *Progress in Brain Research*, 150, 251–269.
- LaGraize, S. C., LaBuda, C. J., Rutledge, M. A., Jackson, R. L., & Fuchs, P. N. (2004). Differential effect of anterior cingulate cortex lesion on mechanical hyperalgesia and escape/avoidance behavior in an animal model of neuropathic pain. *Experimental Neurology*, 188, 139–148.
- Levy, N. (2008). *Cognition and Consciousness* (in press).
- Lewis, D. (1980). Mad pain and Martian pain. In N. Block (Ed.), *Readings in the philosophy of psychology* (Vol. 1, pp. 216–222). Cambridge, MA: Harvard University Press.
- Lidstone, S.C., & Stoessl, A.J. (2007). Understanding the placebo effect: Contributions from neuroimaging. *Molecular Imaging & Biology*, 9, 176–185.
- Locke, J. (1959). *Essay concerning human understanding* (A.C. Fraser, Ed.). New York: Dover.
- Logothetis, N.K., & Wandell, B.A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, 66, 735–769.
- Mill, J. S. (1979). Chapter XII: The psychological theory of the belief in matter, how far applicable to mind. In J. M. Robson (Ed.), *The collected works of John Stuart Mill, Volume IX—An examination of William Hamilton's philosophy and of the principal philosophical questions discussed in his writings*. (Introduction by A. Ryan). Toronto: University of Toronto Press. London: Routledge and Kegan Paul. Retrieved December 13, 2007, from <http://oll.libertyfund.org/title/240/40871>.

- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435–450.
- Porro, C. A. (2003). Functional imaging and pain: behavior, perception, and modulation. *Neuroscientist*, 9, 354–969.
- Ryle, G. (1949). *The concept of mind*. Chicago: University of Chicago Press.
- Shriver, A. (2006). Minding mammals. *Philosophical Psychology*, 19, 443–442.
- Singer, P. (1990). *Animal liberation*, revised edition. New York: Random House.
- Singer, P. (1999). *Practical ethics*, 2nd edition. Cambridge, UK: Cambridge University Press.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157–1162.
- Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68, 141–156.