

The First Neuroethics Meeting: Then and Now

Essays by Jonathan D. Moreno, Patricia Smith Churchland, and
Kenneth F. Schaffner

Editor's Note: It wasn't until 2002 that more than 150 neuroscientists, bioethicists, doctors of psychiatry and psychology, philosophers, and professors of law and public policy came together to chart the boundaries, define the issues, and raise some of the ethical implications tied to advances in brain research. On the 15th anniversary of the Neuroethics: Mapping the Field conference in San Francisco, we asked three of the original speakers to reflect on how far the neuroethics field has come in 15 years—and where the field may be going in the next 15.

The Brains Behind Morality

By Patricia Smith Churchland, B.Ph.

Have we learned more about the brain basis for moral behavior in the years since that memorable neuroethics meeting in 2002? Yes, and remarkably so. Two branches of brain research that are especially relevant to this understanding have blossomed. First, social neuroscience is revealing the underpinnings of social bonding and why we trust and care about family and friends. This is the fundamental platform for morality. Second, research on the wiring that supports reinforcement learning is revealing how we acquire norms and values, along with the powerful feelings that accompany them. These are the social behaviors that take shape on the platform. What remains poorly understood is social problem solving—how social norms emerge or are modified in response to ecological and other pressures. These are the social institutions that give direction and predictability in a culture.

Linking Social Bonding and Morality

Moral values originate, not in the gods we invent, not in some magical “pure reason” detached from all feelings, but in the neurobiology of attachment. Bonding is powerful in mammals and birds, where the young are born immature and helpless. During their evolution, the brains of mammals and birds were rewired to ensure that the mother, and in some species also the father, bonded with and cared for their infants until they were independent. Caring for others, more generally, is a disposition that is expressed in a host of behavioral ways—provisioning, defending, cuddling, grooming, and being together. As a species adapts to its environment, these “ways” become established as values and norms.

Oxytocin is essential in the process of bonding. The mother mammal, whose brain is awash in oxytocin at the birth of her offspring, feels pain when separated from them, and relief when the babies are together with her. As babies are suckled and cuddled, their brains are awash in oxytocin, and they bond too, feeling distress when separated from the mother, and pleasure when close. Thus, stress hormones come into play as well. It is as though the ambit of self-care expands to embrace those to whom we are bonded. This seems to be the roots of all social glue in mammals and birds.

The suite of neurochemicals involved in bonding is not limited to oxytocin and stress hormones, but includes also the endogenous opioids and cannabinoids, which are associated with feelings of pleasure. Then there are galanin and vasopressin, which regulate aggression; and dopamine, the lynchpin of the reward system, which is essential for learning about one's social world. It is through the interplay of these compounds that we discover what is acceptable and what is not, which is nasty and who is nice, and how to get on in the social world. Anatomically, the ancient structures of the reward system, such as the hypothalamus and basal ganglia, are tightly linked to the neocortex, that amazing structure that gives mammals a level of flexibility not seen in social insects whose behavior is under tight genetic control.

Oxytocin research suggests that its roles are complex, not only in guiding attention to socially relevant goings-on and, probably, orchestrating the feelings accompanying bonding, but also in lowering levels of anxiety. In a first approximation of the process, as oxytocin levels rise, stress hormones fall; and with the increase in oxytocin levels come feelings associated with safety, friendliness, and sociality. This means that when I feel you are a friend, my stress hormones decrease and I can relax—I trust you and need not feel the vigilant anxiety that you may hurt me. We can cooperate; we can count on each other. Mammals that are bonded to each other are apt to console one another, want to hang out together, and perhaps share food and defend each other. The hypothesis, briefly, is that the platform for morality is attachment and bonding, which in some species, extends to mates, kin, friends, and possibly strangers. In part I am drawing on a metaphor from computer engineering, where the platform--a computer's architecture and operating system--allows the software to run. Here, the neurobiology of attachment and bonding provides a motivational and emotional substructure that allows the scaffolding of social practices, moral inhibitions, and obligations to find expression. If mammals did not feel the powerful need to belong and be included, did not care about the well-being of offspring, kin and kith, then moral responsibility and moral concern could not take hold.

We can cooperate; we can count on each other. Mammals that are bonded to each other are apt to console one another, want to hang out together, and perhaps share food and defend each other.

The hypothesis, briefly, is that the platform for morality is attachment and bonding, which in some species, extends to mates, kin, friends, and possibly strangers. In part I am drawing on a metaphor from computer engineering, where the platform--a computer's architecture and operating system--allows the software to run. Here, the neurobiology of attachment and bonding provides a motivational and emotional substructure that allows the scaffolding of social practices, moral inhibitions, and obligations to find expression. If mammals did not feel the powerful need to belong and be included, did not care about the well-being of offspring, kin and kith, then moral responsibility and moral concern could not take hold.

Learning Norms and Values

The brains of all mammals and birds are powerfully organized to construct and model their physical and social environments through learning. Learning takes off with a vengeance at birth, especially, but not only, in the cortex. Approval is highly rewarding, disapproval is the opposite. Here, too, evolutionarily ancient subcortical structures are crucial: the hippocampus and the basal ganglia provide the platform for reward learning. Links to the frontal cortex add sophistication and complexity in self-control and in sizing up a social complex situation, in planning and in evaluating consequences.

Is Morality Hard-Wired?

"Hard-wired" is an expression that makes me think of my toaster. It is "hard-wired" to toast bread and bagels. Period. But the mammalian nervous system is remarkable in that its genes turn off and on owing to environmental interactions, enabling changes in the very structure of its wiring. Brains adapt to local circumstances, in both the physical and the social domains.

Had I been born 250,000 years ago, I would have learned, as I grew up, to fit into a local social and physical environment very different from what we face today. With my acquisition of social knowledge managed by my basal ganglia and its reinforcement wiring, I might have found very hairy, smelly men exceptionally attractive, for example. (This does not mean that human brains are

infinitely malleable—I can never learn to play hockey like Wayne Gretzky or sense odors like my dog Duff.) We might say my brain is “soft-wired” for sociality. I was born with a capacity to imitate, to want to be with others, to dislike shunning and disapproval. In a typical social environment of family and clan, those dispositions flourish and become very strong. In an abusive environment, they are apt to develop atypically.

Different cultures do have different practices regarding many aspects of social life, such as when it is wrong to stare or to laugh, when it is a social misstep to offer help, or the conditions under which one must sacrifice family needs for the needs of others. Such pluralism regarding moral practices is an important aspect of what we cope with as humans living together in a highly interconnected world.

Plasticity and Moral Practices and Values

Variability is, of course, a hallmark of biology, and there is certainly variability among humans in their predispositions to sociality, common sense, and temperament. Contrast the loner gold prospector, whose level of sociability extends no further than his old mutt, with the kindergarten teacher who is intensely social and loves to be in the middle of her loving brood. (Such differences in attachment needs notwithstanding, the old prospector might like to come to a pub once a year and the teacher might appreciate the quiet of her solitary cup of tea at the end of the day.)

Exposure from birth to close social interactions means that that the brain is shaped by the world right off the bat. When those worlds differ between individuals—as between that of an Inuit living in the Arctic and an Aztec living in Mesoamerica, there will be corresponding differences in what brains learn. But despite all that variability, it still seems likely that the basic urge to be with others—needed in all mammalian species for parents to tend infants and for infants to want to stay close to parents—is deep and strongly conserved. It is the platform for complex forms of sociality that we find pleasurable as well as profitable.

This means that we are connected to one another in a very basic way. It has long been observed that various social skills we acquire as we grow up are typically “transportable” to other families, other clans, other cultures, with perhaps a bit of tweaking. The brains of strangers can quickly

synchronize owing to common networks and neurohormones, suggesting a powerful connecting thread that links us all together.

But certain differences persist: even when people agree on facts, they may rationally disagree on the relative importance of certain goals— about the value of freedom versus prosperity, for example, or of commitment to family versus dedication to all. Even if, as I suggest, the deepest level of value—the social platform—is shared, we may not agree on the relative value of higher-level practices, because learned values supporting those practices may differ. As I see it, these stubborn moral dilemmas are best addressed in a highly traditional way—through discussion and conversation, with respect and understanding, and by knowing as much as possible about the relevant history of different practices. My feeling is that knowing more about neuroscience, for example, is not going to solve these kinds of evaluative differences, whereas institutions that incentivize mutual understanding may at least encourage peaceful progress.

Bio

Patricia Smith Churchland, B.Ph., is a professor emerita of Philosophy at the University of California, San Diego, and an adjunct professor at the Salk Institute. Her research focuses on the interface between neuroscience and philosophy. She is author of the pioneering book, *Neurophilosophy*, and co-author with T. J. Sejnowski of *The Computational Brain*. Her current work focuses on morality and the social brain; *Braintrust: What Neuroscience tells us about Morality*. *Touching a Nerve* portrays how to get comfortable with this fact: I am what I am because my brain is as it is. She has been president of the American Philosophical Association and the Society for Philosophy and Psychology, and won a MacArthur Prize in 1991, the Rossi Prize for neuroscience in 2008, and the Prose Prize for science for the book, *Braintrust*. She was chair of the philosophy department at the University of California San Diego from 2000-2007. Churchland received a Bachelor of Philosophy from Oxford University, a Master of Arts degree from the University of Pittsburgh, and a Bachelor of Arts from the University of British Columbia.

Past and Present Considerations

By Kenneth F. Schaffner, M.D., Ph.D.

The field of neuroethics has flourished in the 15 years since the 2002 landmark conference, “Neuroethics: Mapping the Field,” and its subsequent publication—an expansion that includes the founding of the International Neuroethics Society, two journals (*Neuroethics* and *AJOB Neuroscience*), several related websites, and a number of significant books and collections as well as a recent federal research grants program from the “BRAIN Initiative: Research on the Ethical Implications of Advancements in Neurotechnology and Brain Science.” This vibrant evolution has been nourished by the extraordinary development of neuroscience and neuroimaging and by the hopes and concerns generated by the prospect of new brain interventions.

At the 2002 conference, I proposed several distinctions related to the neuroethics questions of reductionism and free will and, in the subsequent 15 years, have further explored these issues. But more needs to be said and speculated on involving scientific and philosophical developments in the past 15 years—and for the next 15 as well.

We can begin with the “standard model,” the “split-level view” or the “hierarchical model,” all of which are partly dependent on a development of the well-known theory of Harry Frankfurt and Gerald Dworkin—two celebrated American philosophy professors—on free will and autonomy.

Frankfurt’s view notes that humans have the capacity to entertain “second-order desires,” which is analogous to an actor deliberating about choice and decision. When an individual acts, so that both first- and second-order desires agree, that individual exhibits “free action.” Another way of looking at that agreement is that the action reflects the individual’s “true self,” and that the second-order desire is one in which the individual identifies and is happy to own.

Frankfurt’s account has generated extensive criticism, but an elaboration of his approach has become a sort of “default position” in neuroethics, with the agreement of first and second order desires being supplemented with requirements of rationality and information sufficiency, as well as

the absence of external pressures or excusing conditions such as mental illness. In the neuroethics literature, such action involving free will is frequently termed “agency.”

Still, more examination is required, and this free-will perspective suggests that a deep account of a person’s “self” might offer an approach to the intertwined notions of choice, freedom, control, and the fundamental value of autonomy. In my view, one might consider *self* in terms of what is stable over the long term—what “grounds” a person and what he or she is like. Their “self-identity” or “personal identity” or “who we are” is constituted by a temporarily evolving but roughly stable continuous narrative. From extensive psychological research as well as human genetics investigations, we may well best obtain a reasonable understanding, or at least a first attempt at such, by examining the nature of an individual’s personality. The self, however, surely encompasses many other facets, a view developed originally by the 19th century American philosopher and psychologist William James, and modernized by Ulric Neisser, a German-born American psychologist and author of *Cognitive Philosophy* in 1967.

Ethics and Neuromodulation

In recent years, neuroethics concerns in the areas of choice, personality, and the self have been complicated by the development of brain neuromodulation techniques. Closely related issues involve personality alterations and the limitations of choice posed by such conditions as addiction and psychopathy, which may adversely affect both the individual and society and become entangled with legal questions. It is in the complex area related to neuromodulation that I envision significant neuroethics advances and challenges emerging.

Three neuromodulation interventions thus far have been shown to affect motor diseases, psychiatric disorders, executive processes, and personality traits. All are currently under investigation in research contexts and two are used clinically. Deep brain stimulation (DBS), in which electrodes are implanted deep in the brain and connected to a permanently implanted electrical source, is invasive, but also the most effective. Externally applied magnetic and electrical interventions—repetitive transcranial magnetic stimulation (rTMS) and transcranial direct current stimulation (tDCS)—are noninvasive. The latter techniques have had less consistent results than DBS but offer promise, and one (rTMS) has been approved for treatment of depression.

Questions surrounding DBS, which has proven quite useful in the treatment of Parkinson's disease, depression, and obsessive-compulsive disorder (OCD), have arisen with the appearance of marked and unusual side effects. One of the most striking examples in the literature involved a 60-year old OCD patient who developed a sudden and distinct musical preference for Johnny Cash following DBS, but who reverted to his usual, more eclectic tastes when not under neuromodulation. A more distressing case involved a 62-year old patient with severe Parkinson's disease that was alleviated with DBS, but who developed severe mania while stimulated, resulting in the ethical dilemma to the patient of choosing to be either bed-ridden or manic. He made the latter choice.

Ethical questions are also raised by the application of DBS, now in research trials, to the treatment of addiction behavior, and by the possibility (currently under pre-clinical investigation) of using it for personality disorders, including psychopathy or its more restrictive DSM-related construct, antisocial personality disorder. This type of intervention has been proposed for potential treatment of incarcerated but willing psychopathic individuals.

(Though it is not known exactly how DBS works, recent advances in optogenetics—a precision technique developed by Stanford University professor Karl Deisseroth in 2005, in which light-responsive proteins allow specific neurons to be turned on or off—may resolve these questions.)

DBS, moreover, can produce significant neuroanatomical circuit remodeling (as recently shown in a mouse model of depression), raising troublesome questions concerning reversibility. To resolve such questions, much more will need to be known about brain circuits and how their variations affect thought and behavior. Advances in this area are likely to come from the National Institute of Mental Health (NIMH) [RDoC initiative](#), already under way.

For such reasons, neuromodulation interventions have already generated concerns and policy proposals in neuroethics, and will continue to do so with increasing urgency. Some philosophical and psychological developments offer assistance. One argument states that psychological and neuroscientific findings amplify our understanding of choice behavior but undercut the “standard model” cited above, showing it to be illusory. Others, however, see that something like this

standard model can be retained, albeit with further refinement. Dartmouth College philosophy professor Adina L. Roskies has rightly suggested that “agency” in a DBS context—and of the type approximated in the standard model—needs much more multidimensional development to reflect its true complexities.

Recommended Philosophers

With these approaches in mind, neuroethicists should examine the extensive philosophical literature on the “self,” such as found in books edited by Shaun Gallagher, and especially in philosophical analyses of self-identity. These inquiries need to be coupled with neuroscience advances as pioneered by neuroscientist-philosophers like Antonio Damasio and Georg Northoff, in a combined top-down and bottom-up pincer strategy. Neurobiological uncertainty about the self—in particular, its anatomical location—makes this an especially challenging task. Commenting on localization, Gallagher has written that “there seems to be overwhelming evidence that the self is everywhere and nowhere in the brain.”

Further, I believe they should incorporate the burgeoning [psychological literature](#) on the “true self” which is a kind of conscience or super-ego. This notion overlaps with the concept of the cluster of second-order desires in the standard model, and shares with that concept features both of subjectivity, normativity, and complexities of verifiability. Including a “true self” in this mix, however, with subjective and necessarily person-variable features, will be controversial. And perhaps, along with the notion of self-identity, even constitute a *diachronic* version of the much-discussed but still unsolved “hard problem” of what consciousness is, as initially formulated by philosopher David Chalmers.

This possibility that we will confront the “hard problem” of consciousness in neuroethics in this regard seems very likely, since first-order desires are perceived as first-person experiences, and second-order desires are similar, even if intertwined with memories and critical executive evaluations. Given its centrality in the issues of choice and free will and the standard model, and though difficult and understudied, including consideration of a “true self” in these analyses seems worthy of sustained exploration. It is these types of investigations into a deep account of a person’s “self” that I believe will advance this significant set of interrelated neuroethical issues.

Down the Line

This, then, is my expectation: neuroethics will increasingly be concerned with ethical, psychological, psychiatric, and legal issues in the intertwined areas of choice, personality, the self, and consciousness as affected by continuing developments in brain neuromodulation techniques. Refinement of these techniques, especially DBS, is likely to come from integrated optogenetic capacities that will provide better knowledge, specificity, and control. More generally, advances in the understanding of brain circuits and higher-level functions including consciousness, arising out of research such as the NIMH's RDoC initiative, will foster progress in the field.

References to specific citations supporting most of the neuroethics advances mentioned are available from the author by writing him at kfs@pitt.edu.

Bio

Kenneth F. Schaffner, M.D., Ph.D., is Distinguished University Professor of History and Philosophy of Science emeritus at the University of Pittsburgh. He taught previously at the University of Chicago (1965-1972) and The George Washington University (1991-2005), where he is University Professor of Medical Humanities emeritus. At Pitt he served as co-director of the Center for Medical Ethics (1986-1991), and was editor-in-chief of *Philosophy of Science* (1975-1980). His most recent book is *Behaving: What's Genetic and What's Not and Why Should We Care?* (Oxford University Press, 2016), and he is currently at work on a sequel: *Choosing: What Can We Learn about Choice and Flourishing from Behavioral Neurogenetics?* Schaffner holds a Ph.D. in philosophy from Columbia University and an M.D. from the University of Pittsburgh.