

The Impact of Neuroscience on Philosophy

Patricia Smith Churchland^{1,*}

¹Philosophy Department, University of California, San Diego, La Jolla, CA 92093, USA

*Correspondence: pschurchland@ucsd.edu

DOI 10.1016/j.neuron.2008.10.023

In the last two decades, neuroscience has profoundly transformed how we understand learning, decision making, self, and social attachment. Consequently, traditional philosophical questions about mind and morality have been steered in new directions.

Philosophy, in its traditional guise, addresses questions where experimental science has not yet nailed down plausible explanatory theories. Thus, the ancient Greeks pondered the nature of life, the sun, and tides, but also how we learn and make decisions. The history of science can be seen as a gradual process whereby speculative philosophy cedes intellectual space to increasingly well-grounded experimental disciplines—first astronomy, but followed by physics, chemistry, geology, biology, archaeology, and more recently, ethology, psychology, and neuroscience. Science now encompasses plausible theories in many domains, including large-scale theories about the cosmos, life, matter, and energy. The mind's turn has now come.

The classical "mind" questions center on free will, the self, consciousness, how thoughts can have meaning and "aboutness," and how we learn and use knowledge. All these matters interlace with questions about morality: where values come from, the roles of reason and emotion in choice, and the wherefore of responsibility and punishment.

The vintage mind/body problem is a legacy of Descartes: if the mind is a completely nonphysical substance, as he thought, how can it interact causally with the physical brain? Since the weight of evidence indicates that mental processes actually are processes of the brain, Descartes' problem has disappeared. The classical mind/body problem has been replaced with a range of questions: what brain mechanisms explain learning, decision making, self-deception, and so on. The replacement for "the mind-body problem" is not a single problem; it is the vast research program of cognitive neuroscience.

The dominant methodology of philosophy of mind and morals in the twentieth

century was conceptual analysis. Pilloried by philosophers of science as *know-nothing philosophy*, conceptual analysis starts with what introspection reveals about the allegedly unassailable truths of folk psychology. Then, via reflection and maybe a thought experiment, you figure out what must be true about the mind.

A frankly a priori strategy, conceptual analysis ran up against a torrent of neuropsychological results that clashed with the "truths" of folk intuition. Among the surprises were patients with split brains or blindsight or hemineglect or alien hand. Their deficits and residual capacities confounded the designated "conceptual truths." Because the data are the data, in place of these alleged "truths" arose empirical questions about brain mechanisms.

In a general way, therefore, the impact of neuroscience and psychology has been profound. Like the world, the mind turns out to be rather different from how it appears to us to be. The Earth seems flat, the moon seems about the size of a small barn, and boils seem to be God's punishment for sin. Intuitions notwithstanding, it is not so. Like folk physics and folk biology, folk psychology embodies much misdirection, despite being moderately serviceable in day-to-day business. Though introspection is useful, the brain is not rigged to directly know much about itself, such as why we are depressed or in love or that factors such as serotonin levels influence our decisions.

Once philosophers appreciated that the seemingly invulnerable truths of intuition were all too vulnerable, conceptual analysis as a method stumbled to its knees. Currently, the most productive philosophers of the mind/brain are steeped in the relevant empirical sciences. Predictably, the style of their work varies: exper-

imental, synthetic or integrative, theoretical or speculative.

Despite advances from the behavioral and brain sciences, moral philosophers in general continued to reassure students that philosophical inquiry into values and moral rules has essentially nothing to learn from brain research. *Moral* philosophy, at least, is safe from neuroscience.

This too, is an illusion. Over the last several decades, research on social behavior has ushered in a naturalistic framework for looking at human morality and decision making. My aim here is to tell the story, about as condensed as the one-minute *Hamlet*, of the impact of neuroscience on our understanding of morality. The story is told against the immensely rich backdrop of results in the biological and social sciences. It begins with the now-legendary research on the neurobiology of mate attachment in voles (Insel and Fernald, 2004; Carter et al., 2008).

Pair-bonding varies across different species of vole: prairie voles mate for life; montane voles display no partner preference. Male prairie voles guard the female and the nest and share parenting of the pups. In montane voles, only females rear the pups. General levels of sociability are also distinct. Placed randomly in a large room, prairie voles tend to cluster in fairly chummy proximity; montane voles are loners. What is the brain basis for these striking differences in sociality?

The main neurobiological contrast is that prairie voles have a much higher density of receptors for the sibling neuropeptides arginine vasopressin (AVP) and oxytocin (OT) in the ventral pallidum and the nucleus accumbens, respectively, than do montane voles (Lim et al., 2004). Although all mammals have both OT and AVP centrally, it is the receptor density in these specific and highly interconnected

regions that marks the crucial difference in behavior.

The profile of receptor density seen in prairie voles extends to other monogamous species—for example, to marmosets and the mouse (*Peromyscus californicus*). By contrast, nonmonogamous species, such as rhesus monkey and the mouse *Peromyscus leucopus*, have an OT and AVP receptor profile similar to that of the nonmonogamous voles. The data for humans are not yet available.

OT is released during positive social interactions and has been shown to inhibit defensive behaviors, such as fighting, fleeing, and freezing. It interacts with the hypothalamic-pituitary-adrenal axis to inhibit activity in the amygdala and to down-regulate autonomic responses originating in the brainstem. But its effects are context sensitive. OT administered to male rats increases aggression to an intruder but decreases aggression toward pups. Less is known about the role of AVP.

Philosophically, these results were alarming. Monogamy seemed to be a complex life choice, requiring rational adherence to a universal rule and conscious self-control. It was commonly argued that one had a moral duty to be monogamous and that this duty is owed to moral deliberation and reason or perhaps to God's commands. The very possibility that pair-bonding in humans might be significantly underpinned, or even modestly affected, by the density of receptors for the simple peptides OT and AVP in brain-specific regions seemed difficult to square with the high-minded requirements of moral duty.

The celebrated caution to acknowledge here is that human sociality is not identical to that of voles or marmosets. Quite so, but like them, most humans do form long-term attachments with mates, offspring, kin, and others, and like them, our reward systems mediate learning local practices. Moreover, evolution is remarkably conservative; brain organization and chemistry is shared across mammals. Consequently, it would not be surprising to find that OT and AVP play a significantly similar role in social attachment in humans. Although much remains to be discovered, available data point in that direction.

A recent Swedish study indicated significant pair-bonding differences between adult human males who carried the so-

called “polygamous” variant of the gene for the AVP receptor and those who did not (Walum et al., 2008). Manipulations of OT have also produced significant results. Using a nasal spray, Kosfeld et al. (2005) administered OT to human subjects before they began playing Investor, a neuroeconomics game where the degree of trust between the investor and the trustee affects the level of monetary winnings. OT investors showed higher levels of trust than controls.

Studies on prairie voles and on humans have shown that OT is important in development of normal social behavior, including the capacity for later formation of stable bonds with mates and others. Wismer Fries et al. (2005) showed that children raised in orphanages and deprived of normal cuddling as infants had significantly lower levels of OT following interactions with their adoptive mothers than did control children interacting with their mothers.

A diminished capacity to form and maintain trusting bonds with others forestalls the many benefits of cooperation. King-Casas et al. (2008) studied subjects identified as having borderline personality disorder (BPD) as they played the investor game. In the investor role, BPD subjects were poor in maintaining a trusting relationship and poor in signaling trustworthiness to repair a trust rupture, even when given an incentive to do so. As investors, they do less well in the game, and they also self-report lower levels of trust than do normal controls.

Although these sociality data need to be widely known because they bear upon how humans choose, they do not automatically imply that our standards for responsibility must be relaxed. An explanation does not entail an excuse, though it is relevant to our understanding of behavior (Churchland, 2006).

But so what, the moral philosopher may ask. What does social attachment have to do with morality? The hypothesis on offer is that attachment, and its cohort, trust, are the anchors of morality; the reward systems tune up behavioral responses. Social animals, including humans, have a powerful urge to be with those to whom they have become attached. We feel safe in their company and anxious when separated.

These emotions spur the brain to find harmonious solutions to the complexities

of social life. Attachments per se do not specify exactly what action should be performed in what condition. They may be best conceived as dispositions that contour social-problem space. Relative to context, these dispositions might be expressed by grooming a consort, attacking intruders, or nurturing a baby. Come the time when action is required, a range of factors can come into play: perceptions, other emotions such as fear of nearby predators, drives such as hunger, and levels of hormones.

The brain's networks continuously face constraint satisfaction problems, both social and otherwise. In dilemmas, some considerations are not mutually satisfiable; e.g., saving one child versus saving another. Typically, constraints are not measurable against each other; e.g., how do we measure the value of training soldiers to kill against the cost to them of becoming killers? To a first approximation, the constraints will include immediate desires, but also the force of habits, reputations, the expectations of others, and evaluation of relevant options. As the relevant constraints weigh in, the networks settle into a solution—the brain's decision. The exact nature of the process whereby networks settle is a largely unsolved problem in computational neuroscience. But the representation of rules and their applicability to the situation at hand seems to be only one constraint among others. According to my hypothesis, practical reasoning mainly consists in finding a good solution to a constraint satisfaction problem. Deduction—the sentimental favorite of logicians—plays at most a minor or post hoc role (Churchland, 2008).

Despite the neuroendocrine and wiring similarities between humans and other social animals, it may be argued that *only humans have genuine morality*. One reason given is that human morality extends to all humans, in a way in which chimp morality does not extend to all chimps.

Whether human morality is really as universal or as exalted as this argument presumes is controversial, owing to the history of tribal and national warfare and common out-group hostility (Wrangham and Peterson, 1996). It is worth noting that the idea that human rights apply equally to all humans, though laudable by our standards, appears to be a fairly recent invention (Hunt, 2007).

Setting aside the issue of historical fact, it is true that human groups can be large and that kindly behavior may extend beyond the circle of kin and even beyond the community. Traditional moral philosophers are apt to attribute this phenomenon to a unique relationship with God, to the greater intrinsic goodness of humans, to our greater intelligence, or to some combination of these. Though these may be implicated, it is worth considering that biologically rooted dispositions explain extending social attachment beyond kin and clan.

Bowles (2006) has argued that altruism and lethal competition between human groups coevolved. Just as a chimp troop is apt to expand its territory and resources by killing off members of a neighboring troop, early hominins probably found it paid to raid weaker hominin clans and divide the spoils in a sufficiently fair-ish way to ensure loyalty. Able manpower to defend and attack would be an important consideration in enlarging the group and extending attachments.

Even so, amalgamation is a risky business, since problematic newcomers could undermine the welfare or stability of the group. Will they be a social boon or burden? Before accepting a newcomer, the group needs assurance that he can bond normally and is not socially or emotionally handicapped. The hypothesis is that, as a first-pass filter for trustworthiness, *unconscious mimicry* serves rather well.

Psychological studies on unconscious mimicry in humans show that the posture, mannerisms, prosody, and words of the experimenter are unknowingly mimicked by the experimental subject as the two engage on a shared task. Additionally, subjects whom the experimenter mimics tend to evaluate the experimenter more favorably than if they were not mimicked (Chartrand and Dalton, 2008). Subjects who experience social stress before beginning the task display a higher level of unconscious mimicry than otherwise. Casual observation of humans getting to know each other supports the science, indicating that unconscious mimicry functions as “social glue.” The production and detection of mimicry requires energy, implying that the brain cares enough to spend the resources on a regular basis. Why? Is it possible that humans use imitative behavior as evidence of normal social capacities?

Humans appear to be vastly more imitative than other primates (Tomasello et al., 2005). When infants begin to imitate, a deeper level of bonding seems to emerge. *Why does infant imitation bring such joy to parents?* One factor among others is that imitative performance predicts that the child has the neural wherewithal to learn what he needs to learn to survive, both socially and in the wider world. Negatively put, if the infant fails to imitate, the failure is a worrisome predictor that the brain lacks what the infant needs to get on in the social world. In the ancestral condition, parental investment may be reduced accordingly. Mimicry, I suggest, serves as a social signal because it indicates the presence of a crucial social capacity, namely the capacity to “read minds”—know what others intend, believe, expect, and feel. If mimicry can be used to evaluate infants, so also strangers.

The idea is that adults respond positively to mimicry in social situations because imitative behavior is a powerful signal of social competence that inaugurates trust or assures the continuation of trust. If the newcomer is trustworthy, in this sense, he will probably behave in a way that is consistent with good citizenry. This means that mimicry, even if unconsciously produced and unconsciously detected, is a safety signal. The level of OT, and hence the level of trust, probably increase; defensive behavior and autonomic arousal decrease. Mimicry is not a fail-safe predictor of social competence, and full acceptance will be gradual. As a first-pass filter, however, it may weed out the worst. As a first-pass filter, it may also set the stage for trade and cooperation with other clans.

Some strangers with evil intent may pretend so thoroughly that they do unconsciously mimic. Others may not, thus tipping off the insiders that something is amiss. The occasional sociopaths may easily gain entry, though the old hands may read groveling behavior as too good to be true.

If values are rooted in biology and the social emotions, can we just settle social/moral questions by looking at our biology? Can the neurobiology of social behavior give us specific answers, such as whether we ought to have a military draft or legalize cocaine? No indeed, but no one seriously

supposes so anyhow. Solving social problems is an awesomely complex business, requiring relevant facts, including facts about cultural practices, about what brains do value, and fact-based predictions about consequences. Fundamentally, moral/social problems are *constraint-satisfaction problems* at the many-brain level, just as most individual choices are constraint-satisfaction problems at the single-brain level. As Aristotle and Hume well recognized, they are problems where moral fervor or absolute rules often get us into more trouble than calm, collective constraint-satisfying negotiation (Churchland, 2008).

ACKNOWLEDGMENTS

Thanks to A.K. Churchland, M.M. Churchland, P.M. Churchland, and E. McAmis.

REFERENCES

- Bowles, S. (2006). *Science* 314, 1569–1572.
- Carter, C.S., Grippio, A.J., Pournajafi-Nazarloo, H., Ruscio, M.G., and Porges, S.W. (2008). *Prog. Brain Res.* 170, 331–336.
- Chartrand, T.L., and Dalton, A. (2008). In *The Oxford Handbook of Human Action*, E. Morsella, J. Bargh, and P. Gollwitzer, eds. (Cambridge, MA: Oxford University Press).
- Churchland, P.S. (2006). In *Neuroethics*, J. Illes, ed. (Cambridge, MA: Oxford University Press), pp. 3–16.
- Churchland, P.S. (2008). In *Oxford Handbook of Philosophy and Neuroscience*, J. Bickle, ed. (Cambridge, MA: Oxford University Press).
- Hunt, L. (2007). *Inventing Human Rights: A History* (New York: Norton).
- Insel, T.R., and Fernald, R.D. (2004). *Rev. Neurol.* 27, 697–722.
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lorenz, T., Fonagy, P., and Montague, R. (2008). *Science* 321, 806–810.
- Kosfeld, M., Heinrichs, M., Fischbacher, U., and Fehr, E. (2005). *Hum. Nat.* 435, 673–676.
- Lim, M.M., Murphy, A.Z., and Young, L.J. (2004). *J. Comp. Neurol.* 468, 555–570.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). *Behav. Brain Sci.* 28, 675–691.
- Walum, H., Westberg, L., Henningsson, S., Neiderhiser, J.M., Reiss, D., Igl, W., Ganiban, J.M., Spotts, E.L., Pedersen, N.L., Eriksson, E., et al. (2008). *Proc. Natl. Acad. Sci. USA* 105, 14153–14156.
- Wismar Fries, A.B., Ziegler, T.E., Kurian, J.R., Jaccor, S., and Pollak, S.D. (2005). *Proc. Natl. Acad. Sci. USA* 102, 17237–17240.
- Wrangham, R., and Peterson, D. (1996). *Demonic Males* (Boston: Houghton-Mifflin).