7/10/10

To appear in *The Handbook of Social Neuroscience,* edited by Jean Decety & John Cacioppo

Ethical, legal and societal issues in social neuroscience

Martha J. Farah

Contents

I. The emergence of neuroethics and social neuroscience.

"Neuroethics" has become the short-hand term for the wide array of ethical, legal and societal issues raised by progress in neuroscience. Social neuroscience has contributed many of the most interesting and important of these issues because of its direct relevance to our personal lives. Consider the ways in which the emergence of social neuroscience has multiplied the roles that neuroscience can play in society.

The achievements of human neuroscience in the 20th century primarily concerned nonsocial neuroscience. Tremendous progress was made in understanding the perceptual, motor and elementary memory systems of the brain (e.g., Kandel, Schwartz & Jessell, 2000). This progress certainly had the potential to influence human life. For example, it had consequences for the diagnosis and treatment of neurological and psychiatric disorders, and opened the door to the investigation of genetic and pharmacologic manipulations of learning and memory in healthy brains (Marshall, 2004). Nevertheless, much of what we care about as human beings – our personalities, moods, relationships, and our lives outside of the medical context – were relatively untouched by these advances. In contrast, neuroscience that addresses the mechanisms of social processes, including person perception and memory, personality, motivation, moral judgment, attachments and biases, presents many more opportunities to affect human life.

Presumably for these reasons, neuroethics has emerged as a field more or less in tandem with social neuroscience. The first meetings on neuroethics were held at the University of Pennsylvania and Stanford University in February and May of 2002, just following the first meeting on "Social Cognitive Neuroscience" at UCLA in April of 2001. The first journals devoted to neuroethics, *American Journal of Bioethics – Neuroscience* and *Neuroethics*, were established in 2007 and 2008, respectively, just following the inaugural issues of *Social Neuroscience* and *Social, Cognitive and Affective Neuroscience* in 2006.

This chapter reviews the neuroethical issues that are most closely related to social neuroscience. The next two sections will address the issues that emerge from neuroscience-

based technologies, in other words relatively pragmatic issues concerning how the fruits of social neuroscience can and should be applied. These include ethical, legal and social challenges raised by newfound abilities to image the brain and thereby obtain information about mental states and personal traits, as well as by our growing ability to intervene in individuals' brain function to alter these states and traits. The final section will address neuroethical issues that emerge from the impact of social neuroscience on our understanding of human beings. In this section it is the knowledge per se, not its technological applications, that is the focus of the review. This section will include the ways in which our evolving understanding of the human person challenges our long-held beliefs about morality and spirituality.

II. Neuroimaging of socially relevant traits and states

Discussions of imaging in neuroethics often seem to suffer from split personality. On the one hand, one finds claims of imminent mind-reading and the end of mental privacy, especially in the popular press. On the other, functional neuroimaging is frequently denigrated as misleading, oversold, and incapable of delivering true information about an individual's psychology. Not surprisingly, the truth about neuroimaging falls somewhere in between. Challenges to mental privacy and the dangers of uncritical acceptance are both important neuroethical issues. We begin by reviewing the state of the art in imaging mental traits and mental states.

"Images of mind:" from cognitive to social neuroscience and from groups to individuals. The now ubiquitous use of functional brain imaging in human neuroscience began in the late 1980's with PET studies of language and memory (e.g., Petersen, Fox, Posner, Mintun & Raichle, 1989; Squire, Ojemann, Miezin, Petersen, Videen & Raichle, 1992). As described in the book "Images of Mind" (Posner & Raichle, 1994), within ten years PET was largely displaced by the more widely available and noninvasive fMRI, which was quickly applied to a wide range of cognitive processes. By the turn of the 20th century this work had extended to include affective and social processes. Landmark studies of such topics as theory of mind (see Stone, this volume),

empathy (see Singer & Decety, this volume), unconscious attitudes (see Cunningham, this volume) and emotion regulation (see Ochsner, this volume) ushered in the era of social neuroscience.

The ethical, legal and social implications of this research did not become significant until the focus shifted from groups of subjects, studied in order to generalize about typical brain function, to individual subjects, studied to understand individual variation. Typical of the earlier studies were findings that localized a mental process, for example the experience of empathic pain, by comparing the brain activation of a group of subjects in an empathic pain condition with the brain activation of the same subjects in a control condition (see Singer & Decety, this volume). Prospects for applying this research to the measurement or classification of individuals are, of course, inherently limited. However, within a few years the same researchers had begun to relate degrees of brain activation differences or effects in these experimental paradigms to individual differences in personal traits of potentially great interest to society. Although the statistical basis for some of these analyses has been criticized (Vul et al., 2009), there is no doubt that moderately strong correlations have been documented in many cases between imaging measures and psychological traits.

Imaging psychological traits. A number of socially relevant psychological traits have been found to correlate moderately strongly with brain activity in fMRI studies, including aggression, altruism and racial attitudes. For example, an early and influential study by Liz Phelps and collaborators found that white subjects' amygdala activation correlates with the degree of unconscious negative evaluation of black faces (Phelps et al., 2000). Specifically, the discrepancy between amygdala activation to black and white faces correlated r=0.576 with the magnitude of unconscious bias against blacks measured in the Implicit Association Test (Greenwald, McGhee & Schwartz, 1998).

Coccaro et al. (2007) showed subjects with and without a history of impulsive aggression photos of faces displaying different emotions, while measuring neural responses to these photos with fMRI. In addition to finding overall differences between aggressive and

nonaggressive subjects in their response to the sight of an angry face, including greater activation of the amygdala and less activation of the presumably regulatory orbitofrontal cortex, they also found a correlation between amygdala activation and aggression. The more aggressive one's behavior, measured over one's lifetime, the higher the activation of the left amygdala to angry faces, r=0.546.

Turning to a more desirable trait, altruistic cooperation, Rilling et al. (2002) scanned subjects while they played an iterated "Prisoner's Dilemma" game and assessed the relationship between the tendency to prolong mutually cooperative play and the activation of reward-related brain areas by such cooperation. They found a correlation of r=0.70 between cooperative behavior in the scanner and the activation it evoked in the ventral striatum, an area associated with the enjoyment of rewards from money to chocolate (O'Dougherty, 2004).

These studies and others like them involve what Poldrack (2006) has called "forward inference." That is, the direction of inference is from psychological process (manipulated by the researcher experimentally or by selection of subject characteristics) to brain activation (measured by the researcher; see Cacioppo & Tassinary, 1990). Each of these studies was part of a research program aimed at understanding the neural bases of the particular psychological construct under study. This can be contrasted with research whose aim is detecting or measuring personality traits or attitudes, which would be a form of "reverse inference," going from brain activity to psychological trait. Nevertheless, the data collected in these studies are also applicable to the goal of measurement. A recent review of the published literature and statistical reanalysis of published data concluded that a modest degree of measurement ability already resides in the experimental paradigms of social neuroscience research (Farah, Smith, Gawuga, Lindsell & Foster, 2008). To the extent that neuroimaging can be used to measure traits of personal or social significance, it raises issues of privacy, which will be addressed in more detail in the final part of this section.

Imaging mental states. Whereas the studies of trait-related differences in brain activation have generally been undertaken within the context of basic research programs, imaging studies of mental states often have a closer relationship to applied research goals such as the measurement or detection of specific states with real-world relevance. The field of neuromarketing is a prime example of the use of reverse inference for an applied goal.

The emotions and motivations of consumers are crucial for many marketing decisions, from brand identity to pricing, but consumers are notoriously poor at reporting these aspects of their own psychology. The prospect of directly "reading" consumers' brain states is therefore of great interest to marketers. In addition, brain imaging is relatively well-suited to this type of reverse inference. Compared to some psychological states, states of liking and wanting have a relatively straightforward relation to patterns of brain activity. EEG and fMRI have therefore become widely used tools in market research.

Published research in the field of neuromarketing has illuminated the ways in which packaging design, price, brand identity, spokesman celebrity and other marketing factors separate from the product itself affect neural responses to the product, and how accurately those neural responses predict purchasing decisions (for reviews see Hubert & Kenning, 2008; Lee, Broderick & Chamberlain, 2006). The success of neuromarketing as a business tool is harder to assess, but the list of companies paying for neuromarketing suggests that many corporate decision-makers have faith in it. Forbes Magazine reported that this list includes Chevron, Disney, Ebay, Google, Hyundai, Microsoft, Pepsico and Yahoo (Burkitt, 2009).

The techniques of neuromarketing are not limited to selling products and services. They have also been used to study preferences for health behaviors (Langleben et al., 2009) and political candidates (Westen et al., 2006). The firm FKF Applied Research published advice to American presidential candidates for the 2008 election in the New York Times Op Ed pages, based on their fMRI studies (Iacoboni et al., 2007). Their advice received widespread attention in the media and online (Aron et al., 2007; Farah, 2007; see also Iacoboni, 2008, Poldrack, 2008). Less

public attempts to understand voters' reactions to candidates based on measures of brain function have reportedly been carried out at the request of specific political campaigns (Linstrom, 2008).

Another type of mental state that researchers have attempted to read from brain activation is the state of intentional deception. Early studies of deception were aimed at the basic science goal of characterizing the differences in brain activation between lying and truth-telling (e.g. Langleben et al., 2002), and showed that the anterior cingulate cortex (ACC) as well as regions of prefrontal and parietal cortex were more active during lies (see Bles & Haynes, 2008; Christ et al, 2009, for reviews). More recent research on deception with fMRI has been aimed at explicitly at the reverse inference of determining the truth value of statements based on brain activation (e.g., Davatzikos et al., 2005) and at least two companies offer fMRI lie detection services. Among the potential applications cited by them are reduction of "risk in dating," vindication "if your word, reputation or freedom is in dispute," and a substitute for drug screening, resume validation and security background checks" in employment screening (Cephos, No Lie MRI urls).

To date such methods have not been admitted as evidence in a court of law. A different type of brain-based lie detection, based on event-related potentials (ERPs) has been admitted as evidence in the US (Harrington v State of Iowa), and in India. Indeed, in India the method has helped convict at least two defendants of murder (Aggarwal, 2009).

Ethical, legal and societal issues in brain imaging. Concerns about the ethics of brain imaging generally focus either on privacy concerns or concerns about the illusory accuracy and objectivity of brain imaging. To the extent that brain imaging can actually deliver information about a person's mental states or traits, the issue of privacy is important. To the extent that it cannot, but people believe that it can, the issue of public misunderstanding is important.

A number of writers have commented on the potential threat to privacy posed by functional neuroimaging (e.g., Committee on Science and Law, ABCNY, 2005; Hyman, 2004). On the face

of things, brain imaging poses a novel challenge to privacy in that it can in principle deliver information about thoughts, attitudes, beliefs and traits, even when someone offers no behavioral responses. More concretely, and perhaps more importantly, imaging-based psychological investigations lend themselves to stealth uses in ways that more conventional paper-and-pencil or other low-tech methods do not. Both structural and functional brain images can be obtained with consent for one purpose but later analyzed for other purposes. Furthermore, in many studies the stimuli and instructions do not reveal the nature of the psychological information being sought. For example, in two of the studies cited earlier, unconscious racial attitudes and impulsive aggression were both correlated with brain activity evoked by simply viewing pictures of faces (Coccaro et al., 2007; Phelps et al., 2000). Hence in principle it seems possible to obtain information about racial attitudes and aggressive tendencies without subjects' knowledge or consent, by misleading them into thinking the study concerns face perception.

At present the problem of public misunderstanding of neuroimaging is a more immediate challenge than the problem of mental privacy. A number of authors have suggested that laypersons may attribute greater objectivity and certainty to brain images than to other types of information about the human mind (Dumit, 2004; McCabe & Castell, 2007; Racine, Bar-Ilan & Illes, 2005; Roskies, 2008). This may contribute to the premature commercialization of brain imaging for various real-world applications, including lie detection.

Different applications of brain-based lie detection call for different levels of protection for the consumers or citizens involved (Farah et al., submitted). Although laboratory research with the kinds of methods used by these companies documents impressively high levels of accuracy approaching 90% under laboratory conditions, this is not sufficient accuracy for high-stakes decision making in contexts such as employment screening, business disputes, legal cases or national security, the very contexts for which the companies recommend their methods. In addition, the accuracies reported in the fMRI literature are based on laboratory tasks that differ in many ways from the real-world situations in which lie detection is employed. These include

the participation of cooperative subjects (especially important for pattern classification methods that require extensive training of the system; see Haynes & Rees, 2006), the artificiality of the situation (e.g., subjects lie because they are instructed to), and the inconsequential nature of the lies.

Within the field of neuromarketing, demonstrations of good performance predicting simple buying decisions (e.g., Knutson et al., 2007) may not extend to more complex social decision making. Yet, as already mentioned, neuromarketers are venturing into the prediction of preferences in more complex situations such as political campaigns.

The risks of premature adoption of these methods, encouraged by the aura of cutting-edge science that surrounds them, can be grave. Qualified people may be denied job opportunities, guilty people may be exonerated and innocent people may be found guilty. At the same time, the more frightening aspects of science fiction mind-reading should not be allowed to cloud our judgment regarding efforts to develop and validate these methods.

III. Manipulation of socially relevant brain traits and states

Although most developments in psychopharmacology and brain stimulation have been motivated by the need to treat neurological and psychiatric illnesses, some drugs can also be used to improve or change aspects of healthy brain function. The use of psychoactive drugs for manipulating normal brain function has been of interest and concern within neuroethics mainly in connection with neurocognitive enhancement (e.g., see Farah et al., 2004; Greely et al., 2008). What about the enhancement of traits and states that are more directly relevant to social phenomena?

Neuroscientists have succeeded in manipulating of normal levels of mood, personality, empathy, trust, aggression and so forth, although most of this work has yet to be translated into practically useful methods. The history of what could be called social-emotional brain enhancement goes back at least as far as the mid-20th century, with medications such as Miltown and Dexadrine, also known as “mother’s little helpers.” These drugs were explicity

marketed for the enhancement of social and emotional functioning in medical journals and elsewhere. Ads touted them as solutions to the problems of modern life, helping housewives better handle the stresses of their 24/7 responsibilities and helping hen-pecked husbands assert themselves at home and at the office. Yet they proved to be addictive and are now rarely prescribed to enhance social or emotional functioning of healthy individuals.

The introduction of Prozac in the 1980's, followed by a string of other selective serotonin reuptake inhibitors (SSRIs), offered a much needed new treatment option for patients suffering from depression and anxiety disorders, but had wider societal effects as well. Peter Kramer foretold much of current neuroethics in his book Listening to Prozac (19??), in which he discussed the use of Prozac by patients who appreciated the drug's effect on their personalities, puzzled over the relationship between brain and self, and coined the term "cosmetic psychopharmacology."

Any discussion of brain enhancement must address the question of where to draw the line between enhancement and treatment. For cognitive enhancement, the question is usually framed in terms of diagnostic boundaries between everyday distractibility and ADHD, or between normal cognitive aging and dementia. In the case of SSRIs for social-emotional enhancement the question is more complex, partly because there are so many therapeutic uses of SSRIs – including depression, premenstrual dysphoria, general anxiety, social anxiety, obsessive-compulsive disorder – and partly because the relevant diagnostic boundaries appear to have shifted because of the SSRIs themselves. In the case of depression, antidepressant medications before Prozac had more troublesome side effects and were therefore reserved for patients with major depression. The greater tolerability of SSRIs, combined with pharma's energetic marketing to patients and doctors, led to a larger number of less ill patients using these drugs and to a revision of diagnostic categories (Healy, 2004). As the line between pathology and health moves to include more people on the pathological side of the line, uses of medication that would have been considered enhancement become therapy.

Ultimately, whether one labels the use of SSRIs by functional and seemingly well-adjusted people as therapy or enhancement, this definitional issue matters less than the fact that large numbers of people are using them. A recent study reported that antidepressants are now the most widely used class of drugs in the US, with an estimated 10% of the population having received a prescription for them in the year 2005 (Olfson & Marcus, 2009). This is particularly relevant to the neuroethics of social neuroscience because SSRIs subtly alter personality. A recent study in depressed patients found that the SSRI paroxetine affects personality above and beyond its effect on depression (Tang et al., 2009). The most pronounced effect on personality was on the trait of neuroticism, the tendency to experience negative emotions. In studies that have examined the effects of SSRIs in nondepressed subjects, their main effect appears to be the diminution of negative affect or neuroticism (Furlan et al, 2004 Knutson, 1998;). For example, Knutson and colleagues (1998) administered paroxetine or placebo for 4 weeks and assessed the effects of the drug on personality and social behavior. The drug reduced negative affect, particularly hostility, and increased affiliative behaviors. For example, subjects on the drug spoke fewer commands and instead made more suggestions to their partners in a problem-solving exercise. Among the subjects who received the drug, plasma levels correlated with changes in negative affect and social behavior.

In subjects selected for criminal behavior rather than psychiatric diagnosis or lack thereof, SSRIs have demonstrated potential for another socially relevant use: promotion of prosocial and law-abiding behavior. Impulsive violence is associated with abnormalities in seratonergic systems, and SSRIs reliably decrease aggression in individuals prone to violence (Berman, McCloskey, Fanning, Schumacher & Coccaro, 2009; Walsh & Dinan, 2001). SSRIs have been found to decrease repeat offending in sex offenders and are used for this purpose, along with hormonal treatments to decrease sex drive (Bourget & Bradford, 2008).

Love, romance and sexuality in healthy normal people constitute another realm for brain enhancements. Drugs that affect these aspects of life through central nervous system mechanisms have not achieved the success of, for example, Viagra, but more limited successes

have been reported. The drug known as “ecstasy” (MDMA) increases feelings of closeness and interpersonal connection and can be used to enhance relationships, although serious risks accompany its use (Sessa, 2007). Hormone supplementation has been used by low-testosterone men and postmenopausal women to increase libido. A number of new drugs, including the serotonin agonist flibanserin, show promise for improving sexual function in young women suffering from low libido, and are under review for this purpose with the US Food and Drug Administration (Fitzhenry & Sandberg, 2005).

In recent years a wealth of new findings has emerged on the role of the hormones oxytocin and vasopressin in human trust, altruism and bonding. Building on basic research with animals, correlational studies have shown that several aspects of interpersonal behavior are related to levels of these hormones or variations in genes governing their action. For example, mothers with higher levels of plasma oxytocin during pregnancy showed better bonding with their infants by several measures of their behavior and thought patterns (Feldman et al., 2007). Players of an economic game who showed more trust of other players and more trustworthiness themselves had higher levels of circulating oxytocin (Zak, Kurzban & Matzner, 2005). Whereas quality of pair-bonding correlates with oxytocin levels in women, it correlates with vasopressin levels in men (Taylor, Saphire-Bernstein & Seeman, 2009).

More relevant to the topic of this section, intravenous or inhaled doses of these hormones have been shown to alter the same range of behaviors. Intranasal oxytocin improves recognition of facial expressions, consistent with a role in social cognition (Domes et al., 2007). Oxytocin has been shown to engender more trusting and generous strategies in economic games (Kosfeld et al., 2005; Zak, Stanton & Ahmadi, 2007). Furthermore, oxytocin appears to interfere with normal responses to maltreatment. Subjects who are betrayed by a partner in a trust game normally reduce their investment on the next trial, an understandable precaution against a selfish or untrustworthy partner. However, no such reduction of investment occurred for subjects on oxytocin (Baumgartner, et al., 2008).

In addition to providing experimental tests of the hypothesis that oxytocin plays a causal role in the behaviors with which it has been correlated in observational studies, this research has obvious potential for translation into a number of applied domains. It provides a proof of concept that could be used to alter the interpersonal relationships between spouses, parents and children, and business associates. It could also be used in diplomatic, forensic and security contexts. Not surprisingly, a quick search online will turn up numerous companies selling oxytocin, although without evidence that the formulation being offered is effective.

Ethical, legal and societal issues in manipulating the brain

Athough the technologies just reviewed are a far cry from mind control, they do offer a number of new ways to influence the feelings and behaviors of others. We should therefore consider how these technologies could be used, what kinds of uses we would consider socially acceptable, and how we might discourage unacceptable uses.

For example, the results of research on oxytocin suggest that neuroendocrine manipulation could be a profitable, if unethical, business strategy. By increasing trust, generosity and forgiveness in one's opponents it appears possible to influence the outcomes of financial, political or other negotiations. This technology also has the potential to enhance interrogation in law enforcement and national security contexts. Could the benefits ever outweigh the harms in any of these cases? Drugging an unsuspecting business associate for financial advantage seems clearly wrong, but what if we could obtain socially valuable information from an unwilling informant without causing physical or psychological pain?

What about encouraging a successful resolution of difficult negotiations by enhancing feelings of bonding and brotherhood in both parties? With their informed consent? What if we chose to increase our own resistance to the natural persuasion of others by defensively blocking our own hormonal capacity for trust and altruism?

From a consequentialist point of view, sufficiently high benefits to society should tip the balance in favor of oxytocinizing interrogees or opponents in a political conflict, even without

their consent. Yet most of us sense a troubling violation of personhood in these scenarios. It is not just the assault on autonomy inherent in influencing people without their knowledge, but the co-opting of our highest moral emotions for instrumental purposes. After all, part of what makes these emotions so precious, to individuals and society, is precisely that they guide us away from selfishness, from the pursuit of our own selfish ends. The prospect of someone else harnessing them for their own ends is therefore especially repugnant.

Pharmacologic treatment of criminal offenders presents us with another set of tradeoffs between benefits and risks. If SSRIs or hormone treatments can enable offenders to live outside of prison and can protect society against crime, then the "benefit" side of the equation is substantial. However, state-imposed psychopharmacology poses a relatively new kind of risk to offenders' autonomy and privacy, different in kind from the restrictions on autonomy and privacy imposed by incarceration.

Although the most obvious neuroethical issues concern the imposition of brain interventions on others, there are also issues that arise with voluntary self-enhancement. The SSRIs have been the subject of two critiques that focus on the ways in which enhancing mood and personality in oneself can harm oneself and society. The first of these is by Carl Elliott, the best-known critic of treating the angst of normal life with medications such as SSRIs. In a number of thoughtful essays he has written of the value of such angst or alienation in alerting one to the need for a more meaningful life (e.g., Elliott, 2004). Whereas Elliott focuses on the dangers of pacifying oneself rather than experiencing misgivings about the true state of one's life or one's world, Fukuyama (2002) has expressed concern over the possibility that SSRIs could inappropriately raise the self esteem of the user, undermining an important source of motivation. He asks if Caesar and Napoleon would have created their empires had they been able to raise their self esteem simply by popping a pill (Fukuyama, 2002, p. 46)! In both cases, the risks seen by these authors have consequences for the individual user and for society. The user is deprived of an authentic view of reality. By failing to see just how problematic the world is, or just how lacking one's own accomplishments are, we will be less motivated to improve these things for the benefit of all.

IV. Social neuroscience: Enlightenment or nihilism?

Some of the most profound ethical challenges from social neuroscience come not from new technologies, but from new understandings. In place of the folk psychology with which we have traditionally understood ourselves and each other, neuroscience is offering us increasingly detailed physical mechanisms. Personality, empathy, altruism and love have all become subjects of study in neuroscience. Starting with the Enlightenment in the 18th century, the natural sciences have become the dominant way of understanding the world around us. In this century it seems likely that neuroscience will complete this process by naturalizing our understanding of ourselves.

As scientists we of course welcome this progress. Yet this progress will pose challenges. The idea that human beings are no more than physical objects, albeit very complex objects containing powerful computational networks, seems to threaten our most fundamental beliefs about the value of human life and the possibility of human agency. The final type of neuroethical issue to be reviewed in this chapter is the challenge of assimilating neuroscience's increasingly complete physical explanation of human behavior without lapsing into nihilism.

At the root of the neuroethical challenges discussed in this final section is the erosion, by neuroscience, of a fundamental distinction that underlies many of our moral intuitions: the distinction between persons and objects. Advances in basic science are revealing the necessary and sufficient neural processing that underlies our mental life, the aspect of persons that most definitively distinguishes them from objects. Advances in applied science, including the manipulation of mental states and traits, reinforce the view that we are physical objects. As we increasingly manipulate our own and each others' brain functions in order to change abilities, moods and personality traits, we will be living with increasingly frequent reminders of the physical nature of the human person.

Moral responsibility. The person-object distinction is more than just a metaphysical abstraction; it is central to our intuitions about morality. We view persons as having agency

and therefore generally hold them responsible for their actions. Although many people believe that, in principle, human behavior is the physical result of a causally determined chain of biophysical events, we tend to put that aside when making moral judgments. We do not say, "But he had no choice—the laws of physics made him do it!" However, as the neuroscience of decision making and impulse control begins to offer a more detailed and specific account of the physical processes leading to irresponsible or criminal behavior, the amoral deterministic viewpoint will probably gain a stronger hold on our intuitions. Criminal law is largely consistent with our intuitions about human behavior, and as such Greene and Cohen (2004) have argued that progress in neuroscience will eventually force changes to the law. They suggest that the difficulty of assigning moral blame may supplant retributive punishment by incentive-based punishments and therapies.

There is already ample evidence for the idea that physical, neuroscience explanations of negative behavior can change attitudes towards them. Most readers of this chapter will be too young to remember first-hand the stigma carried by mental illness before widespread use of psychiatric drugs. In the course of marketing drugs for mental illnesses, the pharmaceutical industry introduced the public to the idea that these conditions have a neurochemical basis. Organizations such as the National Alliance on Mental Illness (NAMI) have fought against stigma in large part based on the biological bases of these illnesses. As a result, depressed individuals are less likely to face blame for being lazy or unwilling to pull themselves out of it, and children with ADHD are less likely to be viewed as bad boys and girls who refuse to behave. The role of neuroscience in shifting discourse from blame to the need for therapy is nowhere more apparent than in society's view of addiction. In 1997, while he was director of the National Institute of Drug Abuse, Alan Leshner wrote an article entitled "Addiction Is a Brain Disease – and It Matters." In it, he cited neuroscience to argue against the view that "drug addicts are weak or bad people, unwilling to lead moral lives and to control their behavior and gratifications." Of course, the public has yet to fully absolve addicts from responsibility for their behavior, but the disease model of addiction has nevertheless been influential in reducing the amount of blame born by addicts.

Human rights. In addition to viewing persons as having agency and hence responsibility, we also view them as having special moral status and hence certain rights. Whereas we value objects for what they can do—a car because it transports us, a book because it contains information, a painting because it looks beautiful—the value of persons transcends their abilities, knowledge, or attractiveness. Persons have what Kant called "dignity," meaning a special kind of intrinsic value that trumps the value of any use to which they could be put (Kant, 1996). One way of expressing this is to say that persons have rights just because they are persons. However, the categorical distinction between persons and objects is difficult to maintain on the view that human beings are no more than physical objects. In that case, why should it matter what becomes of any of us? Why should the fate of objects containing human brains matter more than the fate of other natural or manmade objects?

Religion and spirituality. One way in which people have traditionally justified the special status of human beings, apart from other objects in the natural world, is in terms of the spiritual. Most religions include the belief that a human person is part physical and part nonphysical, the latter part referred to as a soul or spirit. This accords well with our intuition that each of us has some essence above and beyond the hundred or two pounds of matter that we can see and touch. However, if we extrapolate the ongoing progress of neuroscience towards explaining all aspects of our mental lives in terms of physical processes, there is no logical reason to believe in any immaterial aspect of human beings. Neuroscience may therefore contribute to a shift in religious belief away from dualist conceptions of the human person (Farah & Murphy, 2009).

Self and personal identity. Finally, our intuitive understanding of persons includes the idea that they have a unique essence that persists over time. The changes wrought by normal development and life experience are understood as elaborations on a foundational personal identity that is constant throughout life. Yet this belief does not fit with the idea that a person is just his or her brain. As physical objects, brains can and do change in countless ways in response to injury, disease and drugs and, less commonly but no less realistically, implants,

grafts and other surgical interventions. There is no principled limit to the ways in which a brain can physically change, and thus no immutable core to the neural substrates of a person. How can this fact be squared with the notion of an enduring personal identity or essence?

In sum, neuroscience is calling into question our age-old understanding of the human person. Our traditional ways of thinking about responsibility and blame, human rights, spirituality and personal identity will all undergo change in the process of accommodating neuroscience's view of humanity. Whether the result is nihilism or enlightenment remains to be seen. Folk psychology and dualism do not necessarily make the world more humane. The disease model of drug addiction has not so far given addicts a free pass for their destructive behavior but rather has increased society's commitment to providing therapy to addicts. Nancy Murphey (2010) has suggested that Christianity's focus on saving souls may have detracted from its incentive to improve life for those suffering in the physical world of here-and-now. There is thus reason to hope that the insights of neuroscience we will help make us more, not less, able to understand, appreciate and care for one another.

References Cited

Aron, A. et al. (2007). “Politics and the Brain.” New York Times. New York.

Baumgartner T, Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. Neuron. 2008 May 22;58(4):639-50.

Bles M, Haynes JD. Detecting concealed information using brain-imaging technology. Neurocase. 2008;14(1):82-92.

Burkitt, L. (2009). “Neuromarketing: Companies Use Neuroscience for consumer Insights.” Forbes, Forbes.

Cacioppo JT, Tassinary LG. (1990). Inferring psychological significance from physiological signals. American Psychology. Jan;45(1):16-28.

Coccaro EF, et al. (2007). Amygdala and orbitofrontal reactivity to social threat in individuals with impulsive aggression. Biological Psychiatry. Jul 15;62(2):168-78. Epub 2007 Jan 8.

Christ SE, et al. (2009). The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. Cereb Cortex. Jul;19(7):1557-66. Epub 2008 Nov 2.

Committee on Science and Law. National Bar Association. (2005). Are your thoughts your own? Neuroprivacy and the legal implications of brain imaging. New York.

Davatzikos C, et al. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. Neuroimage. 2005 Nov 15;28(3):663-8. Epub 2005 Oct 5.

Domes, Gregor, et al. Oxytocin Improves “Mind-Reading” in Humans. Biological psychiatry 15 March 2007 Volume 61 issue 6 Pages 731-733.

Dumit, Joseph, *Picturing Personhood: Brain Scans and Biomedical Identity,* Princeton University Press, 2004.

Farah MJ, et al. (2004). Neurocognitive enhancement: what can we do and what should we do? Nat Rev Neurosci. May;5(5):421-5.

Farah, M. (2007). “This is Your Brain on Politics?” Neuroethics and Law Blog. A. Kolber. 2009.

Farah, M. J. and N. Murphy (2009). "Neuroscience and the Soul." Science 323(5918): 1168.

Fitzhenry, D. and L. Sandberg (2005). "Female sexual dysfunction." Nature Reviews Drug Discovery 4(2): 99-100.

Feldman R, Weller A, Zagoory-Sharon O, Levine A. (2007). Evidence for a neuroendocrinological foundation of human affiliation: plasma oxytocin levels across pregnancy and the postpartum period predict mother-infant bonding. Psychol Sci. 2007 Nov;18(11):965-70.\

Furlan PM, et al (2004). SSRIs do not cause affective blunting in healthy elderly volunteers. Am J Geriatr Psychiatry. 2004 May-Jun;12(3):323-30.

Greely H, Sahakian B, Harris J, Kessler RC, Gazzaniga M, Campbell P, Farah MJ. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature* 2008, 456(7223):702-705

Greene J, Cohen J. For the law, neuroscience changes nothing and everything. Philos Trans R Soc Lond B Biol Sci. 2004 Nov 29;359(1451):1775-85.

Haynes, J. D. and G. Reese (2006). "Decoding mental states from brain activity in humans." Nature Reviews Neuroscience 7(7): 523-534.

Healy D., (2004), Let Them Eat Prozac. New York University Press, New York.

Hyman, S. E. (2004). "Introduction: the brain's special status." U 6(4): 9-12.

Iacobani, M. (2007). “This is you brain on politics.” New York Times. New York, November 11, 2007.

Iacoboni, M. (2008). “Iacoboni Responds to Neuropolitics Criticism.” Neuroethics and Law Blog. A. Kolber.

Kandel ER, Schwartz JH, Jessell TM. (2000). *Principles of Neural Science, 4th ed.* McGraw-Hill, New York. ISBN 0-8385-7701-6

Kant, I. (1996). Critique of Pure Reason. Indianapolis, Hackett Publishing Company

Kosfeld, M., M. Heinrichs, et al. (2005). Oxytocin increases trust in humans. Nature 435(7042): 673-676.

Knutson B, et al. (2007). Neural predictors of purchases. Neuron. 2007 Jan 4;53(1):147-56.

Knutson, B., Momenan, R., Rawlings, R. R., Fong, G. W., & Hommer, D. (2001). Negative association of neuroticism with brain volume ratio in healthy humans. Biological Psychiatry, 50, 685-690.

Peter D. Kramer, *Listening to Prozac*. Penguin Books, New York, 1993.

Langleben DD, et al. (2002). Brain activity during simulated deception: an event-related functional magnetic resonance study. Neuroimage. Mar;15(3):727-32.

Lindström, M. (2008). Buy ology : Truth and Lies About Why We Buy. New York, Doubleday.

Marshall I et al (2004). Repeatability of motor and working-memory tasks in healthy older volunteers: assessment at functional MR imaging. Radiology. 2004 Dec;233(3):868-77. Epub 2004 Oct 21.

McCabe, D. P., & Castel., A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*, 343-352

Olfson Mark, and Steven Marcus. (2009) National Patterns in Antidepressant Medication Treatment. *Arch Gen Psychiatry.* 2009;66(8):848-856

Phelps EA, et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. Journal Cognitive Neuroscience. Sep;12(5):729-38.

Poldrack, R. (2008). "Poldrack Replies to Iacoboni Neuropolitics Discussion." Neuroethics and Law Blog. A. Kolber. 2010.

Poldrack RA.(2008). The role of fMRI in cognitive neuroscience: where do we stand? Current Opinion Neurobiology. Apr;18(2):223-7. Epub 2008 Aug 7.

Posner, Michael I.; Raichle, ME (1994). *Images of Mind*. Scientific American Books.

Posner MI, Petersen SE (1990). The attention system of the human brain. Annual Review Neuroscience; 13:25-42.

Rilling. J. A Neural Basis for Social Cooperation. *Neuron*, Volume 35, Issue 2, Pages 395-405

Racine E, Bar-Ilan O, Illes J. fMRI in the public eye. Nature Reviews Neuroscience. 2005 Feb;6(2):159-64.

Roskies A. Neuroscientific challenges to free will and responsibility. Trends in Cognitive Science. 2006 Sep;10(9):419-23. Epub 2006 Aug 8.

Sessa B. (2007). Is there a case for MDMA-assisted psychotherapy in the UK? Journal of Psychopharmacology. 2007 Mar;21(2):220-4.

Tang, Tony Z., et al. (2009). Personality Change During Depression Treatment: A Placebo-Controlled Trial, *Arch Gen Psychiatry.* 66(12):1322-1330.

Taylor, S. E., Saphire-Bernstein, S., & Seeman, T. E. (2010). Are plasma oxytocin in women and plasma vasopressin in men biomarkers of distressed pair-bond relationships? *Psychological Science, 21*, 3-7.

Vul, Edward, et al (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognitio. Perspectives on Psychological Science 4.3. 274-290 DOI: 10.1111/j.1745-6924.2009.01125.x

Zak PJ, Kurzban R, Matzner WT. (2005). Oxytocin is associated with human trustworthiness. Horm Behav. Dec;48(5):522-7. Epub 2005 Aug 18.

Zak, P. J., A. A. Stanton, et al. (2007). Oxytocin increases generosity in humans. PLoS One 2(11): e1128.